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

Alison M. Etheridge[*]   Amandine Véber[†]   Feng Yu[‡]

Abstract

We consider the spatial Λ-Fleming-Viot process model 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 first prove that the model is well-defined and provide a measure-valued dual process encoding the locations of the "potential ancestors" of a sample taken from such a population, in the same spirit as the dual process for the SLFV without natural selection [7]. We then 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 dimension. 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 consider the process of "potential ancestors" of a sample of individuals taken from these populations, which we see as (the empirical distribution 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 their collision local time. 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.

[*]Department of Statistics, University of Oxford, UK. E-mail: etheridg@stats.ox.ac.uk
[†]CMAP, CNRS, Institut Polytechnique de Paris, France. E-mail: amandine.veber@polytechnique.edu
[‡]School of Mathematics, University of Bristol, UK. E-mail: Feng.Yu@bristol.ac.uk

The spatial Lambda-Fleming-Viot process with selection

Contents

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 [23] and dubbed "the pain in the torus". The *spatial Λ-Fleming-Viot process* (SLFV), introduced in [7, 15], 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 [6, 17, 18, 19, 25], 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 [24, 32] 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 [18, 19] which address the interaction of natural selection and genetic drift when local population densities are small.

The Fisher-KPP equation

$$\partial_t p = \frac{\sigma^2}{2}\Delta p + sp(1-p) \tag{1.1}$$

was introduced independently by Fisher [24], specifically to model the spread of an advantageous gene through a spatially distributed population, and Kolomogorov, Petrovsky & Piskunov [32], 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 has been on 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 (1.1) 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 [14] 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 (1.1), 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 [31], and studied in more generality in, for

example, [47]. The (continuous space) stochastic Fisher-KPP equation can be obtained from the discrete space counterpart through rescaling (*cf.* [5], 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 [39]. 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.* [13, 37, 38] and a huge body of closely related work inspired by work of Brunet, Derrida and coworkers, *e.g.* [4]. 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. In particular, the rescaled process $\overline{M}^n$ introduced in Section 1.3 of this work provides a tractable analogue in dimension $d \geq 2$ to the one-dimensional stochastic Fisher-KPP equation with small noise when n is large.

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 [7] 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 is replaced by offspring of a parent chosen from the population immediately before the event (a precise definition of the process is given below). The Poisson structure renders the process particularly amenable to analytic study. In the neutral setting, which has been studied rather extensively (see [8] 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 [16]). Of course there are many ways to modify the selection mechanism. For example, as in Definition 1.3 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.* [17].

Let us turn to a precise definition. All the random objects in this section are defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

First we describe the state space of the process, borrowing some results from [49] 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

$$\mathcal{M}_\lambda := \left\{ M \text{ measure on } \mathbb{R}^d \times \{0,1\} \colon \forall f \in C_c(\mathbb{R}^d), \int_{\mathbb{R}^d \times \{0,1\}} f(x) M(\mathrm{d}x, \mathrm{d}\kappa) = \int_{\mathbb{R}^d} f(x) \mathrm{d}x \right\} \tag{1.2}$$

of such measures is equipped with the topology of vague convergence, which makes it a compact set (*cf.* Lemma 1.1 in [49]). Here $C_c(\mathbb{R}^d)$ denotes the space of all compactly supported continuous functions on $\mathbb{R}^d$. A standard decomposition theorem (see *e.g.* [29], p. 561) gives us the existence of a measurable mapping $w_t : \mathbb{R}^d \to [0, 1]$ such that

$$M_t(\mathrm{d}x, \mathrm{d}\kappa) = \big(w_t(x)\delta_0(\mathrm{d}\kappa) + (1 - w_t(x))\delta_1(\mathrm{d}\kappa)\big)\,\mathrm{d}x. \tag{1.3}$$

Morally, $w_t(x)$ represents the local fraction of individuals of type 0 at site $x \in \mathbb{R}^d$ at time t, and we abuse notation and call it the "density" of M. 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 $(M_t)_{t \geq 0}$.

For every $f \in C_c(\mathbb{R}^d)$ and every $F \in C^1(\mathbb{R})$ (the space of all continuously differentiable functions on $\mathbb{R}$), let us set

$$\langle w, f \rangle := \int_{\mathbb{R}^d} w(x) f(x) \mathrm{d}x \tag{1.4}$$

and let us define the function $\Psi_{F,f}$ on $\mathcal{M}_\lambda$ by

$$\Psi_{F,f}(M) := F(\langle w, f \rangle) = F\left(\int_{\mathbb{R}^d \times \{0,1\}} f(x) \mathbb{1}_{\{0\}}(\kappa) M(\mathrm{d}x, \mathrm{d}\kappa) \right), \tag{1.5}$$

where w is any representative of the density of M. These functions will prove particularly useful for the following reason.

Lemma 1.1. *The set of functions of the form $\Psi_{F,f}$, $F \in C^1(\mathbb{R})$ and $f \in C_c(\mathbb{R}^d)$, is dense in $C(\mathcal{M}_\lambda)$ for the supremum norm topology.*

Proof of Lemma 1.1. Since we endow $\mathcal{M}_\lambda$ with the topology of vague convergence, the set of all functions of the form

$$M \mapsto G\left(\int_{\mathbb{R}^d \times \{0,1\}} \varphi(x, \kappa) M(\mathrm{d}x, \mathrm{d}\kappa) \right), \tag{1.6}$$

with $G \in C^1(\mathbb{R})$ and $\varphi \in C_c(\mathbb{R}^d \times \{0,1\})$ is dense in $C(\mathcal{M}_\lambda)$. But if w is a representative of the density of M, we can write

$$\int_{\mathbb{R}^d \times \{0,1\}} \varphi(x,\kappa) M(\mathrm{d}x, \mathrm{d}\kappa) = \int_{\mathbb{R}^d} \varphi(x,0) w(x) \mathrm{d}x + \int_{\mathbb{R}^d} \varphi(x,1)(1 - w(x)) \mathrm{d}x$$

$$= \int_{\mathbb{R}^d} (\varphi(x,0) - \varphi(x,1)) w(x) \mathrm{d}x + \int_{\mathbb{R}^d} \varphi(x,1) \mathrm{d}x,$$

and so the mapping (1.6) can be rewritten in the form $F(\langle w, f \rangle)$, with

$$F(y) = G\left(y + \int_{\mathbb{R}^d} \varphi(x,1) \mathrm{d}x \right) \qquad \text{and} \qquad f(x) = \varphi(x,0) - \varphi(x,1).$$

By construction, in the above we have $F \in C^1(\mathbb{R})$ and $f \in C_c(\mathbb{R}^d)$. The set of functions of the form (1.6) is thus included in the set of functions of the form (1.5) and the conclusion follows. $\qquad \square$

In order to gain a feeling for the process, let us first give a non-rigorous description based on the two independent Poisson point processes of "neutral" and "selective" events mentioned above. This intuitive idea of how the SLFV with fecundity selection should evolve suggests a natural choice of operator $\mathcal{L}$ on functions of the form (1.5), see (1.9), and we shall show in Theorem 1.2 that for any probability measure P on $\mathcal{M}_\lambda$ describing the law of the initial condition, the martingale problem for $(\mathcal{L}, P)$ has a unique solution on the space of all measurable $\mathcal{M}_\lambda$-valued paths. Furthermore, this solution is a Markov process with a.s. càdlàg paths, and it has the Feller property. The SLFV with selection, with initial distribution P, can then be defined as the unique solution to this well-posed martingale problem (see Definition 1.3).

So first, the idea. 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

$$\int_0^\infty r^d \int_0^1 u \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) < \infty, \qquad \text{and} \qquad \int_0^\infty r^d \int_0^1 u \nu'_r(\mathrm{d}u) \mu'(\mathrm{d}r) < \infty. \qquad (1.7)$$

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 $\mathrm{d}t \otimes \mathrm{d}x \otimes \mu'(\mathrm{d}r)\nu'_r(\mathrm{d}u)$. Let $M_0 \in \mathcal{M}_\lambda$ be the (for now, deterministic) initial value of the process. The dynamics of $(M_t)_{t \geq 0}$ are 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

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

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

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

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

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

Let us now introduce the operator that will encode this dynamics. For every potential density $w : \mathbb{R}^d \to [0, 1]$, $x \in \mathbb{R}^d$, $r > 0$ and $u \in [0, 1]$, let us define

$$\Theta_{x,r,u}^+(w) := \mathbb{1}_{B(x,r)^c} w + \mathbb{1}_{B(x,r)}((1 - u)w + u), \qquad \text{and}$$

$$\Theta_{x,r,u}^-(w) := \mathbb{1}_{B(x,r)^c} w + \mathbb{1}_{B(x,r)}(1 - u)w. \tag{1.8}$$

These quantities will correspond to the value of the density immediately after an event (t, x, r, u) if the offspring are of type 0 or type 1 respectively.

Assuming that the above description corresponds to a well-posed martingale problem, we would expect the corresponding operator $\mathcal{L}$ to act on functions of the form (1.5) as follows (recall that V_r stands for the volume of a ball of radius r): for every $M \in \mathcal{M}_\lambda$,

$$\mathcal{L}\Psi_{F,f}(M) = \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \Big[w(y) F(\langle \Theta_{x,r,u}^+(w), f \rangle) \tag{1.9}$$

$$+ (1 - w(y)) F(\langle \Theta_{x,r,u}^-(w), f \rangle) - F(\langle w, f \rangle) \Big] dy \, \nu_r(du) \, \mu(dr) \, dx$$

$$+ \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)^2} \frac{1}{V_r^2} \Big[w(y)w(z) F(\langle \Theta_{x,r,u}^+(w), f \rangle)$$

$$+ (1 - w(y)w(z)) F(\langle \Theta_{x,r,u}^-(w), f \rangle) - F(\langle w, f \rangle) \Big] dy \, dz \, \nu_r'(du) \, \mu'(dr) \, dx.$$

Note that $\mathcal{L}\Psi_{F,f}(M)$ can also be expressed without referring to the density of M:

$$\mathcal{L}\Psi_{F,f}(M) \tag{1.10}$$

$$= \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r) \times \{0,1\}} \frac{1}{V_r} \Big[\mathbb{1}_{\{0\}}(\kappa) F\Big(\langle M, \mathbb{1}_{B(x,r)^c \times \{0\}} f \rangle + (1 - u)\langle M, \mathbb{1}_{B(x,r) \times \{0\}} f \rangle$$

$$+ u \int_{B(x,r)} f(x')dx' \Big) + \mathbb{1}_{\{1\}}(\kappa) F\Big(\langle M, \mathbb{1}_{B(x,r)^c \times \{0\}} f \rangle + (1 - u)\langle M, \mathbb{1}_{B(x,r) \times \{0\}} f \rangle \Big)$$

$$- F\Big(\langle M, \mathbb{1}_{\mathbb{R}^d \times \{0\}} f \rangle \Big) \Big] M(dy, d\kappa) \, \nu_r(du) \, \mu(dr) \, dx$$

$$+ \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{(B(x,r) \times \{0,1\})^2} \frac{1}{V_r^2} \Big[\mathbb{1}_{\{0\}}(\kappa) \mathbb{1}_{\{0\}}(\kappa') F\Big(\langle M, \mathbb{1}_{B(x,r)^c \times \{0\}} f \rangle$$

$$+ (1 - u)\langle M, \mathbb{1}_{B(x,r) \times \{0\}} f \rangle + u \int_{B(x,r)} f(x')dx' \Big)$$

$$+ \big(1 - \mathbb{1}_{\{0\}}(\kappa)\mathbb{1}_{\{0\}}(\kappa')\big) F\Big(\langle M, \mathbb{1}_{B(x,r)^c \times \{0\}} f \rangle + (1 - u)\langle M, \mathbb{1}_{B(x,r) \times \{0\}} f \rangle \Big)$$

$$- F\Big(\langle M, \mathbb{1}_{\mathbb{R}^d \times \{0\}} f \rangle \Big) \Big] M(dy, d\kappa) M(dy', d\kappa') \, \nu_r'(du) \, \mu'(dr) \, dx,$$

where we have used the bracket notation for some of the integrals to ease the notation.

Let $B_{\mathcal{M}_\lambda}[0,\infty)$ (*resp.*, $D_{\mathcal{M}_\lambda}[0,\infty)$) denote the space of all paths (*resp.*, càdlàg paths) with values in $\mathcal{M}_\lambda$. When needed, $D_{\mathcal{M}_\lambda}[0,\infty)$ is endowed with the standard Skorokhod topology and the associated Borel σ-field. Our first main result is the following.

Theorem 1.2. *Suppose that Condition* (1.7) *holds. Then, for every probability measure P on $\mathcal{M}_\lambda$ we have:*

(i) The $B_{\mathcal{M}_\lambda}[0,\infty)$-martingale problem for $(\mathcal{L}, P)$ is well-posed. That is, there exists a unique measurable process $(M_t)_{t\geq 0}$ with values in $\mathcal{M}_\lambda$ such that M_0 has law P and for every function $\Psi_{F,f}$ of the form (1.5),

$$\left(\Psi_{F,f}(M_t) - \Psi_{F,f}(M_0) - \int_0^t \mathcal{L}\Psi_{F,f}(M_s)\mathrm{d}s \right)_{t\geq 0} \tag{1.11}$$

is a martingale.

(ii) The process $(M_t)_{t\geq 0}$ in (i) is a Markov process and its semigroup is Feller. Moreover, it has càdlàg paths almost surely.

We can finally define the SLFV with fecundity selection in a rigourous way, assuming that Condition (1.7) is satisfied.

Definition 1.3 (SLFV with fecundity selection (SLFVS))**.** *Let P be a probability measure on $\mathcal{M}_\lambda$. We call* spatial Λ-Fleming-Viot process with fecundity selection, *with initial distribution P, the unique solution $(M_t)_{t\geq 0}$ to the martingale problem for $(\mathcal{L}, P)$ obtained in Theorem 1.2(i). In particular, by Theorem 1.2(ii), the SLFVS is a strong Markov process with càdlàg paths a.s.*

The proof of Theorem 1.2, given in Section 2 to ease the exposition, proceeds as follows. First, the result would be an obvious consequence of the Poisson point process formulation if we had chosen a compact set E 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. We thus start from this simple case and take a sequence of Poisson point processes whose intensities converge to the (possibly infinite) intensities of Π^N and Π^S on $\mathbb{R} \times E \times (0,\infty) \times [0,1]$. We then take a sequence of hypercubes growing to $\mathbb{R}^d$, and construct the process $(M_t)_{t\geq 0}$ of Theorem 1.2 as a potential limit for the corresponding processes. Uniqueness of such a limit is proved via a duality relation between any solution to the martingale problem (1.11) and a given family of solutions to the martingale problem satisfied by the particle system $(\Xi_t)_{t\geq 0}$ introduced in Section 3. This duality argument is a natural analogue of the argument guaranteeing uniqueness of the neutral SLFV [7], for which the dual process is a system of coalescing random walks interpreted as tracing the locations of the ancestors of individuals in a sample from the population. In the case with selection, we shall see the dual process as a system of branching and coalescing random walks that describes the locations of all *potential* ancestors of individuals in a sample from the population modelled by $(M_t)_{t\geq 0}$. The technical Condition (1.7) corresponds to Assumption 2.4 in [7] and expresses the fact that each "ancestral lineage" is affected by an event 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 (*cf.* [3, 26, 36] in the non-spatial setting).

We note that [20] 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 (1.7). The proof of existence is given (only) in the neutral case, but uniqueness remains open. The second construction, which requires Condition (1.7), 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 measure-valued dual process of "potential ancestors"

In this section, we first introduce a process $(\Xi_t)_{t \geq 0}$ with values in the set of all finite point measures on $\mathbb{R}^d$, whose evolution is driven by an independent copy of the Poisson point processes Π^N and Π^S. In Section 1.2.2, we state a duality relation between any solution to the martingale problem (1.11) and the process Ξ starting from suitable initial distributions. This duality is the analogue of the relation between the neutral SLFV and its "genealogical process" (see Theorem 4.2 in [7] for a general version of this relation, and Equation (8) in [9] for the particular case of two types of individuals). This is the content of Proposition 1.7, whose proof is deferred to Section 3 to ease the exposition. Although the duality presented here is very reminiscent of the standard notion of duality between two martingale problems (see [21], pp. 188–189, with $\alpha = \beta = 0$ for us), it differs in that the natural duality function

$$f(M, \Xi) = \prod_{i=1}^{k} w(x_i)$$

(for every $\Xi = \sum_{i=1}^{k} \delta_{x_i}$ and $M \in \mathcal{M}_\lambda$ with "density" w) suggested by classical population genetics is not well defined (see the discussion at the beginning of Section 1.2.2). Indeed, another representative w' of the density of M may differ from w at some of the x_i, yielding a different value for $f(M, \Xi)$. Consequently, we must modify the Ethier & Kurtz approach to duality, but Relation (1.25) stated in Proposition 1.7 will still take the same form as Relation (4.35) in [21].

1.2.1 Definition of the dual process

In contrast with the strategy adopted in Section 1.1 to construct the SLFVS, here we do not base the definition of the dual process on a martingale problem but, instead, we provide an explicit construction of this finite rate jump process in Definition 1.4. In Proposition 1.5, we show that this definition gives rise to a well-defined Markov process which also solves a martingale problem. This will be sufficient to obtain the duality relation stated in Proposition 1.7 and which is required to prove uniqueness of the solution to the martingale problem for $(\mathcal{L}, P)$ stated in (1.11).

Let us start with some heuristics on the form and dynamics of the dual process before formulating Definition 1.4. 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 $t-$. 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 informally 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 is 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 $x_1, \ldots, x_k \in \mathbb{R}^d$ at time 0, "the present", assuming that the population has been evolving for some very large time (that we do not specify). We want to trace back the locations of the "ancestors" of the individuals in the sample: that is, we want to go back into the past and describe at every earlier time t the set of locations in $\mathbb{R}^d$ from which the collection of types seen in our sample may have originated. To motivate the introduction of the process $(\Xi_t)_{t \geq 0}$ below, let us first analyse from a genealogical perspective what happens during each reproduction event. If the event is neutral (*i.e.*, belongs to Π^N), when an ancestor finds itself in the region affected by the event just after the latter has occurred, the probability that it belongs to the fraction u of the local population replaced during the event is precisely u. In this case, the "parent" of this ancestor was the "individual" whose type was chosen to be the one reproduced during the event, and as expounded above, the location of this "parent" is uniformly distributed over the affected area. Consequently, precisely at the time of this event in the past, the ancestral lineage corresponding to the ancestor found in this area jumps onto the location of the "parent". On the other hand, if (with probability $1 - u$) the ancestor does not belong to the fraction replaced, it is not an offspring of the "parent" and its ancestral lineage is not affected by the event (*i.e.*, it remains at the same spatial location). Finally, if there is more than one ancestor in this area, each of them belongs to the fraction of the population just replaced with probability u independently of each other, and the ancestral lineages of all those (and only those) who lie in this "offspring" population merge into a single ancestral lineage located at the position of the "parent". Note that this procedure is independent of the type of the "parent". During a selective event (*i.e.*, belonging to Π^S), this can no longer be the case; since we only follow the spatial locations from which the sampled types originate, and not their types, we are unable to decide which of the two "potential" parents is the true parent of the event. Instead we follow the locations of all "potential" ancestors. More precisely, as in a neutral event, every ancestor present in the area of the event just after it occurred belongs to the fraction of the local population just replaced with probability u, independently of each other. At the time of the event in the past, the ancestral lineages corresponding to the ancestors who belong to the "offspring" population merge, since they all have the same "parent". However, we do not know *a priori* from which of the two potential "parents" they inherit their types and so the new ancestral lineage instantly splits into two potential lineages, starting from the positions of the two potential "parents", independently and uniformly distributed over the area covered by the event. 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 [33, 40]. We shall sometimes use this informal description to see our dual process as a system of branching and coalescing jump processes, although this interpretation will appear much clearer when we describe the limiting "ancestral" processes that arise in the regimes of parameters on which we focus in Theorems 1.13 and 1.16.

We now give a formal definition of the process $(\Xi_t)_{t \geq 0}$ which will keep track of the locations of the potential ancestors of a sample taken from the current state of the population. To this end, observe that the time-reversed point processes

$$\overleftarrow{\Pi}^i := \big\{(-t, x, r, u) \, : \, (t, x, r, u) \in \Pi^i\big\}, \qquad i \in \{N, S\}, \tag{1.12}$$

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. The way in which events happen in both directions of time is thus the same in distribution. Hence, let $\widetilde{\Pi}^N$ and $\widetilde{\Pi}^S$ be independent copies of Π^N and Π^S respectively, defined on another probability space $(\mathbf{\Omega}, \mathcal{F}', \mathbf{P})$ (and so is the process Ξ introduced below).

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. The process $(\Xi_t)_{t \geq 0}$ will take its values in $\mathcal{M}_p(\mathbb{R}^d)$: each atom of Ξ_t will represent the location of a potential ancestor t units of time in the past.

Definition 1.4. *Let Ξ^0 be an $\mathcal{M}_p(\mathbb{R}^d)$-valued random variable, and let us define the process $(\Xi_t)_{t \geq 0}$ with initial value Ξ^0 as follows. We set $\Xi_0 = \Xi^0$ and, for convenience, at every time $t \geq 0$ we write*

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

where $N_t = \Xi_t(\mathbb{R}^d)$ and some of the ξ_t^i may be identical (by Lemma 2.3 in [28], the elements of this decomposition are measurable functions of Ξ_t). Note that the ordering by $1, \ldots, N_t$ of the atoms is arbitrary and will play no role in the updating of Ξ_t.

Then:

For every $(t, x, r, u) \in \widetilde{\Pi}^N$:

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

2. *If at least one atom ξ_{t-}^i is marked, to form Ξ_t we remove all the marked atoms from Ξ_{t-} and we add a Dirac mass at a location which is drawn uniformly at random from within $B(x, r)$.*

For every $(t, x, r, u) \in \widetilde{\Pi}^S$:

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

2. *If at least one atom ξ_{t-}^i is marked, to form Ξ_t we remove all the marked atoms from Ξ_{t-} and we add two Dirac masses at locations which 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 (unless $\Xi_0 = 0$), since any removal is accompanied by the insertion of at least one new atom.

Before stating the result showing that this definition gives rise to a well-defined Markov process, let us introduce the operator $\mathcal{G}$ which will turn out to be the extended generator of $(\Xi_t)_{t \geq 0}$ (*i.e.*, the operator on which the martingale problem satisfied by Ξ is based). Let $C_b^1(\mathbb{R})$ denote the set of all functions on $\mathbb{R}$ which are bounded, of class C^1 and whose first derivatives are bounded. Let also $\mathcal{B}_b(\mathbb{R}^d)$ denote the set of all bounded measurable functions on $\mathbb{R}^d$. For every $F \in C_b^1(\mathbb{R})$ and $f \in \mathcal{B}_b(\mathbb{R}^d)$, we define the function $\Phi_{F,f}$ by

$$\Phi_{F,f}(\Xi) := F(\langle \Xi, f \rangle), \qquad \forall \Xi \in \mathcal{M}_p(\mathbb{R}^d), \tag{1.13}$$

where $\langle \Xi, f \rangle = \int f(x) \Xi(\mathrm{d}x)$. Finally, we define the function $\mathcal{G}\Phi_{F,f}$ as follows. For every $\Xi = \sum_{i=1}^{l} \delta_{x_i} \in \mathcal{M}_p(\mathbb{R}^d)$,

$$
\begin{aligned}
\mathcal{G}\Phi_{F,f}(\Xi) := & \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \Bigg[\sum_{\substack{D \subseteq I_{x,r}(\Xi) \\ |D| \geq 1}} u^{|D|}(1-u)^{|I_{x,r}(\Xi) \setminus D|} \\
& \times \left(F\left(\langle \Xi, f \rangle - \sum_{i \in D} f(x_i) + f(z) \right) - F(\langle \Xi, f \rangle) \right) \Bigg] \mathrm{d}z \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x \\
& + \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)^2} \frac{1}{V_r^2} \Bigg[\sum_{\substack{D \subseteq I_{x,r}(\Xi) \\ |D| \geq 1}} u^{|D|}(1-u)^{|I_{x,r}(\Xi) \setminus D|} \\
& \times \left(F\left(\langle \Xi, f \rangle - \sum_{i \in D} f(x_i) + f(z) + f(z') \right) - F(\langle \Xi, f \rangle) \right) \Bigg] \mathrm{d}z\mathrm{d}z' \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \mathrm{d}x,
\end{aligned}
\tag{1.14}
$$

where

$$
I_{x,r}(\Xi) = \{ i \in \{1, \ldots, l\} \, : \, x_i \in B(x,r) \}
\tag{1.15}
$$

is the set of atoms of Ξ sitting in the closed ball $B(x,r)$ and by convention, the sum over $D \subset I_{x,r}(\Xi), |D| \geq 1$ is set to 0 if $I_{x,r}(\Xi)$ is empty. Note again that by Lemma 2.3 in [28], the elements $l, x_1, \ldots, x_l$ of the decomposition of Ξ are measurable functions of Ξ, and so the mapping $\mathcal{G}\Phi_{F,f}$ is a well-defined measurable function on $\mathcal{M}_p(\mathbb{R}^d)$.

Proposition 1.5. *The process $(\Xi_t)_{t \geq 0}$ of Definition 1.4 is a well-defined Markov jump process with values in $\mathcal{M}_p(\mathbb{R}^d)$. In addition, if there exists $K > 0$ such that $\mathbf{P}[\Xi^0(\mathbb{R}^d) \leq K] = 1$, then for every $F \in C_b^1(\mathbb{R})$ and $f \in \mathcal{B}_b(\mathbb{R}^d)$, the process*

$$
\left(\Phi_{F,f}(\Xi_t) - \Phi_{F,f}(\Xi_0) - \int_0^t \mathcal{G}\Phi_{F,f}(\Xi_s) \, \mathrm{d}s \right)_{t \geq 0}
\tag{1.16}
$$

is a martingale.

Proof of Proposition 1.5. Let us first argue that the process $(\Xi_t)_{t \geq 0}$ of Definition 1.4 is well defined for all time $t \geq 0$. Let us focus on a given atom in Ξ_t (for some $t \geq 0$), say at $z \in \mathbb{R}^d$. Since it is affected by a reproduction event only if it lies in the area of the event and if it is marked (which happens with a prescribed probability u), by construction the rate at which this atom is impacted by an event is given by

$$
\begin{aligned}
& \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}_{\{|z-x| \leq r\}} u \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x + \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}_{\{|z-x| \leq r\}} u \, \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \mathrm{d}x \\
& = \int_0^\infty \int_0^1 V_r u \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) + \int_0^\infty \int_0^1 V_r u \, \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) := C_0 < \infty
\end{aligned}
\tag{1.17}
$$

(where the finiteness of C_0 comes from Condition (1.7)), and so the total rate at which any of the atoms of Ξ_t is affected, and hence Ξ jumps, is bounded from above by $C_0 \Xi_t(\mathbb{R}^d)$. Furthermore, the number of atoms in Ξ_t can increase only during an event of $\widetilde{\Pi}^S$, and by at most one (if only one atom is erased and two atoms are created during a selective event). Consequently, the total number of atoms in Ξ_t is stochastically bounded by the number of particles in a Yule process starting with $\Xi_0(\mathbb{R}^d)$ particles, each of which splits into two at constant rate

$$
\int_0^\infty \int_0^1 V_r u \, \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r),
\tag{1.18}
$$

independently of each other. Combining the above with the fact that $\Xi_0(\mathbb{R}^d)$ is finite a.s., we obtain that with probability one the total mass of Ξ_t is finite for every $t \geq 0$ and there is no accumulation of jumps in finite time. That is, $(\Xi_t)_{t \geq 0}$ is a finite rate jump process defined for all time $t \geq 0$. The fact that Ξ is Markovian then comes from the Poisson point process structure of its evolution.

Let us now give a bound on $\mathcal{G}\Phi_{F,f}(\Xi)$, defined in (1.14), to prove first that the operator $\mathcal{G}$ is well-defined on the set of test functions considered, and second that $(\Xi_t)_{t \geq 0}$ is indeed solution to the martingale problem (1.16). To this end, let $F \in C_b^1(\mathbb{R})$, $f \in \mathcal{B}_b(\mathbb{R}^d)$ and $\Xi \in \mathcal{M}_p(\mathbb{R}^d)$. Denoting the sup norm by $\|\cdot\|$ and applying Taylor's theorem to the function F, we can write

$$\left| \mathcal{G}\Phi_{F,f}(\Xi) \right| \tag{1.19}$$
$$\leq \|F'\| \Xi(\mathbb{R}^d) \|f\| \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \left[\sum_{\substack{D \subseteq I_{x,r}(\Xi) \\ |D| \geq 1}} u^{|D|} (1-u)^{|I_{x,r}(\Xi) \setminus D|} \right] \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x$$
$$+ \|F'\| (\Xi(\mathbb{R}^d) + 1) \|f\| \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \left[\sum_{\substack{D \subseteq I_{x,r}(\Xi) \\ |D| \geq 1}} u^{|D|} (1-u)^{|I_{x,r}(\Xi) \setminus D|} \right] \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \mathrm{d}x.$$

Next, using the bounds $|I_{x,r}(\Xi)| = \Xi(B(x,r)) \leq \Xi(\mathbb{R}^d)$,

$$\sum_{\substack{D \subseteq I_{x,r}(\Xi) \\ |D| \geq 1}} u^{|D|} (1-u)^{|I_{x,r}(\Xi) \setminus D|} = 1 - (1-u)^{|I_{x,r}(\Xi)|} \leq u \, |I_{x,r}(\Xi)| \leq u \, \Xi(\mathbb{R}^d) \mathbb{1}_{\{\Xi(B(x,r)) > 0\}},$$

and

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 u \, \Xi(\mathbb{R}^d) \mathbb{1}_{\{\Xi(B(x,r)) > 0\}} \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x$$
$$\leq \Xi(\mathbb{R}^d) \int_0^\infty \int_0^1 u \, \mathrm{Vol}(\mathrm{Supp}(\Xi) + B(0,r)) \nu_r(\mathrm{d}u) \mu(\mathrm{d}r)$$
$$\leq C_d \, \Xi(\mathbb{R}^d)^2 \int_0^\infty \int_0^1 u r^d \nu_r(\mathrm{d}u) \mu(\mathrm{d}r),$$

where $\mathrm{Supp}(\Xi)$ denotes the (discrete) support of Ξ and C_d is the volume of a d-dimensional ball of radius 1, we obtain that

$$\left| \mathcal{G}\Phi_{F,f}(\Xi) \right| \leq C_d \|F'\| \|f\| \Xi(\mathbb{R}^d)^2 \left(\Xi(\mathbb{R}^d) + 1 \right)$$
$$\times \left(\int_0^\infty \int_0^1 u r^d \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) + \int_0^\infty \int_0^1 u r^d \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \right). \tag{1.20}$$

From this we can first conclude that the operator $\mathcal{G}$ is indeed well-defined on the set of functions of the form $\Phi_{F,f}$, with $F \in C_b^1(\mathbb{R})$ and $f \in \mathcal{B}_b(\mathbb{R}^d)$. It is then straightforward to see that for every such test function and every $\Xi \in \mathcal{M}_p(\mathbb{R}^d)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{E}_\Xi \left[\Phi_{F,f}(\Xi_t) \right] \Big|_{t=0} = \mathcal{G}\Phi_{F,f}(\Xi). \tag{1.21}$$

Since $\Xi^0(\mathbb{R}^d) \leq K$ a.s., the Yule process with branching rate given in (1.18) that dominates the number of particles in Ξ has finite moments at any time $t \geq 0$ (see Equation (5) in [50] for the original derivation of the distribution of the number of individuals at any time t in a Yule process, which is negative binomial for any initial number of individuals),

and so the expression on the r.h.s. of (1.20) applied to Ξ_t is integrable for any $t \geq 0$. Combined with the boundedness of F and Fubini's theorem, this yields that

$$\Phi_{F,f}(\Xi_t) - \Phi_{F,f}(\Xi_0) - \int_0^t \mathcal{G}\Phi_{F,f}(\Xi_s)\,\mathrm{d}s$$

is integrable for every $t \geq 0$. Together with (1.21), this allows us to conclude that Ξ is indeed a solution to the martingale problem (1.16) with initial distribution the law of Ξ^0. $\qquad\square$

1.2.2 Duality relation between $(M_t)_{t \geq 0}$ and $(\Xi_t)_{t \geq 0}$

A key feature of our model, that we shall use repeatedly, is the fact that the processes $(M_t)_{t \geq 0}$ and $(\Xi_t)_{t \geq 0}$ are dual to each other if we restrict our attention to initial distributions on $\mathcal{M}_p(\mathbb{R}^d)$ of a particular form (in essence, the atoms of Ξ_0 should be random and have a distribution absolutely continuous with respect to Lebesgue measure – see below). As in the neutral case [7, 9], this will allow us to transfer the information we obtain on $(M_t)_{t \geq 0}$ onto $(\Xi_t)_{t \geq 0}$, and vice versa. Because we want to use this property in the proof of existence of $(M_t)_{t \geq 0}$ in Section 2 (more precisely, to show that there is at most one solution to the martingale problem for $(\mathcal{L}, \delta_{M^0})$), Proposition 1.7 is phrased in a more general way and relates $(\Xi_t)_{t \geq 0}$ to any solution to the martingale problem for $\mathcal{L}$.

The difficulty that we face is that the density of any element of $\mathcal{M}_\lambda$ is only defined Lebesgue a.e. and so the usual test functions used to establish such dualities in population genetics when the underlying geographical space is discrete, which take the form

$$D(M, \Xi) := \exp\left(\int_{\mathbb{R}^d} \ln w(x)\Xi(\mathrm{d}x)\right) = \prod_{i=1}^k w(x_i) \tag{1.22}$$

for $M \in \mathcal{M}_\lambda$ with density w and $\Xi = \sum_{i=1}^k \delta_{x_i}$, will not make sense. However, if, instead of taking deterministic points $x_1, \ldots, x_k$, we take random points, with a distribution which has a density ψ with respect to Lebesgue measure on $(\mathbb{R}^d)^k$, then writing μ_ψ for the law of the random measure constructed in this way, we have for any $M \in \mathcal{M}_\lambda$

$$\int_{\mathcal{M}_p(\mathbb{R}^d)} D(M, X)\mu_\psi(dX) = \int_{(\mathbb{R}^d)^k} \psi(x_1, \ldots, x_k)\left\{\prod_{j=1}^k w(x_j)\right\} \mathrm{d}x_1 \cdots \mathrm{d}x_k \tag{1.23}$$

$$= \int_{(\mathbb{R}^d \times \{0,1\})^k} \psi(x_1, \ldots, x_k)\left\{\prod_{j=1}^k \mathbb{1}_{\{0\}}(\kappa_j)\right\} M(\mathrm{d}x_1, \mathrm{d}\kappa_1) \cdots M(\mathrm{d}x_k, \mathrm{d}\kappa_k),$$

which *is* well-defined (and independent of the representative w of the density of M). The following property will therefore be very useful for the main result of this section, Proposition 1.7.

Lemma 1.6. *Suppose that the distribution of Ξ_0 has the form μ_ψ, for some $k \geq 1$ and some density function ψ on $(\mathbb{R}^d)^k$. Then for every $t \geq 0$ and every $j \in \{1, 2, \ldots\}$, conditionally on $N_t = j$, the law of $(\xi_t^1, \ldots, \xi_t^j)$ is absolutely continuous with respect to Lebesgue measure on $(\mathbb{R}^d)^j$.*

Proof of Lemma 1.6. The desired property follows from the fact that during every event of $\widetilde{\Pi}^N$ or $\widetilde{\Pi}^S$, the distribution of each "potential parent" is uniformly distributed over the area of the event, independently of the current locations of the atoms of Ξ_s. Hence, each time a point from Ξ_s is removed, the one or two atoms that are added have a location whose law is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$, while the

thinning procedure used to remove the points already in Ξ_s preserves the property that the distribution of the locations of the remaining atoms has a density with respect to Lebesgue measure. $\qquad\square$

Setting for every vector of k locations $(x_1, \ldots, x_k) \in (\mathbb{R}^d)^k$

$$\Xi[x_1, \ldots, x_k] := \sum_{i=1}^{k} \delta_{x_i} \in \mathcal{M}_p(\mathbb{R}^d), \tag{1.24}$$

writing $\Xi_0 \sim \mu_\psi$ to denote the fact that the random variable Ξ_0 has law μ_ψ, and recalling that $\mathbb{P}$ (*resp.*, $\mathbf{P}$) is the probability measure on the space on which $(M_t)_{t\geq0}$ (*resp.*, $(\Xi_t)_{t\geq0}$) is defined, we can now state the following result, whose proof is given in Section 3.

Proposition 1.7. *Let $M^0 \in \mathcal{M}_\lambda$, $k \in \{1, 2, \ldots\}$ and let ψ be a density function on $(\mathbb{R}^d)^k$. Then any solution $(M_t)_{t\geq0}$ to the $B_{\mathcal{M}_\lambda}[0, \infty)$-martingale problem for $(\mathcal{L}, \delta_{M^0})$ satisfies: for every $t \geq 0$*

$$\int_{\mathcal{M}_p(\mathbb{R}^d)} \mathbb{E}\big[D(M_t, X) \,|\, M_0 = M^0\big] \mu_\psi(dX) = \mathbf{E}\big[D(M^0, \Xi_t) \,|\, \Xi_0 \sim \mu_\psi\big]. \tag{1.25}$$

Equivalently, by Fubini's theorem and (1.23),

$$\mathbb{E}_{M^0}\bigg[\int_{(\mathbb{R}^d)^k} \psi(x_1, \ldots, x_k) \bigg\{ \prod_{j=1}^{k} w_t(x_j) \bigg\} \, dx_1 \cdots dx_k \bigg]$$

$$= \int_{(\mathbb{R}^d)^k} \psi(x_1, \ldots, x_k) \mathbf{E}_{\Xi[x_1, \ldots, x_k]} \bigg[\prod_{j=1}^{N_t} w^0\big(\xi_t^j\big) \bigg] \, dx_1 \cdots dx_k. \tag{1.26}$$

Remark 1.8. By linearity, (1.26) also holds for every $\psi \in \mathbb{L}^1((\mathbb{R}^d)^k)$. In addition, to keep the notation simple we have restricted our attention to deterministic initial values M^0, but the proof of Proposition 1.7 shows that a similar duality formula holds when M^0 is any $\mathcal{M}_\lambda$-valued random variable. See in particular (3.12), which only needs to be integrated with respect to the law of M^0 to yield the result.

Remark 1.9. One may try to use Proposition 1.7 to prove uniqueness of the solution to the martingale problem for $(\mathcal{G}, \mu_\psi)$, which is satisfied by $(\Xi_t)_{t\geq0}$. To this end, in the statement of Proposition 1.7, we would like to replace the process $(\Xi_t)_{t\geq0}$ of Definition 1.4 by any process $(\widetilde{\Xi}_t)_{t\geq0}$ solving the same martingale problem (1.16). However, in contrast with the explicit construction of Ξ which immediately yields Lemma 1.6, one cannot see from the martingale problem formulation that at any time $t \geq 0$, conditionally on $\widetilde{\Xi}_t(\mathbb{R}^d)$, the law of the locations in $\mathbb{R}^d$ of the atoms of $\widetilde{\Xi}_t$ is absolutely continuous with respect to Lebesgue measure on $(\mathbb{R}^d)^{\widetilde{\Xi}_t(\mathbb{R}^d)}$. But this property is crucial to the proof of Proposition 1.7, and therefore we cannot prove that (1.26) holds more generally than for the process Ξ of Definition 1.4.

1.3 Convergence of the rescaled SLFVS to Fisher-KPP processes

Now that we have introduced the spatial Λ-Fleming-Viot process with selection and its dual process of "potential ancestors", we turn to the main questions of this work: can we recover the solution to the deterministic or the stochastic Fisher-KPP equation as a scaling limit of the SLFVS, and how does the introduction of (a particular form of) rare but geographically extended extinction-recolonisation events impact the law of the limiting process under analogous scaling assumptions? Recall that the Fisher-KPP equation is a classical model for the wave of advance of a slightly favourable allele in a very dense population, in which individuals reproduce locally, so that changes in local

allele frequencies are continuous in time and space. In our framework, this corresponds to focusing on a regime of parameters in which selective events are rare compared to neutral events, the impact of every event (*i.e.*, the fraction of the local population actually affected by the event) is very small, and the event radii have a bounded variance. Therefore, writing n for a parameter that we shall let tend to infinity, in what follows we shall assume that there exist $\delta, \gamma > 0$ such that the relative frequency of selective events to neutral events scales like $n^{-\delta}$, and the impact of every event scales like $n^{-\gamma}$. Furthermore, in the first case that we consider below, all events will have the same radius (but the assumption of bounded radii would lead to the same type of results), and this assumption will be relaxed in the second case we consider. Since we are interested in the patterns of variation that we see under this model if we look over large spatial and temporal scales, we shall need a third parameter $\beta \geq 0$ to describe the relevant spatial scale to be considered: time will be scaled by a factor n when space will be scaled by a factor n^β.

Let us be more precise about our assumptions. First, we concentrate on the particular case in which the intensity measures of the Poisson point processes of reproduction events (see Definition 1.3) satisfy

$$\mu'(\mathrm{d}r)\nu_r'(\mathrm{d}u) = s_n \mu(\mathrm{d}r)\nu_r(\mathrm{d}u) \tag{1.27}$$

for a parameter s_n of the form $\sigma n^{-\delta}$, with $\sigma > 0$ independent of n. That is, the distribution of radii and impacts are the same for neutral and selective events, but neutral events happen n^δ/σ times faster than events during which type 1 individuals are favoured. We also 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, we assume that all events (neutral and selective) have impact $u_n = un^{-\gamma}$, where $u > 0$ is independent of n. In formulae:

$$\nu_r(\mathrm{d}u) = \nu_r'(\mathrm{d}u) = \delta_{u_n}(\mathrm{d}u) \qquad \text{for every } r > 0, \tag{1.28}$$

implying in particular that $\mu' = s_n\mu$ by (1.27). The assumptions that

$$u_n = \frac{u}{n^\gamma}, \qquad \text{and} \qquad s_n = \frac{\sigma}{n^\delta} \tag{1.29}$$

mirror 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 [16]).

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 $\gamma = 1/3$, $\delta = 2/3$, $\beta = 1/3$ and set (in any dimension)

$$\overline{w}_t^n(x) := \frac{1}{V_R} M_{nt}\big(B(n^\beta x, R) \times \{0\}\big) = \frac{1}{V_R} \int_{B(n^{1/3}x, R)} w_{nt}(y)\,\mathrm{d}y, \tag{1.30}$$

  where we recall that V_R stands for the volume of a d-dimensional ball of radius R. 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 $n^\beta = n^{1/3}$ (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

$$\mu(\mathrm{d}r) = \frac{\mathbb{1}_{\{r \geq 1\}}}{r^{d+\alpha+1}} \, \mathrm{d}r, \tag{1.31}$$

and

$$\overline{w}_t^n(x) := \frac{1}{V_1} M_{nt}\big(B(n^\beta x, 1) \times \{0\}\big) = \frac{1}{V_1} \int_{B(n^\beta x, 1)} w_{nt}(y) \, \mathrm{d}y, \tag{1.32}$$

with

$$\beta = \frac{1}{2\alpha - 1}, \quad \gamma = \frac{\alpha - 1}{2\alpha - 1} \quad \text{and} \quad \delta = \frac{\alpha}{2\alpha - 1}. \tag{1.33}$$

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 (1.7) are satisfied; in particular, the indicator function $\mathbb{1}_{\{r \geq 1\}}$ 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. Indeed, it is not simply obtained by a change in space and time coordinates of the measures $(M_t)_{t \geq 0}$ (with parameters s_n, u_n, ...) but its density $\overline{w}_t^n$ at any time is defined as an average over a ball of fixed radius R (or 1 in the stable case) of the density of M_{nt}. Therefore, the law of the "parental" type(s) picked during an event cannot be expressed in terms of a sampling from the current value of $\overline{M}^n$ and, additionally, the change in the value of each $\overline{w}_t^n(y)$ due to an event centered in $B(x,r)$ will depend on the geometry of the intersection $B(n^\beta y, R) \cap B(x,r)$. Hence, the evolution of quantities of the form $\langle \overline{w}_t^n, f \rangle$, with $f \in C_c(\mathbb{R}^d)$, cannot be fully described in terms of $\overline{M}_t^n$.

Remark 1.10. 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 (*cf.* Section 4). 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 4.

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 (1.4). Our main results are as follows, starting with the case of "local" reproduction.

Theorem 1.11 (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 continuous sample paths, starting at $M_0^\infty = M_0$. The limiting process is characterised as follows. Let*

$$\Gamma_R = \frac{1}{V_R} \int_{B(0,R)} \int_{B(x,R)} (z_1)^2 \mathrm{d}z \mathrm{d}x \tag{1.34}$$

(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

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

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

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

(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

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

(independently of the representative chosen at every time t), where $\dot{\mathcal{W}}$ 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

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

Remark 1.12. As we shall explain in Section 4, our choice of $\beta = 1/3 = \gamma$ and $\delta = 2/3$ is obtained by solving

$$1 - \gamma = 2\beta, \qquad 1 - \delta - \gamma = 0, \qquad \text{and } \beta = \gamma.$$

This set of three equations guarantees that in one dimension, the limiting process M^∞ is solution to the stochastic Fisher-KPP equation. If we replace the last condition by the inequality $0 < \beta < \gamma$, then the sequence of processes $(\overline{M}^n)_{n \geq 1}$ still converges, to a limit which is solution to the deterministic Fisher-KPP equation in any dimension (including $d = 1$).

Theorem 1.11 has a counterpart for the correspondingly rescaled dual process. For every $n \in \mathbb{N}$, let $(\Xi_t)_{t \geq 0}$ be the process of Definition 1.4 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 $u_n = u n^{-1/3}$. $(\Xi_t)_{t \geq 0}$ is thus dual to the unscaled process $(M_t)_{t \geq 0}$ with the same parameters, in the sense of Proposition 1.7 (to ease the notation, the dependence on n of these processes is not reported). Now, define the rescaled process $(\Xi_t^n)_{t \geq 0}$ so that for every $t \geq 0$,

$$\Xi_t^n = \sum_{i=1}^{N_t^n} \delta_{\xi_t^{n,i}} := \sum_{i=1}^{N_{nt}} \delta_{n^{-1/3} \xi_{nt}^i}. \tag{1.35}$$

Recall that the space $\mathcal{M}_p(\mathbb{R}^d)$ of finite point measures on $\mathbb{R}^d$ is endowed with the topology of *weak* convergence, and recall also the definition of the law μ_ψ on $\mathcal{M}_p(\mathbb{R}^d)$ given in the paragraph below (1.22).

Theorem 1.13 (Fixed radius – dual)**.** *Let $k \in \{1, 2, \ldots\}$, ψ be a probability density on $(\mathbb{R}^d)^k$ and suppose that for any $n \geq 1$, Ξ_0^n has law μ_ψ. Then, as $n \to \infty$, $(\Xi_t^n)_{t \geq 0}$ converges in distribution in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0, \infty)$ to a limiting Markov process $(\Xi_t^\infty)_{t \geq 0}$ characterised as follows: Ξ_0^∞ has law μ_ψ and*

(i) When $d = 1$, $(\Xi_t^\infty)_{t \geq 0}$ is a system of branching and coalescing Brownian motions, 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. In addition to branching and diffusing, each pair of particles, independently, also coalesces at rate $4R^2u^2$ times their collision local time.

(ii) When $d \geq 2$, $(\Xi_t^\infty)_{t \geq 0}$ is a branching Brownian motion (with no coalescence), 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.

To state the corresponding results for stable radii, we need some more notation. 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, for every $f \in C_c^\infty(\mathbb{R}^d)$ we set

$$\mathcal{D}^\alpha f(y) = u \int_{\mathbb{R}^d} \Phi(|z - y|)(f(z) - f(y)) \mathrm{d}z. \tag{1.36}$$

We shall check in Lemma 6.1 that this defines the infinitesimal generator of a symmetric stable process (that is, it is a constant multiple of the fractional Laplacian). Our results for stable radii are then as follows.

Theorem 1.14 (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 $(M_t^\infty)_{t \geq 0}$ starting at M_0. Furthermore, if $\mathcal{D}^\alpha$ denotes the generator of the symmetric α-stable process defined in (1.36), 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 $f, g \in C_c^\infty(\mathbb{R})$,

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

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 $t \geq 0$,

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

Remark 1.15. Again, our choice of values for β, γ and δ is obtained by solving

$$1 - \gamma = \alpha\beta, \qquad 1 - \delta - \gamma = 0, \qquad \text{and } (\alpha - 1)\beta = \gamma.$$

(see Section 4) in order to obtain a limiting process M^∞ which is stochastic in one dimension. If we replace the last condition by the inequality $0 < (\alpha - 1)\beta < \gamma$, then (in any dimension) $(\overline{M}^n)_{n \geq 0}$ converges to a deterministic limit which is characterised as in the statement of Theorem 1.14(ii).

Likewise, letting $(\Xi_t)_{t \geq 0}$ be the $\mathcal{M}_p(\mathbb{R}^d)$-valued process which is dual to the unscaled process $(M_t)_{t \geq 0}$ corresponding to the case of stable radii with parameters $u_n = u/n^\gamma$ and $s_n = \sigma/n^\delta$, and defining the rescaled process $(\Xi_t^n)_{t \geq 0}$ in such a way that for every $t \geq 0$,

$$\Xi_t^n = \sum_{i=1}^{N_t^n} \delta_{\xi_t^{n,i}} := \sum_{i=1}^{N_{nt}} \delta_{n^{-\beta}\xi_{nt}^i}, \tag{1.37}$$

(with the values of β, γ, δ given in (1.33)), we have the following convergence result.

Theorem 1.16 (Stable radii – dual). *Let $k \in \{1, 2, \ldots\}$, ψ be a probability density on $(\mathbb{R}^d)^k$ and suppose that for any $n \geq 1$, Ξ_0^n has law μ_ψ. Then, as $n \to \infty$, $(\Xi_t^n)_{t \geq 0}$ converges in distribution in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0, \infty)$ to a limiting Markov process $(\Xi_t^\infty)_{t \geq 0}$ characterised as follows: Ξ_0^∞ has law μ_ψ and*

(i) When $d = 1$, $(\Xi_t^\infty)_{t \geq 0}$ is a branching and coalescing stable process, in which particles follow 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 fully described by the generator $\mathcal{D}^\alpha$ defined in (1.36). In addition, each pair of particles, independently, coalesces at rate $4u^2/(\alpha - 1)$ times their collision local time.

(ii) When $d \geq 2$, $(\Xi_t^\infty)_{t \geq 0}$ is a branching stable process (with no coalescence), in which particles follow independent symmetric α-stable processes with generator $\mathcal{D}^\alpha$, and branch at rate $u\sigma V_1/\alpha$ into two new particles, started at the location of the parent.

In fact, we shall use knowledge of the limiting "population model" $(M_t^\infty)_{t \geq 0}$ to recover the corresponding limiting results for our rescaled duals. The difficulty with proving Theorems 1.13 and 1.16 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 [9] in the neutral case and in [18, 19] 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 large scale evolution of the spatial pattern of allele frequencies.

We close this section with a few remarks. First, one may observe from the expression of $\mathcal{D}^\alpha$ given in (1.36) 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. 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 [25], 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 (despite the fact that the impact u_n tends to zero, while the dependence on dimension mostly occurs when the impact remains fixed).

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 considers more general selection mechanisms, in which the strength and direction of selection depends on the local frequencies of 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 prove Theorem 1.2. In Section 3, we prove the duality relation stated in Proposition 1.7. In Section 4, we provide heuristic arguments to explain our rescalings. In Section 5, we turn to proving Theorem 1.11, the scaling limit in the case of fixed radii, and Theorem 1.13 which provides the corresponding result for the rescaled duals. In Section 6, we prove Theorems 1.14 and 1.16, the analogous results for stable radii. In Appendices A and B, we obtain continuity estimates for the rescaled SLFVS of Sections 5 and 6. In particular, these rather technical estimates are key ingredients in (and nice complements to) the proofs of Theorems 1.11 and 1.14.

2 Proof of Theorem 1.2 (existence of the SLFVS)

The strategy of the proof is the following. We start with a version of the process in which $\mathbb{R}^d$ is replaced by a hypercube E of finite sidelength and the measures μ and μ' are assumed to be finite. In this case, the total rate at which events happen is finite and the corresponding process is a well-defined measure-valued Markov jump process with a.s. càdlàg trajectories. We then proceed in two steps:

(i) We show existence when E has finite sidelength but μ and μ' are only σ-finite, by taking sequences of finite measures $(\mu^n)_{n \geq 1}$ and $(\mu'^n)_{n \geq 1}$ such that $\mu^n(\mathrm{d}r)$ converges to $\mu(\mathrm{d}r)$ (and the same with primes), and proving that the corresponding sequence of processes converges to a well-defined limit.

(ii) Given (i), we extend to $\mathbb{R}^d$ by considering a sequence of processes obtained by restricting to an increasing family of hypercubes $(E_n)_{n \geq 1}$ which exhaust the space, and proving that this sequence converges to the process $(M_t)_{t \geq 0}$ that we are seeking.

Both steps rely on Theorem 4.8.10 in [21], which states that provided we can show that

(a) The operator $\mathbf{L}$ on which the limiting martingale problem is based is included in the set $C_b(\mathcal{M}_\lambda) \times C_b(\mathcal{M}_\lambda)$, where $C_b(\mathcal{M}_\lambda)$ is the set of all bounded continuous functions on $\mathcal{M}_\lambda$;

(b) The limiting $D_{\mathcal{M}_\lambda}[0, \infty)$-martingale problem for $(\mathbf{L}, P^E)$ (where P^E is the distribution of the limit of $(M_0^{(n)})_{n \geq 1}$, in particular $P^{\mathbb{R}^d} = P$) has at most one solution;

(c) For every n, $M^{(n)}$ is a process with sample paths in $D_{\mathcal{M}_\lambda}[0, \infty)$ (here we follow Ethier and Kurtz in taking $(\mathcal{G}_t^n)_{t \geq 0}$ to be the natural filtration associated to $M^{(n)}$) and the sequence $(M^{(n)})_{n \geq 1}$ is relatively compact;

(d) There exists a countable set $\Gamma \subset [0,\infty)$ such that for every $(F,G) \in \mathbf{L}$,

$$\lim_{n\to\infty} \mathbb{E}\left[\left(F\big(M_{t+s}^{(n)}\big) - F\big(M_t^{(n)}\big) - \int_t^{t+s} G\big(M_u^{(n)}\big)\mathrm{d}u\right)\left(\prod_{i=1}^k h_i\big(M_{t_i}^{(n)}\big)\right)\right] = 0 \quad (2.1)$$

for all $k \geq 0$, $0 \leq t_1 < t_2 < \cdots < t_k \leq t < t+s$ with t_i, t, $t+s \notin \Gamma$, and $h_i \in C_b(\mathcal{M}_\lambda)$;

then there exists a solution $(M_t)_{t\geq 0}$ to the $D_{\mathcal{M}_\lambda}[0,\infty)$-martingale problem for $(\mathbf{L}, P^E)$ and $(M^{(n)})_{t\geq 0} \Rightarrow (M_t)_{t\geq 0}$ as $n \to \infty$. Item (b) will be a consequence of Proposition 1.7 and Remark 1.8, whose proofs are postponed until Section 3 for the sake of clarity. The other items will be checked one by one below. Once this is done, existence of a solution to the martingale problem in $D_{\mathcal{M}_\lambda}[0,\infty)$ will imply the existence of a solution in the larger space $B_{\mathcal{M}_\lambda}[0,\infty)$, and since uniqueness holds in $B_{\mathcal{M}_\lambda}[0,\infty)$ too (see the proof of item (b) below), this will show that the $B_{\mathcal{M}_\lambda}[0,\infty)$-martingale problem for $(\mathbf{L}, P^E)$ is well-posed. That is, Theorem 1.2(i) and the property that the trajectories of $(M_t)_{t\geq 0}$ are càdlàg a.s. will be proved. Furthermore, since (b) is satisfied for any distribution P^E on $\mathcal{M}_\lambda$, we shall be able to deduce from (a), (b) and Theorem 4.4.2(a) in [21] that the limiting process M is a Markov process with respect to its natural filtration. The last step will consist in showing that its semigroup is Feller.

Recall the definitions of $\Theta_{x,r,u}^+(w)$ and $\Theta_{x,r,u}^-(w)$ given in (1.8), and the notation $\|f\|$ (resp., $\|f\|_1$) for the supremum (resp., $\mathbb{L}^1$) norm of the function f. To simplify notation, we shall restrict our attention to initial distributions P of the form δ_{M^0} for some $M^0 \in \mathcal{M}_\lambda$. Indeed, the extension of (b) to a general P is covered by Remark 1.8, the bounds on the elements of the semi-martingale decomposition on which the proof of (c) rely are independent of the choice of the initial values for the processes of interest, and the proof of (d) can easily be generalised by using the linearity of the expectation and integrating all key equations with respect to $P(dM^0)$. From now on, we thus fix $M^0 \in \mathcal{M}_\lambda$.

Proof of (i)

Let E be some hypercube with sidelength ℓ, and let μ, μ' be the σ-finite measures on $(0,\infty)$ of Theorem 1.2. Let $(\mu^n)_{n\geq 1}$ and $(\mu'^n)_{n\geq 1}$ be two sequences of finite measures on $(0,\infty)$ such that, as $n \to \infty$,

$$\int_0^\infty \varphi(r)\mu^n(\mathrm{d}r) \nearrow \int_0^\infty \varphi(r)\mu(\mathrm{d}r) \quad \text{and} \quad \int_0^\infty \varphi(r)\mu'^n(\mathrm{d}r) \nearrow \int_0^\infty \varphi(r)\mu'(\mathrm{d}r), \quad (2.2)$$

for all measurable $\varphi \geq 0$. Let $\mathcal{M}_\lambda(E)$ be the analogue of $\mathcal{M}_\lambda$ (see (1.2)) when the "geographical" space $\mathbb{R}^d$ is replaced by E. That is, $\mathcal{M}_\lambda(E)$ is the set of all measures on $E \times \{0,1\}$ whose first marginal distribution is Lebesgue measure on E. It is also a compact space when endowed with the topology of vague convergence (note that since $E \times \{0,1\}$ is compact, the topology of vague convergence is the same as the topology of weak convergence). Let

$$M_E^0 := M^0\Big|_{E\times\{0,1\}} \in \mathcal{M}_\lambda(E)$$

be the measure induced by M^0 (the initial value fixed above) on $E \times \{0,1\}$, and for every $n \geq 1$, let $\Pi_E^{N,n}$ and $\Pi_E^{S,n}$ be independent Poisson point processes on $\mathbb{R} \times E \times (0,\infty) \times [0,1]$ with respective intensity measures $\mathrm{d}t \otimes \mathrm{d}x \otimes \mu^n(\mathrm{d}r)\nu_r(\mathrm{d}u)$ and $\mathrm{d}t \otimes \mathrm{d}x \otimes \mu'^n(\mathrm{d}r)\nu_r'(\mathrm{d}u)$. Finally, let $(M_t^{(n)})_{t\geq 0}$ be defined as in Definition 1.3, with Π^N replaced by $\Pi_E^{N,n}$, Π^S replaced by $\Pi_E^{S,n}$, and with $M_0^{(n)} = M_E^0$ (what we call the ball $B(x,r)$ in this case is $B_E(x,r) := B(x,r) \cap E$). For a given $n \geq 1$, the total rate at which an event of $\Pi_E^{N,n}$ or

$\Pi_E^{S,n}$ happens is

$$\int_E \int_0^\infty \int_0^1 \nu_r(\mathrm{d}u)\mu^n(\mathrm{d}r)\mathrm{d}x + \int_E \int_0^\infty \int_0^1 \nu'_r(\mathrm{d}u)\mu'^{\,n}(\mathrm{d}r)\mathrm{d}x$$
$$= \mathrm{Vol}(E)\big[\mu^n((0,\infty)) + \mu'^{\,n}((0,\infty))\big] < \infty, \qquad (2.3)$$

and so $M^{(n)}$ is a Markov jump process with jump rates uniformly bounded by the quantity in (2.3) and with càdlàg paths, solution to the martingale problem: $M_0^{(n)} = M_E^0$ and for every $F \in C^1(\mathbb{R})$ and $f \in C(E)$,

$$\left(\Psi_{F,f}\big(M_t^{(n)}\big) - \Psi_{F,f}\big(M_0^{(n)}\big) - \int_0^t \mathcal{L}^{(n)}\Psi_{F,f}\big(M_s^{(n)}\big)\mathrm{d}s\right)_{t\geq 0}$$

is a martingale (for the natural filtration associated to $M^{(n)}$), where $\Psi_{F,f}$ is defined as in (1.5) and the bounded continuous function $\mathcal{L}^{(n)}\Psi_{F,f}$ is defined by

$$\mathcal{L}^{(n)}\Psi_{F,f}(M) = \int_E \int_0^\infty \int_0^1 \int_{B_E(x,r)} \frac{1}{\mathrm{Vol}(B_E(x,r))} \Big[w(y)F(\langle\Theta_{x,r,u}^+(w),f\rangle) \qquad (2.4)$$
$$+ (1-w(y))F(\langle\Theta_{x,r,u}^-(w),f\rangle) - F(\langle w,f\rangle)\Big]\mathrm{d}y\,\nu_r(\mathrm{d}u)\,\mu^n(\mathrm{d}r)\,\mathrm{d}x$$
$$+ \int_E \int_0^\infty \int_0^1 \int_{B_E(x,r)^2} \frac{1}{\mathrm{Vol}(B_E(x,r))^2} \Big[w(y)w(z)F\big(\langle\Theta_{x,r,u}^+(w),f\rangle\big)$$
$$+ (1-w(y)w(z))F(\langle\Theta_{x,r,u}^-(w),f\rangle) - F(\langle w,f\rangle)\Big]\mathrm{d}y\,\mathrm{d}z\,\nu'_r(\mathrm{d}u)\,\mu'^{\,n}(\mathrm{d}r)\,\mathrm{d}x,$$

for every $M \in \mathcal{M}_\lambda(E)$. (As earlier, here w is any representative of the density of M and we have kept the notation $\Theta_{x,r,u}^\pm$ for the change in w during an event even though w is now defined on E only.)

Let us show that as $n \to \infty$, $M^{(n)}$ converges in distribution in $D_{\mathcal{M}_\lambda(E)}[0,\infty)$ to the unique solution $M^{(\infty)}$ to the $D_{\mathcal{M}_\lambda(E)}[0,\infty)$-martingale problem for $(\mathcal{L}^{(\infty)}, \delta_{M_E^0})$, where $\mathcal{L}^{(\infty)}$ is defined as in (2.4) with μ^n and $\mu'^{\,n}$ respectively replaced by μ and μ'. We check items $(a)-(d)$ one by one, with $\mathbf{L} = \mathcal{L}^{(\infty)}$ whose domain $\mathcal{D}(\mathcal{L}^{(\infty)})$ is taken to be the set of all functions of the form $\Psi_{F,f}$ with $f \in C(E)$ and $F \in C^1(\mathbb{R})$.

For item (a), observe that since every mapping w that we consider takes its values in $[0,1]$ (and so does its image by any $\Theta_{x,r,u}^\pm$) and E is compact, for every $f \in C(E)$ and every $x \in E$, $r > 0$ and $u \in [0,1]$, we have

$$\big|\langle\Theta_{x,r,u}^\pm(w),f\rangle - \langle w,f\rangle\big| \leq u\|f\|\mathrm{Vol}(B_E(x,r)). \qquad (2.5)$$

Consequently, for any $F \in C^1(\mathbb{R})$, by Taylor's theorem we have

$$\big|F(\langle\Theta_{x,r,u}^\pm(w),f\rangle) - F(\langle w,f\rangle)\big| \leq \left(\sup_{|z|\leq\|f\|\mathrm{Vol}(E)} |F'(z)|\right) u\|f\|C_d r^d, \qquad (2.6)$$

where C_d is a constant that depends only on the dimension d. Writing

$$\big|w(y)F(\langle\Theta_{x,r,u}^+(w),f\rangle) + (1-w(y))F(\langle\Theta_{x,r,u}^-(w),f\rangle) - F(\langle w,f\rangle)\big|$$
$$\leq w(y)\big|F(\langle\Theta_{x,r,u}^+(w),f\rangle) - F(\langle w,f\rangle)\big| + (1-w(y))\big|F(\langle\Theta_{x,r,u}^-(w),f\rangle) - F(\langle w,f\rangle)\big|$$
$$\leq \left(\sup_{|z|\leq\|f\|\mathrm{Vol}(E)} |F'(z)|\right) u\|f\|C_d r^d,$$

we obtain that

$$
\left| \mathcal{L}^{(\infty)} \Psi_{F,f}(M) \right| \leq \left(\sup_{|z| \leq \|f\|\mathrm{Vol}(E)} |F'(z)| \right) \|f\| C_d \mathrm{Vol}(E) \left(\int_0^\infty \int_0^1 u r^d \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \right.
$$
$$
\left. + \int_0^\infty \int_0^1 u r^d \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \right), \tag{2.7}
$$

and the quantity on the r.h.s. is a finite constant independent of M by Condition (1.7). This result proves that the operator $\mathcal{L}^{(\infty)}$ is indeed well defined on $\mathcal{D}(\mathcal{L}^{(\infty)})$. Since $\mathcal{M}_\lambda(E)$ is a compact subset of the set of all measures on $E \times \{0,1\}$ and since functions of the form $\Psi_{F,f}$ are continuous on the latter, each $\Psi_{F,f}$ belongs to $C_b(\mathcal{M}_\lambda(E))$. Recalling (1.5), we can rewrite $\mathcal{L}^{(\infty)} \Psi_{F,f}(M)$ in terms of the measure M as follows:

$$
\mathcal{L}^{(\infty)} \Psi_{F,f}(M) \tag{2.8}
$$
$$
= \int_E \int_0^\infty \int_0^1 \int_{B_E(x,r)\times\{0,1\}} \left\{ F\left(\int_{E\times\{0,1\}} f(z)\mathbb{1}_{\{0\}}(k)M(\mathrm{d}z,\mathrm{d}k) \right. \right.
$$
$$
- u \int_{B_E(x,r)\times\{0,1\}} f(z)\mathbb{1}_{\{0\}}(k)M(\mathrm{d}z,\mathrm{d}k) + u\mathbb{1}_{\{0\}}(\kappa) \left. \int_{B_E(x,r)} f(z)\mathrm{d}z \right)
$$
$$
- F\left(\int_{E\times\{0,1\}} f(z)\mathbb{1}_{\{0\}}(k)M(\mathrm{d}z,\mathrm{d}k) \right) \left. \right\} \frac{M(\mathrm{d}y,\mathrm{d}\kappa)}{\mathrm{Vol}(B_E(x,r))} \nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x
$$
$$
+ \int_E \int_0^\infty \int_0^1 \int_{(B_E(x,r)\times\{0,1\})^2} \left\{ F\left(\int_{E\times\{0,1\}} f(z)\mathbb{1}_{\{0\}}(k)M(\mathrm{d}z,\mathrm{d}k) \right. \right.
$$
$$
- u \int_{B_E(x,r)\times\{0,1\}} f(z)\mathbb{1}_{\{0\}}(k)M(\mathrm{d}z,\mathrm{d}k) + u\mathbb{1}_{\{0\}}(\kappa \vee \kappa') \left. \int_{B_E(x,r)} f(z)\mathrm{d}z \right)
$$
$$
- F\left(\int_{E\times\{0,1\}} f(z)\mathbb{1}_{\{0\}}(k)M(\mathrm{d}z,\mathrm{d}k) \right) \left. \right\} \frac{M^{\otimes 2}(\mathrm{d}y,\mathrm{d}\kappa,\mathrm{d}y',\mathrm{d}\kappa')}{\mathrm{Vol}(B_E(x,r))^2} \nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x.
$$

Using the Dominated Convergence Theorem and (2.7), it is then straightforward to check that if $(M_l)_{l\geq 0}$ is a sequence in $\mathcal{M}_\lambda(E)$ converging to M, then $\mathcal{L}^{(\infty)} \Psi_{F,f}(M_l)$ converges to $\mathcal{L}^{(\infty)} \Psi_{F,f}(M)$ as $l \to \infty$ and the function $\mathcal{L}^{(\infty)} \Psi_{F,f}$ is (sequentially) continuous on $\mathcal{M}_\lambda(E)$. Together with (2.7), this implies that $\mathcal{L}^{(\infty)} \Psi_{F,f} \in C_b(\mathcal{M}_\lambda(E))$ and item (a) is proved.

Item (b) is a consequence of the exact analogue of Proposition 1.7 in which the "geographical" space $\mathbb{R}^d$ is replaced by E (and $\mathcal{M}_\lambda$ by $\mathcal{M}_\lambda(E)$). Indeed, by Lemma 2.1(c) in [49], the linear span of the set of constant functions and of functions of the form

$$
M \mapsto \int_{(\mathbb{R}^d)^k} \psi(x_1,\ldots,x_k) \left\{ \prod_{j=1}^k w(x_j) \right\} \mathrm{d}x_1 \cdots \mathrm{d}x_k, \tag{2.9}
$$

(where M has density w) for $k \geq 1$ and $\psi \in \mathbb{L}^1((\mathbb{R}^d)^k) \cap C((\mathbb{R}^d)^k)$, is dense in the set of all continuous functions on the compact space $\mathcal{M}_\lambda(E)$ (and the same holds with $E = \mathbb{R}^d$). This set of functions is therefore separating on the space of all probability distributions on $\mathcal{M}_\lambda(E)$. We can then proceed exactly as in the proof of Proposition 4.4.7 in [21] to conclude that for every $M^0 \in \mathcal{M}_\lambda(E)$, uniqueness holds for the $B_{\mathcal{M}_\lambda}[0,\infty)$-martingale problem for $(\mathcal{L}^{(\infty)}, \delta_{M^0})$ (or more generally for any distribution for the initial value M_0). Indeed, in short (1.26) allows us to conclude that any two solutions to the martingale problem have the same one-dimensional distributions, and then Theorem 4.4.2 in [21] gives us that these two solutions necessarily have the same finite dimensional distributions and thus uniqueness in $B_{\mathcal{M}_\lambda}[0,\infty)$ holds. Item (b) is proved.

We now turn to item (c), the relative compactness of $(M^{(n)})_{n\geq 1}$. Since $\mathcal{M}_\lambda(E)$ (equipped with the topology of vague convergence) is a compact space and since by

Lemma 1.1 the set $\mathcal{D}(\mathcal{L}^{(\infty)})$ is dense in $C(\mathcal{M}_\lambda(E))$, by Theorem 3.9.1 in [21] the relative compactness of $(M^{(n)})_{n\geq 1}$ is equivalent to the relative compactness of the sequence of real-valued processes $(\Psi_{F,f}(M^{(n)}))_{n\geq 1}$ for all $\Psi_{F,f} \in \mathcal{D}(\mathcal{L}^{(\infty)})$. Thus let $F \in C^1(\mathbb{R})$ and $f \in C(E)$. Using the standard Aldous-Rebolledo criterion [2, 43] and writing $(\Phi_t^n)_{t\geq 0}$ for the predictable finite variation part of $\Psi_{F,f}(M^{(n)})$ and $(Q_t^n)_{t\geq 0}$ for the predictable quadratic variation of its martingale part, we only have to show that

(1) For every $t \geq 0$, the sequence $(\Psi_{F,f}(M_t^{(n)}))_{n\geq 1}$ is tight.

(2) For every $T > 0$, given a sequence of stopping times $(\tau_n)_{n\geq 1}$ bounded by T, for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} \sup_{\theta\in[0,\delta]} \mathbb{P}\left[\left|\Phi_{\tau_n+\theta}^n - \Phi_{\tau_n}^n\right| > \varepsilon\right] \leq \varepsilon, \tag{2.10}$$

and

$$\limsup_{n\to\infty} \sup_{\theta\in[0,\delta]} \mathbb{P}\left[\left|Q_{\tau_n+\theta}^n - Q_{\tau_n}^n\right| > \varepsilon\right] \leq \varepsilon. \tag{2.11}$$

(1) is straightforward, since for any potential density w we have $|\langle w, f\rangle| \leq \|f\|\mathrm{Vol}(E)$ and F is continuous, which implies that $\Psi_{F,f}(M_t^{(n)})$ is bounded uniformly in n and t. To deal with (2), for every time t we fix a representative $w_t^{(n)}$ of the density of $M_t^{(n)}$. Since each $(M_t^{(n)})_{t\geq 0}$ is a Markov jump process with bounded jump rates, we have

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

and

$$Q_t^n = \int_0^t \int_E \int_0^\infty \int_0^1 \int_{B_E(x,r)} \frac{1}{V(B_E(x,r))} \left\{ w_s^{(n)}(y)\left[F(\langle\Theta_{x,r,u}^+(w_s^{(n)}), f\rangle) - F(\langle w_s^{(n)}, f\rangle)\right]^2 \right.$$
$$\left. + (1 - w_s^{(n)}(y))\left[F(\langle\Theta_{x,r,u}^-(w_s^{(n)}), f\rangle) - F(\langle w_s^{(n)}, f\rangle)\right]^2 \right\} \mathrm{d}y\,\nu_r(\mathrm{d}u)\,\mu^n(\mathrm{d}r)\,\mathrm{d}x\,\mathrm{d}s$$
$$+ \int_0^t \int_E \int_0^\infty \int_0^1 \int_{B_E(x,r)^2} \frac{1}{V(B_E(x,r))^2} \left\{ w_s^{(n)}(y)w_s^{(n)}(z)\left[F(\langle\Theta_{x,r,u}^+(w_s^{(n)}), f\rangle)\right.\right.$$
$$\left. - F(\langle w_s^{(n)}, f\rangle)\right]^2$$
$$+ (1 - w_s^{(n)}(y)w_s^{(n)}(z))\left[F(\langle\Theta_{x,r,u}^-(w_s^{(n)}), f\rangle) - F(\langle w_s^{(n)}, f\rangle)\right]^2 \right\} \mathrm{d}y\,\mathrm{d}z\,\nu_r'(\mathrm{d}u)\,\mu'^n(\mathrm{d}r)\,\mathrm{d}x\,\mathrm{d}s.$$

Using the expression for $\mathcal{L}^{(n)}$ given in (2.4) and the bound (2.6), as in (2.7) we obtain that for every $M \in \mathcal{M}_\lambda(E)$,

$$\left|\mathcal{L}^{(n)}\Psi_{F,f}(M)\right| \leq C\left(\sup_{|z|\leq\|f\|\mathrm{Vol}(E)} |F'(z)|\right)\|f\|\left(\int_E \int_0^\infty \int_0^1 ur^d\nu_r(\mathrm{d}u)\mu^n(\mathrm{d}r)\mathrm{d}x\right.$$
$$\left. + \int_E \int_0^\infty \int_0^1 ur^d\nu_r'(\mathrm{d}u)\mu'^n(\mathrm{d}r)\mathrm{d}x\right)$$
$$\leq C'\left(\int_0^\infty \int_0^1 ur^d\nu_r(\mathrm{d}u)\mu(\mathrm{d}r) + \int_0^\infty \int_0^1 ur^d\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\right), \tag{2.12}$$

where we have used that E has finite volume and, by assumption, $\mu^n(\mathrm{d}r) \nearrow \mu(\mathrm{d}r)$ (and the corresponding statement with primes). By Condition (1.7), the expression on the r.h.s. is finite (and independent of M), and so with probability 1 we have for every $\theta > 0$

$$\left|\Phi_{\tau_n+\theta}^n - \Phi_{\tau_n}^n\right| \leq C''\theta,$$

where C'' is independent of n (and even of T). Hence, it suffices to choose $\delta > 0$ small enough for (2.10) to hold.

Similarly, the integrands in the expression of Q_t^n are bounded by

$$C'' \left(\sup_{|z| \leq \|f\| \mathrm{Vol}(E)} |F'(z)| \right)^2 \|f\|^2 u^2 \min(r^{2d}, \ell^{2d}), \tag{2.13}$$

where ℓ is the sidelength of E. But $u^2 \leq u$ and there exists a constant $\mathcal{C}_E < \infty$ such that $\min(r^{2d}, \ell^{2d}) \leq \mathcal{C}_E r^d$, and so the same reasoning shows that (2.11) holds too for $\delta > 0$ small enough. The relative compactness of every sequence $(\Psi_{F,f}(M^{(n)}))$ is proved, and the relative compactness of $(M^{(n)})_{n \geq 1}$ in $D_{\mathcal{M}_\lambda(E)}[0, \infty)$ thus follows.

Finally, we prove item (d). Let $F \in C^1(\mathbb{R})$ and $f \in C(E)$. Let also $k \geq 0$, $0 \leq t_1 < t_2 < \cdots < t_k \leq t < t + s$ and $h_i \in C_b(\mathcal{M}_\lambda(E))$ for all $i \in \{1, \ldots, k\}$. Since $M^{(n)}$ satisfies the martingale problem for $(\mathcal{L}^{(n)}, \delta_{M_E^0})$, we have

$$\mathbb{E}\left[\left(\Psi_{F,f}\big(M_{t+s}^{(n)}\big) - \Psi_{F,f}\big(M_t^{(n)}\big) - \int_t^{t+s} \mathcal{L}^{(n)} \Psi_{F,f}\big(M_u^{(n)}\big) \mathrm{d}u \right) \left(\prod_{i=1}^k h_i\big(M_{t_i}^{(n)}\big) \right) \right] = 0. \tag{2.14}$$

We can thus write

$$\mathbb{E}\left[\left(\Psi_{F,f}\big(M_{t+s}^{(n)}\big) - \Psi_{F,f}\big(M_t^{(n)}\big) - \int_t^{t+s} \mathcal{L}^{(\infty)} \Psi_{F,f}\big(M_u^{(n)}\big) \mathrm{d}u \right) \left(\prod_{i=1}^k h_i\big(M_{t_i}^{(n)}\big) \right) \right]$$
$$= 0 + \mathbb{E}\left[\left(\int_t^{t+s} \left(\mathcal{L}^{(n)} \Psi_{F,f}\big(M_u^{(n)}\big) - \mathcal{L}^{(\infty)} \Psi_{F,f}\big(M_u^{(n)}\big) \right) \mathrm{d}u \right) \left(\prod_{i=1}^k h_i\big(M_{t_i}^{(n)}\big) \right) \right]. \tag{2.15}$$

Using (2.2) and applying the estimate (2.12) with the positive measures $\mu(\mathrm{d}r)\nu_r(\mathrm{d}u) - \mu^n(\mathrm{d}r)\nu_r(\mathrm{d}u)$ and $\mu'(\mathrm{d}r)\nu'_r(\mathrm{d}u) - \mu'^n(\mathrm{d}r)\nu'_r(\mathrm{d}u)$, we see that we can use the Dominated Convergence Theorem to argue that the second term on the r.h.s. of (2.15) converges to 0 as $n \to \infty$. Hence, (2.1) holds true with $\Gamma = \emptyset$ and item (d) is proved.

As detailed at the beginning of this section, we can therefore conclude that there exists a unique solution $M^{(\infty)}$ to the $B_{\mathcal{M}_\lambda(E)}[0, \infty)$-martingale problem for $(\mathcal{L}^{(\infty)}, \delta_{M_E^0})$ and this process is Markov with respect to its natural filtration and has càdlàg paths a.s.

Concerning the Feller property of the semigroup of $M^{(\infty)}$, the fact that for every $t \geq 0$ and every $\varphi \in C(\mathcal{M}_\lambda(E))$, $M \mapsto \mathbb{E}_M[\varphi(M_t^{(\infty)})]$ is a continuous function is a consequence of the continuity in M^0 of the quantity on the r.h.s. of (1.26) (which is more easily seen in (1.23) when we replace ψ by the density at time t – conditional on N_t – of the locations of the atoms $\xi_t^1, \ldots, \xi_t^{N_t}$ ordered in some arbitrary way) and the property already mentioned in item (b) that the linear span of the set of constant functions and of functions of the form (2.9) is dense in $C(\mathcal{M}_\lambda(E))$. The strong continuity of the semigroup is a consequence of the fact, proved in item (d), that for every $\Psi_{F,f} \in \mathcal{D}(\mathcal{L}^{(\infty)})$, we have for every $t \geq 0$ and every $M \in \mathcal{M}_\lambda(E)$

$$\mathbb{E}_M\big[\Psi_{F,f}\big(M_t^{(\infty)}\big)\big] - \Psi_{F,f}(M) = \mathbb{E}_M\left[\int_0^t \mathcal{L}^{(\infty)} \Psi_{F,f}\big(M_s^{(\infty)}\big) \mathrm{d}s \right].$$

Together with the uniform bound (2.7), it shows that there exists $C_{F,f} > 0$ such that for every $t \geq 0$

$$\sup_{M \in \mathcal{M}_\lambda(E)} \big| \mathbb{E}_M\big[\Psi_{F,f}\big(M_t^{(\infty)}\big)\big] - \Psi_{F,f}(M) \big| \leq C_{F,f} t,$$

and the quantity on the l.h.s. indeed converges to 0 as $t \to 0$. This property can then be extended to any $\varphi \in C(\mathcal{M}_\lambda(E))$ by Lemma 1.1.

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 Q^n becomes more delicate. The resolution is to exploit the fact that f has compact support S_f.

Let μ, μ', ν and ν' satisfy (1.7) and let $\{E_n\}_{n \geq 1}$ be a sequence of hypercubes increasing to $\mathbb{R}^d$. We embed each $\mathcal{M}_\lambda(E_n)$ into $\mathcal{M}_\lambda = \mathcal{M}_\lambda(\mathbb{R}^d)$ by setting $w(x) \equiv 0$ outside E_n. For every $n \geq 1$, let $M^{[n]}$ denote the $\mathcal{M}_\lambda$-valued Markov process obtained by imposing that $M^{[n]}\big|_{E_n \times \{0,1\}}$ should evolve like the SLFVS on $E_n \times \{0,1\}$ obtained in (i), and $M^{[n]}\big|_{E_n^c \times \{0,1\}} = \mathrm{d}x\big|_{E_n^c} \otimes \delta_1(d\kappa)$ (*i.e.*, $w^{[n]} \equiv 0$ on E_n^c). For each $n \in \mathbb{N}$, $M^{[n]}$ is an a.s. càdlàg process and we assume that it starts from the measure $M_{E_n}^0$ obtained by restricting M^0 to E_n (as in (i)) and by assuming that its "density" $w_{E_n}^0$ is 0 outside E_n (obviously, $M_{E_n}^0$ converges vaguely to M^0 as $n \to \infty$). According to the previous paragraph, it satisfies the property that for every $F \in C^1(\mathbb{R})$ and every $f \in C_c(\mathbb{R}^d)$,

$$\left(\Psi_{F,f}\big(M_t^{[n]}\big) - \Psi_{F,f}\big(M_0^{[n]}\big) - \int_0^t \mathcal{L}^{[n]}\Psi_{F,f}\big(M_s^{[n]}\big)\mathrm{d}s \right)_{t \geq 0} \tag{2.16}$$

is a martingale, where

$$
\begin{aligned}
&\mathcal{L}^{[n]}\Psi_{F,f}(M) \\
&= \int_0^\infty \int_{(S_f + B(0,r)) \cap E_n} \int_0^1 \int_{B_{E_n}(x,r)} \frac{1}{\mathrm{Vol}(B_{E_n}(x,r))}\Big[w(y)F(\langle \Theta_{n,x,r,u}^+(w), f \rangle) \\
&\qquad\qquad + (1 - w(y))F(\langle \Theta_{n,x,r,u}^-(w), f \rangle) - F(\langle w, f \rangle) \Big] \mathrm{d}y\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \\
&+ \int_0^\infty \int_{(S_f + B(0,r)) \cap E_n} \int_0^1 \int_{B_{E_n}(x,r)^2} \frac{1}{\mathrm{Vol}(B_{E_n}(x,r))^2}\Big[w(y)w(z)F\big(\langle \Theta_{n,x,r,u}^+(w), f \rangle\big) \\
&\qquad\qquad + (1 - w(y)w(z))F\big(\langle \Theta_{n,x,r,u}^-(w), f \rangle\big) - F(\langle w, f \rangle) \Big] \mathrm{d}y\, \mathrm{d}z\, \nu_r'(\mathrm{d}u)\, \mathrm{d}x\, \mu'(\mathrm{d}r). \tag{2.17}
\end{aligned}
$$

Here we have written $S_f + B(0,r) := \{x + y \,:\, x \in S_f, y \in B(0,r)\}$ (motivated by the fact that if the centre of an event of radius r does not belong to this set, then the event does not intersect the support of f and therefore it does not affect the value of $\Psi_{F,f}(M)$) and we have chosen to report the dependence of the operations $\Theta_{n,x,r,u}^{\pm}$ on n since they modify the value of w only within E_n. The key observation is that

$$|\langle \mathbb{1}_{B(x,r)}w, f \rangle| \leq \|f\|\mathrm{Vol}(S_f \cap B(x,r)) \leq C_1\|f\|(r^d \wedge 1), \tag{2.18}$$

and

$$\mathrm{Vol}(S_f + B(0,r)) \leq C_2(r^d \vee 1), \tag{2.19}$$

where C_1 and C_2 are independent of r and depend only on the support of f. Moreover, the estimate (2.18) 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 part of (2.17) corresponding to neutral events. We split the integral over $(0, \infty)$ at some radius $R_0 > 1$. We have that

$$
\left| \int_{R_0}^{\infty} \int_{(S_f + B(0,r)) \cap E_n} \int_0^1 \int_{B_{E_n}(x,r)} \frac{1}{\mathrm{Vol}(B_{E_n}(x,r))} \Big[w(y) F(\langle \Theta_{n,x,r,u}^+(w), f \rangle) \right.
$$
$$
\left. + (1 - w(y)) F(\langle \Theta_{n,x,r,u}^-(w), f \rangle) - F(\langle w, f \rangle) \Big] \, \mathrm{d}y \, \nu_r(\mathrm{d}u) \, \mathrm{d}x \, \mu(\mathrm{d}r) \right|
$$
$$
\leq C_3(F', f) \int_{R_0}^{\infty} \int_{(S_f + B(0,r)) \cap E_n} \int_0^1 u \mathrm{Vol}(B(x,r) \cap S_f) \, \nu_r(\mathrm{d}u) \, \mathrm{d}x \, \mu(\mathrm{d}r)
$$
$$
\leq C_4(F', f) \mathrm{Vol}(S_f) \int_{R_0}^{\infty} \int_0^1 u r^d \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r), \tag{2.20}
$$

where $C_3(F', f)$ and $C_4(F', f)$ depend only on F' and f and the last line uses (2.19) and the fact that $\mathrm{Vol}(B(x,r) \cap S_f) \leq \mathrm{Vol}(S_f)$. 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

$$
\left| \int_0^{R_0} \int_{(S_f + B(0,r)) \cap E_n} \int_0^1 \int_{B_{E_n}(x,r)} \frac{1}{\mathrm{Vol}(B_{E_n}(x,r))} \Big[w(y) F(\langle \Theta_{n,x,r,u}^+(w), f \rangle) \right.
$$
$$
\left. + (1 - w(y)) F(\langle \Theta_{n,x,r,u}^-(w), f \rangle) - F(\langle w, f \rangle) \Big] \, \mathrm{d}y \, \nu_r(\mathrm{d}u) \, \mathrm{d}x \, \mu(\mathrm{d}r) \right|
$$
$$
\leq C_3(F', f) \int_0^{R_0} \int_{(S_f + B(0,r)) \cap E_n} \int_0^1 u \mathrm{Vol}(B(x,r) \cap S_f) \, \nu_r(\mathrm{d}u) \, \mathrm{d}x \, \mu(\mathrm{d}r)
$$
$$
\leq C_5(F', f) \mathrm{Vol}(S_f + B(0, R_0)) \int_0^{R_0} \int_0^1 u r^d \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r), \tag{2.21}
$$

where we have used (2.18) to bound $\mathrm{Vol}(B(x,r) \cap S_f)$ by Cr^d (independently of x), and then we have bounded the remaining integral of $\mathrm{d}x$ over $(S_f + B(0,r)) \cap E_n$ by $\mathrm{Vol}(S_f + B(0, R_0))$. Observe that both bounds (2.20) and (2.21) are finite by Condition (1.7), and they are independent of M (or w). Exactly the same arguments control the selection part of the generator $\mathcal{L}^{[n]}$. Furthermore, the same bounds apply if we replace the operator $\mathcal{L}^{[n]}$ by $\mathcal{L}$ defined in (1.9). Consequently, we can proceed as in (i) to prove that $\mathcal{L}$ is included in $C_b(\mathcal{M}_\lambda) \times C_b(\mathcal{M}_\lambda)$, which was item (a) to check. Item (b) is a direct consequence of Proposition 1.7 and of the same arguments as in the analogous part of the proof of (i). The fact that each $M^{[n]}$ is a process with sample paths in $D_{\mathcal{M}_\lambda}[0, \infty)$ is part of the conclusion of (i). Next, the estimates (2.20) and (2.21) enable us to proceed as in the proof of item (c) for (i) to show that the sequence $(M^{[n]})_{n \geq 1}$ is relatively compact, which proves item (c) for (ii).

To check item (d), notice that by Condition (1.7), by taking R_0 sufficiently large, the right hand side of (2.20) can be made arbitrarily small, uniformly in M. This is enough to ensure that the missing contribution of the events centered *outside* E_n is negligible, that is that

$$
\left| \int_0^{\infty} \int_{(S_f + B(0,r)) \cap E_n^c} \int_0^1 \int_{B(x,r)} \frac{1}{\mathrm{Vol}(B(x,r))} \Big[w(y) F(\langle \Theta_{x,r,u}^+(w), f \rangle) \right.
$$
$$
\left. + (1 - w(y)) F(\langle \Theta_{x,r,u}^-(w), f \rangle) - F(\langle w, f \rangle) \Big] \, \mathrm{d}y \, \nu_r(\mathrm{d}u) \, \mathrm{d}x \, \mu(\mathrm{d}r) \right|
$$
$$
\leq C \mathrm{Vol}(S_f) \int_{d(S_f, E_n^c)}^{\infty} \int_0^1 u r^d \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \to 0 \tag{2.22}
$$

uniformly in M (or w) as $n \to \infty$, where $d(S_f, E_n^c)$ is the minimal distance between a point of S_f and a point of E_n^c (which tends to infinity as n tends to infinity). The same estimates hold for the selection term, and can also be used to control the error due to the vanishing difference between $\langle \Theta_{x,r,u}^{\pm}(w), f \rangle$ and $\langle \Theta_{n,x,r,u}^{\pm}(w), f \rangle$ (the latter modifying w on $B(x,r) \cap E_n$ only). We can then argue as in (2.15) to conclude.

Since items $(a) - (d)$ are now checked, we can conclude that there exists a unique measurable process $(M_t)_{t \geq 0}$ such that $M_0 = M^0$ and satisfying the martingale problem (1.11), and this process is Markov and has càdlàg paths a.s. The Feller property of its semigroup can then be derived using the same arguments as in the proof of (i), recalling the uniform bound on $\mathcal{L}$ obtained in (2.20) and (2.21) (see the paragraph following (2.21)).

3 Proof of Proposition 1.7 (duality)

To prove Proposition 1.7, we first show that we can extend the operator $\mathcal{L}$ to a larger class of functions on $\mathcal{M}_\lambda$ and that any solution to the $B_{\mathcal{M}_\lambda}[0,\infty)$-martingale problem for $(\mathcal{L}, \delta_{M^0})$ stated in (1.11) satisfies the corresponding extended martingale problem. We then use this new set of test functions to complete the proof of Proposition 1.7.

For every $k \in \{1, 2, \ldots\}$ and $\psi \in \mathbb{L}^1((\mathbb{R}^d)^k)$, let us define the function D_ψ by: for every $M \in \mathcal{M}_\lambda$ with density w,

$$
\begin{aligned}
D_\psi(M) &:= \int_{(\mathbb{R}^d \times \{0,1\})^k} \psi(x_1, \ldots, x_k) \left\{ \prod_{j=1}^k \mathbb{1}_{\{0\}}(\kappa_j) \right\} M(\mathrm{d}x_1, \mathrm{d}\kappa_1) \cdots M(\mathrm{d}x_k, \mathrm{d}\kappa_k) \\
&= \int_{(\mathbb{R}^d)^k} \psi(x_1, \ldots, x_k) \left\{ \prod_{j=1}^k w(x_j) \right\} \mathrm{d}x_1 \cdots \mathrm{d}x_k.
\end{aligned}
\tag{3.1}
$$

Let us also set

$$
\begin{aligned}
&\mathcal{L}D_\psi(M) \\
&:= \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \int_{(\mathbb{R}^d)^k} \psi(x_1, \ldots, x_k) \left\{ \prod_{j \in I^c} w(x_j) \right\} \left[w(y) \prod_{j \in I} ((1-u)w(x_j) + u) \right. \\
&\quad \left. + (1 - w(y)) \prod_{j \in I} ((1-u)w(x_j)) - \prod_{j \in I} w(x_j) \right] \mathrm{d}x_1 \ldots \mathrm{d}x_k \mathrm{d}y \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)^2} \frac{1}{V_r^2} \int_{(\mathbb{R}^d)^k} \psi(x_1, \ldots, x_k) \left\{ \prod_{j \in I^c} w(x_j) \right\} \\
&\quad \times \left[w(y)w(z) \prod_{j \in I} ((1-u)w(x_j) + u) + (1 - w(y)w(z)) \prod_{j \in I} ((1-u)w(x_j)) - \prod_{j \in I} w(x_j) \right] \\
&\qquad \mathrm{d}x_1 \ldots \mathrm{d}x_k \mathrm{d}y \mathrm{d}z \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \mathrm{d}x,
\end{aligned}
\tag{3.2}
$$

where we use the notation $I := \{i : x_i \in B(x,r)\}$ and the convention that a product over the empty set is equal to 1. Note that for every $\psi \in C_c(\mathbb{R}^d)$, the function D_ψ coincides with the function $\Psi_{Id,\psi}$ defined in (1.5) and, likewise, the function $\mathcal{L}D_\psi$ defined in (3.2) coincides with the function $\mathcal{L}\Psi_{Id,\psi}$ defined in (1.9). To see this, let us observe that the first part of (3.2) can be rewritten (using the convention that the product over an empty

set is equal to 1, and then Fubini's theorem to pass from the first line to the next)

$$
\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \int_{\mathbb{R}^d} \psi(x_1) \mathbb{1}_{\{x_1 \in B(x,r)\}} \big[w(y)\big((1-u)w(x_1) + u\big)
$$
$$
+ (1 - w(y))(1-u)w(x_1) - w(x_1) \big] \, dx_1 dy \nu_r(du) \mu(dr) dx
$$
$$
= \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \Big\{ w(y) \langle \mathbb{1}_{B(x,r)}\big((1-u)w + u\big), \psi \rangle + (1 - w(y)) \langle \mathbb{1}_{B(x,r)}(1-u)w, \psi \rangle
$$
$$
- \langle \mathbb{1}_{B(x,r)} w, \psi \rangle \Big\} \, dy \nu_r(du) \mu(dr) dx,
$$

which is equal to the first integral on the r.h.s. of (1.9) when $F = Id$ and $f = \psi$. The same reasoning can be made on the second part of (3.2).

The following result shows that the expression in (3.2) in fact extends the operator $\mathcal{L}$ defined in (1.9) to all functions of the form D_ψ.

Lemma 3.1. *Let $(M_t)_{t \geq 0}$ be a solution to the martingale problem (1.11). Then for every $k \geq 1$ and $\psi \in \mathbb{L}^1((\mathbb{R}^d)^k)$,*

$$
\left(D_\psi(M_t) - D_\psi(M_0) - \int_0^t \mathcal{L} D_\psi(M_s) ds \right)_{t \geq 0} \tag{3.3}
$$

is a martingale.

Proof of Lemma 3.1. We have already checked that the desired property held true for $k = 1$ and $\psi \in C_c(\mathbb{R}^d)$, since then $D_\psi = \Psi_{Id,\psi}$ and $\mathcal{L} D_\psi = \mathcal{L} \Psi_{Id,\psi}$. We first extend the result to $\psi \in \mathbb{L}^1(\mathbb{R}^d)$ by a density argument, and then consider the case $k \geq 2$.

To complete both parts of the programme, we need a general bound on functions of the form $\mathcal{L} D_\psi$ that we derive now. Let $m \geq 1$ and $\widetilde{\psi} \in \mathbb{L}^1((\mathbb{R}^d)^m)$. For every $M \in \mathcal{M}_\lambda$ (with density w), the expression for $\mathcal{L} D_{\widetilde{\psi}}(M)$ given in (3.2) can be rewritten as

$$
\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \int_{(\mathbb{R}^d)^m} \widetilde{\psi}(x_1, \ldots, x_m) \bigg\{ \prod_{j \in I^c} w(x_j) \bigg\}
$$
$$
\times \bigg[\bigg(\sum_{J \subset I, J \neq I} (1-u)^{|J|} u^{|I \setminus J|} \prod_{j \in J} w(x_j) \bigg) w(y) + \big((1-u)^{|I|} - 1\big) \prod_{j \in I} w(x_j) \bigg]
$$
$$
dx_1 \ldots dx_m dy \nu_r(du) \mu(dr) dx
$$
$$
+ \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)^2} \frac{1}{V_r^2} \int_{(\mathbb{R}^d)^m} \widetilde{\psi}(x_1, \ldots, x_m) \bigg\{ \prod_{j \in I^c} w(x_j) \bigg\}
$$
$$
\times \bigg[w(y)w(z) \bigg(\sum_{J \subset I, J \neq I} (1-u)^{|J|} u^{|I \setminus J|} \prod_{j \in J} w(x_j) \bigg) + \big((1-u)^{|I|} - 1\big) \prod_{j \in I} w(x_j) \bigg]
$$
$$
dx_1 \ldots dx_k dy dz \nu_r'(du) \mu'(dr) dx.
$$

Bounding w by 1 and using the facts that

$$
\sum_{J \subsetneq I} (1-u)^{|J|} u^{|I \setminus J|} = 1 - (1-u)^{|I|} \leq u|I| \mathbb{1}_{\{|I| \geq 1\}} \leq m u \mathbb{1}_{\{B(x,r) \cap \{x_1, \ldots, x_m\} \neq \emptyset\}}
$$

and

$$
\mathrm{Vol}\big(\{x \in \mathbb{R}^d : B(x,r) \cap \{x_1, \ldots, x_m\} \neq \emptyset\}\big) \leq m C_d r^d
$$

for a constant C_d depending only on d, we obtain that the first term in $|\mathcal{L}D_{\widetilde{\psi}}(M)|$ is bounded by

$$\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\int_{(\mathbb{R}^d)^m}|\widetilde{\psi}|(x_1,\ldots,x_m)\mathbb{1}_{\{B(x,r)\cap\{x_1,\ldots,x_m\}\neq\emptyset\}}2mu$$
$$\mathrm{d}x_1\cdots\mathrm{d}x_m\mathrm{d}y\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x$$
$$\leq\int_0^\infty\int_0^1\int_{(\mathbb{R}^d)^m}|\widetilde{\psi}|(x_1,\ldots,x_m)2m^2C_dr^du\,\mathrm{d}x_1\cdots\mathrm{d}x_m\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)$$
$$=2m^2C_d\|\widetilde{\psi}\|_1\int_0^\infty\int_0^1 r^du\,\nu_r(\mathrm{d}u)\mu(\mathrm{d}r). \tag{3.4}$$

We can then bound the second part of $|\mathcal{L}D_{\widetilde{\psi}}(M)|$ in a similar way and obtain that

$$|\mathcal{L}D_{\widetilde{\psi}}(M)|\leq 2m^2C_d\|\widetilde{\psi}\|_1\left(\int_0^\infty\int_0^1 r^du\,\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)+\int_0^\infty\int_0^1 r^du\,\nu'_r(\mathrm{d}u)\mu'(\mathrm{d}r)\right), \tag{3.5}$$

and the expression on the r.h.s. is finite by Condition (1.7).

Using the fact that $C_c(\mathbb{R}^d)$ is dense in $\mathbb{L}^1(\mathbb{R}^d)$ (for the $\mathbb{L}^1$ norm), the bound (3.5) and dominated convergence, we can then use the same approach as in the proof of items (d) in Section 2 (see Equations (2.14) and (2.15)) to conclude that the process in (3.3) is indeed a martingale when $\psi\in\mathbb{L}^1(\mathbb{R}^d)$.

Let us now consider $k\geq 2$. Any integrable function ψ on $(\mathbb{R}^d)^k$ can be approximated (in $\mathbb{L}^1$ norm) by linear combinations of functions of the product form $\psi_1(x_1)\cdots\psi_k(x_k)$ with $\psi_i\in C_c(\mathbb{R}^d)$ for every i. Furthermore, by polarisation, the test function

$$D_{\otimes\psi_i}(M)=\prod_{i=1}^k\left(\int_{\mathbb{R}^d\times\{0,1\}}\psi_i(x_i)\mathbb{1}_{\{0\}}(\kappa_i)M(\mathrm{d}x_i,\mathrm{d}\kappa_i)\right)=\prod_{i=1}^k\langle w,\psi_i\rangle \tag{3.6}$$

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 $f\in C_c(\mathbb{R}^d)$, for which we can use (1.9) to obtain

$$\mathcal{L}\Psi_{(\cdot)^m,f}(M)=\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\left[w(y)\langle\mathbb{1}_{B(x,r)^c}w+\mathbb{1}_{B(x,r)}((1-u)w+u),f\rangle^m\right.$$
$$+(1-w(y))\langle\mathbb{1}_{B(x,r)^c}w+\mathbb{1}_{B(x,r)}(1-u)w,f\rangle^m-\langle w,f\rangle^m\Big]\mathrm{d}y\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x$$
$$+\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_r^2}\left[w(y)w(z)\langle\mathbb{1}_{B(x,r)^c}w+\mathbb{1}_{B(x,r)}((1-u)w+u),f\rangle^m\right.$$
$$+(1-w(y)w(z))\langle\mathbb{1}_{B(x,r)^c}w+\mathbb{1}_{B(x,r)}(1-u)w,f\rangle^m-\langle w,f\rangle^m\Big]$$
$$\mathrm{d}y\,\mathrm{d}z\nu'_r(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x. \tag{3.7}$$

On the other hand, taking $\psi(x_1,\ldots,x_m)=\prod_{i=1}^m f(x_i)$ in (3.1), we obtain that

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

Let us thus show that, in this case, the expression on the r.h.s. of (3.2) coincides with (3.7). We focus on the first term on the r.h.s. of (3.2), 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

$$
\int_{(\mathbb{R}^d)^m} f(x_1) \cdots f(x_m) \bigg[\prod_{j:x_j \notin B} w(x_i) \bigg] \bigg[\prod_{j:x_j \in B} ((1-u)w(x_j) + u) \bigg] \mathrm{d}x_1 \ldots \mathrm{d}x_m
$$

$$
= \sum_{J \subseteq \{1,\ldots,m\}} \int_{(\mathbb{R}^d)^m} f(x_1) \cdots f(x_m) \bigg[\prod_{j \in J^c} \mathbb{1}_{\{x_j \notin B\}} w(x_j) \bigg] \bigg[\prod_{j \in J} \mathbb{1}_{\{x_j \in B\}} ((1-u)w(x_j) + u) \bigg]
$$

$$
\mathrm{d}x_1 \ldots \mathrm{d}x_m
$$

$$
= \sum_{J \subseteq \{1,\ldots,m\}} \big\langle \mathbb{1}_{B^c} w, f \big\rangle^{m-|J|} \big\langle \mathbb{1}_B ((1-u)w + u), f \big\rangle^{|J|}
$$

$$
= \sum_{j=0}^m \binom{m}{j} \big\langle \mathbb{1}_{B^c} w, f \big\rangle^{m-j} \big\langle \mathbb{1}_B ((1-u)w + u), f \big\rangle^j = \big\langle \mathbb{1}_{B^c} w + \mathbb{1}_B \big((1-u)w + u \big), f \big\rangle^m,
$$

which coincides with the integrand in the first part of (3.7). Checking that the same holds for the three other parts of (3.7), we can conclude that the two expressions for the action of $\mathcal{L}$ on functions of the form $\langle w, f \rangle^m$ coincide. Consequently, the process in (3.3) is a martingale for ψ form $f \otimes \cdots \otimes f$ with $f \in C_c(\mathbb{R}^d)$, and by linearity for ψ of the form $\psi_1 \otimes \cdots \otimes \psi_k$ with $\psi_i \in C_c(\mathbb{R}^d)$ for every i. Using the same density argument as in the case $k = 1$, together with the bound (3.5), we can finally conclude that the process in (3.3) is a martingale for every $k \geq 1$ and every $\psi \in \mathbb{L}^1((\mathbb{R}^d)^k)$, and Lemma 3.1 is proved. $\qquad \square$

Remark 3.2. Note that either of these two sets of test functions, (1.5) or (3.1), is sufficient to characterise the law of the SLFVS (see Lemma 1.1 for the first set, and Lemma 2.1(c) in [49] for the second), and so we can use them interchangeably. In particular, the family (1.5) 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 (3.1).

Armed with Lemma 3.1, we can now prove Proposition 1.7.

Proof of Proposition 1.7. Despite the fact that Theorem 4.4.11 in [21] does not directly apply, we follow its proof closely (with $\alpha \equiv 0 \equiv \beta$). Let $(M_t)_{t \geq 0}$ be solution to the martingale problem (1.11). For every $s, t \geq 0$, let

$$
F(s,t) := \mathbb{E}_{M^0} \Big[\mathbf{E} \big[D(M_s, \Xi_t) \,|\, \Xi_0 \sim \mu_\psi \big] \Big].
$$

By Lemma 1.6, since Ξ_0 has law μ_ψ, at every time $t \geq 0$ the locations of the atoms of Ξ_t have a joint distribution which is absolutely continuous with respect to Lebesgue measure. Let us write $\psi_t^{(n)}$ for the density of these points conditionally on the event $\{N_t = n\}$. We thus have, by (1.23),

$$
F(s,t) = \mathbb{E}_{M^0} \Bigg[\mathbf{E} \bigg[\int_{(\mathbb{R}^d)^{N_t}} \psi_t^{(N_t)}(x_1, \ldots, x_{N_t}) \bigg(\prod_{j=1}^{N_t} w_s(x_j) \bigg) \mathrm{d}x_1 \cdots \mathrm{d}x_{N_t} \,\bigg|\, \Xi_0 \sim \mu_\psi \bigg] \Bigg] \quad (3.8)
$$

$$
= \mathbf{E} \Big[\mathbb{E}_{M^0} \big[D_{\psi_t^{(N_t)}}(M_s) \big] \,\big|\, \Xi_0 \sim \mu_\psi \Big],
$$

where the last line uses Fubini's theorem. Since for every $n \in \{1, 2, \ldots\}$ we have $\psi_t^{(n)} \in \mathbb{L}^1((\mathbb{R}^d)^n)$, we can use Lemma 3.1 and write that

$$
F(s,t) - F(0,t) = \mathbf{E} \Bigg[\mathbb{E}_{M^0} \bigg[\int_0^s \mathcal{L} D_{\psi_t^{(N_t)}}(M_\tau) \mathrm{d}\tau \bigg] \,\bigg|\, \Xi_0 \sim \mu_\psi \Bigg]. \quad (3.9)
$$

On the other hand, for any fixed $s \geq 0$ and any $t \geq 0$, we can rewrite the expression on the r.h.s. of (3.8) as

$$\mathbb{E}_{M^0}\Bigg[\mathbf{E}\bigg[\prod_{j=1}^{N_t} w_s\big(\xi_t^j\big)\,\bigg|\,\Xi_0 \sim \mu_\psi\bigg]\Bigg] = \mathbb{E}_{M^0}\Big[\mathbf{E}\big[\exp\big(\langle \Xi_t, \ln w_s\rangle\big)\,\big|\,\Xi_0 \sim \mu_\psi\big]\Big]$$

$$= \mathbb{E}_{M^0}\Big[\mathbf{E}\big[\Phi_{\exp,\ln w_s}(\Xi_t)\,\big|\,\Xi_0 \sim \mu_\psi\big]\Big], \qquad (3.10)$$

where we have fixed a representative w_s of the density of M_s (since by (1.23), the r.h.s. of (3.10) is independent of the choice of this representative) and $\Phi_{\exp,\ln w_s}$ is defined as in (1.13) with $F = \exp$ and $f = \ln w_s$. Here we use the convention that $\Phi_{\exp,\ln w_s}(\Xi) = 0$ whenever at least one of the atoms x of Ξ is such that $w_s(x) = 0$. With this convention, the definition of $\mathcal{G}\Phi_{\exp,\ln w_s}(\Xi)$ given in (1.14) still makes sense, the function $\Phi_{\exp,\ln w_s}$ takes its values in $[0,1]$ and the same bound as in (1.19) controls the expectation of $\mathcal{G}\Phi_{\exp,\ln w_s}(\Xi_\tau)$ for every $\tau \in [0,t]$ (since in this case $F'(\langle \Xi_\tau, f\rangle) = \prod_i w_s(\xi_\tau^i) \in [0,1]$). We may therefore use Proposition 1.5 to write that

$$F(s,t) - F(s,0) = \mathbb{E}_{M^0}\Bigg[\mathbf{E}\bigg[\int_0^t \mathcal{G}\Phi_{\exp,\ln w_s}(\Xi_\tau)\mathrm{d}\tau\,\bigg|\,\Xi_0 \sim \mu_\psi\bigg]\Bigg]. \qquad (3.11)$$

It remains to show that for every $s,t \geq 0$,

$$\mathbf{E}\Big[\mathbb{E}_{M^0}\big[\mathcal{L}D_{\psi_{t-s}^{(N_{t-s})}}(M_s)\big]\,\big|\,\Xi_0 \sim \mu_\psi\Big] = \mathbb{E}_{M^0}\Big[\mathbf{E}\big[\mathcal{G}\Phi_{\exp,\ln w_s}(\Xi_{t-s})\,\big|\,\Xi_0 \sim \mu_\psi\big]\Big], \qquad (3.12)$$

so that we may use Lemma 4.4.10 in [21] to conclude that

$$F(t,0) = F(0,t),$$

which is equivalent to (1.25).

Using (1.14) and shortening the notation $I_{x,r}(\Xi_{t-s})$ into I, we have,

$\mathcal{G}\Phi_{\exp,\ln w_s}(\Xi_{t-s})$

$$= \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)} \frac{1}{V_r}\bigg(\prod_{j\in I^c} w_s(\xi_{t-s}^j)\bigg)$$

$$\times\bigg[\sum_{D\subseteq I;|D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\bigg(w_s(y)\prod_{i\in D^c} w_s(\xi_{t-s}^i) - \prod_{i\in I} w_s(\xi_{t-s}^i)\bigg)\bigg]\mathrm{d}y\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x$$

$$+ \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2} \frac{1}{V_r^2}\bigg(\prod_{j\in I^c} w_s(\xi_{t-s}^j)\bigg)\bigg[\sum_{D\subseteq I;|D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}$$

$$\times\bigg(w_s(y)w_s(z)\prod_{i\in D^c} w_s(\xi_{t-s}^i) - \prod_{i\in I} w_s(\xi_{t-s}^i)\bigg)\bigg]\mathrm{d}y\mathrm{d}z\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x$$

$$= \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)} \frac{1}{V_r}\bigg(\prod_{j\in I^c} w_s(\xi_{t-s}^j)\bigg)\bigg[w_s(y)\bigg(\sum_{D\subseteq I;|D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\prod_{i\in D^c} w_s(\xi_{t-s}^i)\bigg)$$

$$- \big(1 - (1-u)^{|I|}\big)\prod_{i\in I} w_s(\xi_{t-s}^i)\bigg]\mathrm{d}y\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x$$

$$+ \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2} \frac{1}{V_r^2}\bigg(\prod_{j\in I^c} w_s(\xi_{t-s}^j)\bigg)\bigg[w_s(y)w_s(z)$$

$$\times\bigg(\sum_{D\subseteq I;|D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\prod_{i\in D^c} w_s(\xi_{t-s}^i)\bigg) - \big(1 - (1-u)^{|I|}\big)\prod_{i\in I} w_s(\xi_{t-s}^i)\bigg]$$

$$\mathrm{d}y\mathrm{d}z\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x. \qquad (3.13)$$

On the other hand, using (3.2) and writing ψ_{t-s} for $\psi_{t-s}^{(N_{t-s})}$ to ease the notation, we obtain

$$
\mathcal{L}D_{\psi_{t-s}}(M_s)
$$
$$
= \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \int_{(\mathbb{R}^d)^{N_{t-s}}} \psi_{t-s}(x_1, \ldots, x_{N_{t-s}}) \Big(\prod_{j \in I^c} w_s(x_j) \Big)
$$
$$
\times \Big[w_s(y) \prod_{j \in I} \big((1-u)w_s(x_j) + u \big) + (1 - w_s(y)) \prod_{j \in I} \big((1-u)w_s(x_j) \big) - \prod_{j \in I} w_s(x_j) \Big]
$$
$$
\mathrm{d}x_1 \ldots \mathrm{d}x_{N_{t-s}} \mathrm{d}y \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x
$$
$$
+ \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)^2} \frac{1}{V_r^2} \int_{(\mathbb{R}^d)^{N_{t-s}}} \psi_{t-s}(x_1, \ldots, x_{N_{t-s}}) \Big(\prod_{j \in I^c} w_s(x_j) \Big)
$$
$$
\times \Big[w_s(y)w_s(z) \prod_{j \in I} \big((1-u)w_s(x_j) + u \big) + (1 - w_s(y)w_s(z)) \prod_{j \in I} \big((1-u)w_s(x_j) \big)
$$
$$
- \prod_{j \in I} w_s(x_j) \Big] \mathrm{d}x_1 \ldots \mathrm{d}x_{N_{t-s}} \mathrm{d}y \mathrm{d}z \nu_r'(\mathrm{d}u) \mu'(\mathrm{d}r) \mathrm{d}x. \tag{3.14}
$$

Now, the first integral in the above is equal to

$$
\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \int_{B(x,r)} \frac{1}{V_r} \int_{(\mathbb{R}^d)^{N_{t-s}}} \psi_{t-s}(x_1, \ldots, x_{N_{t-s}}) \prod_{j \in I^c} w_s(x_j)
$$
$$
\times \Big[w_s(y) \sum_{D \subseteq I; |D| \geq 1} u^{|D|} (1-u)^{|I \setminus D|} \prod_{j \in D^c} w_s(x_j) - \big(1 - (1-u)^{|I|} \big) \prod_{j \in I} w_s(x_j) \Big]
$$
$$
\mathrm{d}x_1 \ldots \mathrm{d}x_{N_{t-s}} \mathrm{d}y \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x. \tag{3.15}
$$

Taking the expectation of (3.15) with respect to $\mathbf{E}[\cdot | \Xi_0 \sim \mu_\psi]$, we obtain that it is equal to the expectation of the first term on the r.h.s. of (3.13). The same holds for the expectation of the second part of (3.14) and that of the second part of (3.13). Taking then the expectation with respect to $\mathbb{E}_{M_0}$ (and using Fubini's theorem), we arrive at the desired equality (3.12). Hence, we can use Lemma 4.4.10 in [21] to conclude that

$$
F(t,0) = F(0,t),
$$

and (1.25) is proved. $\qquad \square$

4 Heuristics for the large-scale behaviour of the SLFVS and its dual process

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 trace in the long-term evolution of the population, without leading to an instantaneous invasion by the favoured allele. In particular, we wish to complement the work of [18, 19], in which the impact (or fraction of the local population replaced during an event) is kept of order $\mathcal{O}(1)$ while the selection coefficient goes to 0 as the scaling parameter n tends to infinity, modelling a population in which the local densities of individuals are low and therefore any event leads to the replacement of a macroscopic fraction of the local population. In our work, we always assume that local population densities are very high and reproduction impacts only a very small fraction of the individuals present in

the affected area (*i.e.*, we assume that $u \ll 1$). We consider long timescales (nt, $t \geq 0$) and identify the orders of magnitude of u_n and s_n, and the corresponding spatial scale ($n^\beta x$, $x \in \mathbb{R}^d$) over which the dynamics of the population will converge as $n \to \infty$ to a Fisher-KPP type evolution. In passing, we shall also see why there is no way to obtain a stochastic limit in more than one dimension when we assume that both the impact u_n of each event and the relative frequency s_n of selective to neutral events tend to 0, and we shall justify the claims made in Remarks 1.12 and 1.15 about the range of parameters leading to a deterministic Fisher-KPP process in one dimension.

Let us start with the case of local reproduction. As usual in the spatial Lambda-Fleming-Viot framework, it is easier to first think about the corresponding scaled dual processes. Chapter 7 in [35] suggests that if there is a regime of parameters in which the SLFVS converges to the solution to the Fisher-KPP equation, and in one dimension to its stochastic counterpart, then the sequence of corresponding dual processes should converge to a branching Brownian motion in which, in dimension 1, pairs of particles coalesce at a rate proportional to their collision local time (we explain the correspondence between our frameworks in the paragraph on uniqueness of the limit in the proof of Theorem 1.11, see Section 5.1). Let us thus analyse the different types of event which can affect the dual process when time is sped up by n and space is scaled down by n^β, focusing first on what happens to a single particle. During a neutral event, if this particle is marked (with probability $u_n = \mathcal{O}(n^{-\gamma})$), then it is removed and replaced by another particle whose distribution is uniformly distributed in the region of the event (of radius $Rn^{-\beta}$ in our new units). We see this as a *jump* of the particle. Because the component of the intensity of Π^N corresponding to the centres of events is Lebesgue measure on $\mathbb{R}^d$, when the particle jumps, the location of the centre of the corresponding event is uniformly distributed in the ball of radius $Rn^{-\beta}$ around the current location x of the particle. Consequently, the position of the particle "after the jump" belongs to $B(x, 2Rn^{-\beta})$ a.s. and has a radially symmetric distribution around x. Summing up the above, in our new units a single particle jumps at rate $\mathcal{O}(n^{1-\gamma})$ and makes mean zero, finite variance, jumps of size bounded by a constant times $1/n^\beta$. For this jump process to converge to some non-trivial process (Brownian motion, in fact) as $n \to \infty$, we thus have to assume that

$$1 - \gamma = 2\beta. \tag{4.1}$$

Now consider what happens at a selective event. Again our particle is marked with probability $\mathcal{O}(n^{-\gamma})$, and in this case, the two particles which replace it are created at a separation of order $\mathcal{O}(n^{-\beta})$. Consequently, they may be overlapped by a new event very quickly (after a time of order $\mathcal{O}(n^{-1})$), and then with probability $u_n^2 = \mathcal{O}(n^{-2\gamma})$ they are both erased and replaced by a single "parental" particle (we see this type of event as a "coalescence"). In the limit as $n \to \infty$, we will only "see" the branching event before it is erased by such a coalescence if the two particles have positive probability of moving apart to a distance of order one before (perhaps) coalescing. Let us find conditions under which we can expect this to hold.

In our new timescale, each particle is overlapped by an event at rate $\mathcal{O}(n)$. The probability that only one of the two particles is marked during such an event, and therefore "jumps" to a location at distance $\mathcal{O}(n^{-\beta})$ while the other stands still, is $un^{-\gamma}(1-\mathcal{O}(n^{-\gamma}))$. Furthermore, as soon as the two particles are at distance larger than $2Rn^{-\beta}$, they cannot be overlapped by the same event and so they jump independently of each other according to a continuous time random walk. Hence, what we actually have to understand is how many times the two particles come back to a separation less than $2Rn^{-\beta}$ before they manage to move apart to a separation of $\mathcal{O}(1)$. Indeed, the same type of analysis as the one carried out in the proof of Lemma 6.6 in [7] (see Lemma 5.6 and below in Section 5 of the present work) shows that when they come together, the

two particles remain at distance less than $2Rn^{-\beta}$ during a number of events *affecting them* of order $\mathcal{O}(1)$ (which translates into a number of events simply *overlapping them* of order $\mathcal{O}(n^\gamma)$). Hence, the probability that they are both affected by an event and coalesce before separating again to a distance more than $2Rn^{-\beta}$ is of the same order as the probability of coalescence during a single event conditionally on at least one of the particles being marked, which is $\mathcal{O}(n^{-2\gamma}/n^{-\gamma}) = \mathcal{O}(n^{-\gamma})$. From this, we can conclude in particular that the two particles will need to come back "together" $\mathcal{O}(n^\gamma)$ times before they have a positive probability of both being affected by the same event and therefore coalescing into common ancestral particle(s).

When they are more than $2Rn^{-\beta}$ apart, the two particles jump independently according to a continuous time symmetric random walk with step sizes of order $\mathcal{O}(n^{-\beta})$, and so the separation between them is also a symmetric random walk with step size of this order. We are interested in the probability that during an excursion away from $B(0, 2Rn^{-\beta})$, the difference walk reaches a distance of order 1, *i.e.* n^β times larger than its initial value. It is convenient to work in our original space units. For a symmetric continuous-time random walk with step size of order $\mathcal{O}(1)$, starting at distance slightly larger than $2R$ from 0, the probability of reaching distance n^β from 0 before reentering $B(0, 2R)$ has the same order as the probability that the number of steps to come back within $B(0, 2R)$ is larger than $n^{2\beta}$. This in turn will have the same order as the corresponding quantity for simple symmetric random walk. Using Proposition 5.1.1 in [34] when $d = 1$, Theorem 1 in [45] when $d = 2$ and the transience of simple symmetric random walk when $d \geq 3$, we obtain that in one dimension, this probability is of order $\mathcal{O}(n^{-\beta})$; in two dimensions, this probability is of order $\mathcal{O}(1/\ln n)$; in dimension $d \geq 3$, this probability tends to $p \in (0, 1)$ as $n \to \infty$. Consequently, when $d \geq 2$ the probability that the two particles come back together $\mathcal{O}(n^\gamma)$ times before they separate to a distance of $\mathcal{O}(1)$ tends to 0 and, in the limit, a given branching event is never followed by instantaneous coalescence. Since branching events happen at a rate $\mathcal{O}(ns_n u_n) = \mathcal{O}(n^{1-\delta-\gamma})$, we need to impose that

$$1 - \delta - \gamma = 0 \tag{4.2}$$

if they are to occur at rate $\mathcal{O}(1)$ in the limit. On the other hand, in one dimension, we see that if $\beta > \gamma$, with probability tending to one, the two particles will coalesce back together before they can separate to a distance of $\mathcal{O}(1)$ and in the limit, all branching events are cancelled (*i.e.*, there is no branching in the limiting dual). If $\beta < \gamma$, the two particles become separated at distance $\mathcal{O}(1)$ before they have any chance to coalesce, and all branching events are conserved in the limit; in contrast coalescence will never be seen in the limit for particles starting at any separation. Finally, when

$$\beta = \gamma \tag{4.3}$$

the particles have positive probability of separating to a distance of $\mathcal{O}(1)$ before coalescing, but coalescence of particles happens in finite time a.s. in the limit. In the last two cases, we also need to impose Condition (4.2) for branching events to occur at rate $\mathcal{O}(1)$ in the limit.

Solving the system given by Conditions (4.1), (4.2) and (4.3), we obtain $\beta = 1/3 = \gamma$ and $\delta = 2/3$ as specified in Section 1.3. Observe that if we keep Conditions (4.1) and (4.2) and replace Condition (4.3) by

$$\beta < \gamma, \tag{4.4}$$

even in one dimension two particles can branch but not coalesce in the limit and, as explained in the paragraph on uniqueness of the limit in the proof of Theorem 1.11 (see Section 5.1), the limit of the SLFVS is the (measure-valued) solution to the deterministic Fisher-KPP equation.

Remark 4.1. In two dimensions, our heuristics suggest that if instead of scaling space by n^β, we were to scale it by a more general factor k_n, then the relation between the parameters allowing the number of excursions necessary for the particles to separate to distance $\mathcal{O}(k_n)$ to be of the same order as the number of times the particles need to come "together" before they coalesce is

$$\ln k_n \approx 1/u_n. \tag{4.5}$$

(Here we use $\approx$ to mean that the two quantities are of the same order of magnitude.) Together with the relations $nu_n \approx k_n^2$ ensuring that the limiting motion is Brownian motion, and $nu_ns_n \approx 1$ guaranteeing that branching occurs at rate $\mathcal{O}(1)$ in the limit, this gives us that

$$\ln k_n \approx \frac{n}{k_n^2}, \qquad \text{and so } k_n \approx \sqrt{\frac{n}{\ln n}}. \tag{4.6}$$

But with this choice of parameters, the bound (5.24) we shall establish for the predictable quadratic variation of $F(\langle \overline{w}^n_\cdot, f \rangle)$ reads

$$u_n^2 nk_n^d \times k_n^{-2d} \approx \frac{k_n^2}{n} \approx \frac{1}{\ln k_n} \to 0 \quad \text{as } n \to \infty. \tag{4.7}$$

Therefore, something more subtle happens here and even with a more general form of scaling of space, the limiting process of allele frequencies is still deterministic.

We now turn to the stable case. As before, we first consider the "jumps" of a single particle. Again, such a jump occurs at rate $\mathcal{O}(n^{1-\gamma})$, and its new position is chosen uniformly over a ball whose radius is given by the intensity measure μ with polynomial decay described in (1.33). Consequently, if we choose $nu_n \propto n^{\alpha\beta}$, *i.e.*

$$1 - \gamma = \alpha\beta, \tag{4.8}$$

then in the limit as $n \to \infty$ the jump process will converge to a symmetric α-stable process with index α. Second, in order to see any branching of particles due to selective events at all, as before we need ns_nu_n to be order one, that is Condition (4.2) to be fulfilled. Finally, let us consider the simultaneous removal of two particles, to be replaced by "parental particle(s)" (what we called a coalescence earlier). Since $u_n \to 0$ as $n \to \infty$, although it is now the case that two particles can always be affected by the same event (the radii of the events are not bounded), "most of the time" they will not and their jumps are almost independent. Consequently, the difference of their positions is also approximately described by an α-stable process. Now, because events of radius $\mathcal{O}(1)$ (in our original units) are much more frequent than events of large radii $\mathcal{O}(n^a)$ for any $a > 0$, and the probability that both particles belong to the fraction of the local population replaced during an event is tiny ($u_n^2 = u^2n^{-2\gamma}$), if coalescence is to happen in the limit, then we expect it to be driven by the smaller events. Note that because there is no bound on the event radii, we can no longer perform the same decomposition into excursions away from some ball as in the fixed radius case. The following argument only gives the intuition behind Lemma 6.4 in Section 6, which allows us to control the coalescence rate of the two particles.

In more than one dimension, the rotation-invariant α-stable processes with $\alpha \in (1, 2)$ are transient (see Example 37.19(ii) in [46]), and so as in the fixed radius case, this tells us that the two particles 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 particles, carried out in the fixed radius case). Instead, the condition

$$\gamma = (\alpha - 1)\beta \tag{4.9}$$

will emerge when we control the second term on the r.h.s. of (6.9), which corresponds to the variance term in the limiting process (and thus to the coalescence term in the dual process). See also Equations (6.15) and (6.19) and the surrounding paragraphs. In the end, we have three equations in three unknowns (in one dimension) and solving gives the values in Equation (1.33). As in the fixed radius case, we may replace Condition (4.9) by

$$(\alpha - 1)\beta < \gamma \tag{4.10}$$

and obtain a limiting dual in which, in any dimension, branching occurs at rate $\mathcal{O}(1)$ but coalescence never occurs.

As a final comment, notice that in this work we have chosen a particular form for the parameters u_n and s_n, given in (1.29), and for the scaling of space (by n^β). We have argued that, within this particular framework (which nonetheless covers a wide range of scenarios), in more than one dimension it was not possible to find values for β, γ, δ such that the martingale problem characterising the limiting process $(M_t^\infty)_{t \geq 0}$ contains a Laplacian (or fractional Laplacian) term, a drift term due to the slight advantage of type 1 individuals *and* a martingale term corresponding to the noise in the Fisher-KPP-like equation satisfied by the process. Of course this does not prove that other forms of parameters and spatial scalings would not yield a stochastic limit even under the assumption that $(u_n)_{n \geq 1}$ and $(s_n)_{n \geq 1}$ tend to 0 as n tends to infinity that we have imposed. One way to investigate this question is to use the correspondence between the fact that the limiting process is stochastic and the property that the events that we loosely call "coalescence of particles" have positive probability to happen in the limiting dual process (a general property of continuous-site stepping-stone models, see Section 5 in [22]). We have not been able to find scalings for which, in the limiting dual, particles may "move in space", "branch" and "coalesce" (even in a non-local way), with positive probability when $d \geq 2$. In fact, due to the facts that rotation-invariant α-stable processes are transient for $d > \alpha$ and that coalescence happens u_n times more slowly than movement of particles before taking the limit, we conjecture that such scalings do not exist, even in the case of α-stable radii. We leave this delicate question to the interested reader.

5 Convergence of the rescaled SLFVS and its dual – the fixed radius case

In this section, we prove Theorem 1.11 and, from it, deduce Theorem 1.13.

5.1 Proof of Theorem 1.11

The proof proceeds in the usual way. First, we show that the sequence of non-Markovian processes $(\overline{M}^n)_{n \geq 1}$ is tight in $D_{\mathcal{M}_\lambda}[0, \infty)$ and that any limit point has a.s. continuous trajectories. Next, we prove that any limit point M^∞ satisfies the martingale problem stated in Theorem 1.11. Finally, we show that there is at most one solution in $D_{\mathcal{M}_\lambda}[0, \infty)$ to this martingale problem (again thanks to a duality argument), which will allow us to conclude that indeed $(\overline{M}^n)_{n \geq 1}$ converges to this solution.

1) Tightness and continuity of the limit.

First, let $n \geq 1$. Since the (unscaled) SLFVS $(M_t)_{t \geq 0}$ with parameters given in (1.27), (1.28), (1.29) and $\mu = \delta_R$ has sample paths in $D_{\mathcal{M}_\lambda}[0, \infty)$ by Theorem 1.2, so has the locally averaged and scaled process $\overline{M}^n$ (recall that its density is defined by Relation (1.30)).

Let us now show that the sequence $(\overline{M}^n)_{n\geq 1}$ is relatively compact. We proceed as in the proof of Theorem 1.2 and refer to the paragraph Proof of (i), item (c), in Section 2 for a more detailed justification of the steps taken below. Again, due to the compactness of $\mathcal{M}_\lambda$ endowed with the topology of vague convergence and the fact that the set of functions of the form $\Psi_{F,f}$ with $F \in C^3(\mathbb{R})$ and $f \in C_c^\infty(\mathbb{R}^d)$ is dense (for the supremum norm) in the set of functions of the form $\Psi_{F,f}$ with $F \in C^1(\mathbb{R})$ and $f \in C_c(\mathbb{R}^d)$, and is therefore dense in $C(\mathcal{M}_\lambda)$ by Lemma 1.1, Theorem 3.9.1 in [21] tells us that it suffices to show the relative compactness of the sequence of real-valued processes $(\Psi_{F,f}(\overline{M}^n))_{n\geq 1}$ for every $F \in C^3(\mathbb{R})$ and $f \in C_c^\infty(\mathbb{R}^d)$. Second, for each such function $\Psi_{F,f}$, we use the Aldous-Rebolledo criterion [2, 43] to reduce the problem to tightness of the sequences of the predictable finite variation parts and of the predictable quadratic variation of the martingale parts of $(\Psi_{F,f}(\overline{M}^n_t))_{t\geq 0}$. More precisely, since $\Psi_{F,f}$ is a bounded function on $\mathcal{M}_\lambda$, we directly have that for every $t \geq 0$, the sequence $(\Psi_{F,f}(\overline{M}^n_t))_{n\geq 1}$ is tight. Writing $(\mathcal{A}^n_t)_{t\geq 0}$ (resp., $(\mathcal{Q}^n_t)_{t\geq 0}$) for the finite variation part (resp., the quadratic variation of the martingale part) of $(\Psi_{F,f}(\overline{M}^n_t))_{t\geq 0}$, it remains to prove that for every $T > 0$, every sequence of stopping times $(\tau_n)_{n\geq 1}$ bounded by T, and every $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\limsup_{n\to\infty} \sup_{\theta\in[0,\eta]} \mathbb{P}\big[|\mathcal{A}^n_{\tau_n+\theta} - \mathcal{A}^n_{\tau_n}| > \varepsilon\big] \leq \varepsilon, \tag{5.1}$$

and

$$\limsup_{n\to\infty} \sup_{\theta\in[0,\eta]} \mathbb{P}\big[|\mathcal{Q}^n_{\tau_n+\theta} - \mathcal{Q}^n_{\tau_n}| > \varepsilon\big] \leq \varepsilon. \tag{5.2}$$

In what follows, we first establish an expression for the processes $\mathcal{A}^n$ and $\mathcal{Q}^n$ (see (5.12) and (5.13)), and then we prove (5.1) and (5.2) by decomposing these expressions into several parts that we control separately.

To find expressions for the predictable finite and quadratic variation parts of $\Psi_{F,f}(\overline{M}^n)$ for any fixed $n \geq 1$, let us begin by considering the unscaled SLFVS $(M_t)_{t\geq 0}$ with reproduction events of fixed radius R, and parameters u_n, s_n (notice that for simplicty we have suppressed the dependence of $(M_t)_{t\geq 0}$ on n in the notation). Also, the function F will be fixed but for a moment we replace the function f by any function $\varphi \in C_c^\infty(\mathbb{R}^d)$. A judicious choice of φ will then yield the desired function of the scaled and locally averaged process $\overline{M}^n$.

Let $\varphi \in C_c^\infty(\mathbb{R}^d)$. By Theorem 1.2, we know that before scaling space and time, the extended generator of the SLFVS, acting on the function $\Psi_{F,\varphi}$, is given by

$$\mathcal{L}\Psi_{F,\varphi}(M) = \int_{\mathbb{R}^d} \int_{B(x,R)^2} \frac{1}{V_R^2} \Big\{ w(y)(1 + s_n w(z))\big[F(\langle \Theta^+_{x,R,u_n}(w), \varphi\rangle) - F(\langle w, \varphi\rangle)\big] \tag{5.3}$$

$$+ (1 - w(y) + s_n(1 - w(y)w(z)))\big[F(\langle \Theta^-_{x,R,u_n}(w), \varphi\rangle) - F(\langle w, \varphi\rangle)\big]\Big\} \mathrm{d}y\mathrm{d}z\mathrm{d}x,$$

where w is a representative of the density of M. This gives us that the predictable finite variation part of $(\Psi_{F,\varphi}(M_t))_{t\geq 0}$ is

$$\mathcal{A}_t = \int_0^t \mathcal{L}\Psi_{F,\varphi}(M_s)\,\mathrm{d}s, \qquad t \geq 0. \tag{5.4}$$

Furthermore, the martingale problem stated in Theorem 1.2 applies to $\Psi_{F^2,\varphi} = F(\langle \cdot, \varphi\rangle)^2$, which allows us to obtain (using Itô's formula) that the predictable quadratic variation of $(\Psi_{F,\varphi}(M_t))_{t\geq 0}$ at any time $t \geq 0$ is given by

$$\mathcal{Q}_t = \int_0^t \int_{\mathbb{R}^d} \int_{B(x,R)^2} \frac{1}{V_R^2} \Big\{ w_s(y)(1 + s_n w_s(z))\big[F(\langle \Theta^+_{x,R,u_n}(w_s), \varphi\rangle) - F(\langle w_s, \varphi\rangle)\big]^2 \tag{5.5}$$

$$+ (1 - w_s(y) + s_n(1 - w_s(y)w_s(z)))\big[F(\langle \Theta^-_{x,R,u_n}(w_s), \varphi\rangle) - F(\langle w_s, \varphi\rangle)\big]^2\Big\} \mathrm{d}y\mathrm{d}z\mathrm{d}x\mathrm{d}s.$$

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

$$B_n(x) = B(x, n^{-1/3}R) \tag{5.6}$$

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 $t \geq 0$

$$\frac{n^{d/3}}{V_R}\int_{B_n(x)} w_t^n(z)\mathrm{d}z = \frac{1}{V_R}\int_{B(n^{1/3}x,R)} w_{nt}(y)\mathrm{d}y = \overline{w}_t^n(x). \tag{5.7}$$

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

$$\mathcal{L}^n \Psi_{F,\varphi}(M)$$
$$= n\int_{\mathbb{R}^d}\int_{B(x,R)^2} \frac{1}{V_R^2}\Big\{w(n^{-1/3}y)\big(1 + s_n w(n^{-1/3}z)\big)\big[F(\langle\Theta^+_{n^{-1/3}x,n^{-1/3}R,u_n}(w),\varphi\rangle)$$
$$- F(\langle w,\varphi\rangle)\big]$$
$$+ \big(1 - w(n^{-1/3}y) + s_n(1 - w(n^{-1/3}y)w(n^{-1/3}z))\big)\big[F(\langle\Theta^-_{n^{-1/3}x,n^{-1/3}R,u_n}(w),\varphi\rangle)$$
$$- F(\langle w,\varphi\rangle)\big]\Big\}\mathrm{d}y\mathrm{d}z\mathrm{d}x$$
$$= n^{1+\frac{d}{3}}\int_{\mathbb{R}^d}\Big\{\overline{w}(x)(1 + s_n\overline{w}(x))\big[F(\langle\Theta^+_{x,n^{-1/3}R,u_n}(w),\varphi\rangle) - F(\langle w,\varphi\rangle)\big]$$
$$+ \big(1 - \overline{w}(x) + s_n(1 - \overline{w}(x)^2)\big)\big[F(\langle\Theta^-_{x,n^{-1/3}R,u_n}(w),\varphi\rangle) - F(\langle w,\varphi\rangle)\big]\Big\}\mathrm{d}x. \tag{5.8}$$

The predictable finite variation part of $(\Psi_{F,\varphi}(M_t^n))_{t\geq 0}$ is thus $(\int_0^t \mathcal{L}^n\Psi_{F,\varphi}(M_s^n)\mathrm{d}s)_{t\geq 0}$. Likewise, its predictable quadratic variation at time t is equal to

$$n^{1+\frac{d}{3}}\int_0^t\int_{\mathbb{R}^d}\Big\{\overline{w}_s^n(x)(1 + s_n\overline{w}_s^n(x))\big[F(\langle\Theta^+_{x,n^{-1/3}R,u_n}(w_s^n),\varphi\rangle) - F(\langle w_s^n,\varphi\rangle)\big]^2$$
$$+ \big(1 - \overline{w}_s^n(x) + s_n(1 - \overline{w}_s^n(x)^2)\big)\big[F(\langle\Theta^-_{x,n^{-1/3}R,u_n}(w_s^n),\varphi\rangle) - F(\langle w_s^n,\varphi\rangle)\big]^2\Big\}\mathrm{d}x\mathrm{d}s. \tag{5.9}$$

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

$$\varphi_f(x) = \frac{n^{d/3}}{V_R}\int_{B_n(x)} f(y)\mathrm{d}y \tag{5.10}$$

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

$$\langle w^n, \varphi_f\rangle = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} w^n(y)\frac{n^{d/3}}{V_R}f(z)\mathbb{1}_{\{|z-y|\leq n^{-1/3}R\}}\mathrm{d}y\mathrm{d}z = \langle\overline{w}^n, f\rangle, \tag{5.11}$$

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

$$\mathcal{A}_t^n = \int_0^t \mathcal{L}^n\Psi_{F,\varphi_f}(M_s^n)\mathrm{d}s, \tag{5.12}$$

with $\mathcal{L}^n\Psi_{F,\varphi_f}$ as in (5.8), and its predictable quadratic variation is given by

$$\mathcal{Q}_t^n = n^{1+\frac{d}{3}}\int_0^t\int_{\mathbb{R}^d}\Big\{\overline{w}_s^n(x)(1 + s_n\overline{w}_s^n(x))\big[F(\langle\Theta^+_{x,n^{-1/3}R,u_n}(w_s^n),\varphi_f\rangle) - F(\langle w_s^n,\varphi_f\rangle)\big]^2$$
$$+ \big(1 - \overline{w}_s^n(x) + s_n(1 - \overline{w}_s^n(x)^2)\big)\big[F(\langle\Theta^-_{x,n^{-1/3}R,u_n}(w_s^n),\varphi_f\rangle) - F(\langle w_s^n,\varphi_f\rangle)\big]^2\Big\}\mathrm{d}x\mathrm{d}s. \tag{5.13}$$

Note that

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

$$\langle \Theta^-_{x,n^{-1/3}R,u_n}(w), \varphi_f \rangle - \langle w, \varphi_f \rangle = -u_n \langle \mathbb{1}_{B_n(x)}w, \varphi_f \rangle, \tag{5.14}$$

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 $|\mathcal{A}^n_t|$ 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 $t \geq 0$,

$$\mathcal{A}^n_t = \int_0^t \big(A_n(s) + B_n(s) + C_n(s) + D_n(s) + E_n(s) \big) \, \mathrm{d}s, \tag{5.15}$$

where

$$A_n(s) = u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f \rangle) \int_{\mathbb{R}^d} \Big[\overline{w}^n_s(x) \langle \mathbb{1}_{B_n(x)}(1-w^n_s), \varphi_f \rangle$$
$$- (1 - \overline{w}^n_s(x)) \langle \mathbb{1}_{B_n(x)} w^n_s, \varphi_f \rangle \Big] \mathrm{d}x,$$

$$B_n(s) = u_n^2 n^{1+\frac{d}{3}} \frac{F''(\langle \overline{w}^n_s, f \rangle)}{2} \int_{\mathbb{R}^d} \Big[\overline{w}^n_s(x) \langle \mathbb{1}_{B_n(x)}(1-w^n_s), \varphi_f \rangle^2$$
$$+ (1 - \overline{w}^n_s(x)) \langle \mathbb{1}_{B_n(x)} w^n_s, \varphi_f \rangle^2 \Big] \mathrm{d}x,$$

$$C_n(s) \leq \mathcal{C} n^{1+\frac{d}{3}} \int_{\mathbb{R}^d} \big(u_n \mathrm{Vol}(B_n(x)) \big)^3 \mathbb{1}_{\{B_n(x) \cap S_f \neq \emptyset\}} \mathrm{d}x,$$

$$D_n(s) = u_n s_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f \rangle) \int_{\mathbb{R}^d} \Big[\overline{w}^n_s(x)^2 \langle \mathbb{1}_{B_n(x)}(1-w^n_s), \varphi_f \rangle$$
$$- (1 - \overline{w}^n_s(x)^2) \langle \mathbb{1}_{B_n(x)} w^n_s, \varphi_f \rangle \Big] \mathrm{d}x,$$

$$E_n(s) \leq \mathcal{C}' n^{1+\frac{d}{3}} s_n u_n^2 \int_{\mathbb{R}^d} \mathrm{Vol}(B_n(x))^2 \mathbb{1}_{\{B_n(x) \cap S_f \neq \emptyset\}} \mathrm{d}x, \tag{5.16}$$

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 ($H\varphi_f = DD\varphi_f$). Recall that S_f denotes the compact support of f. Then

$$A_n(s) = u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f \rangle) \int_{\mathbb{R}^d} \Big[\overline{w}^n_s(x) \langle \mathbb{1}_{B_n(x)}, \varphi_f \rangle - \langle \mathbb{1}_{B_n(x)} w^n_s, \varphi_f \rangle \Big] \mathrm{d}x \tag{5.17}$$

$$= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f \rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int\int \mathbb{1}_{\{|y-x| \leq n^{-1/3}R\}} \mathbb{1}_{\{|z-x| \leq n^{-1/3}R\}} w^n_s(y)$$
$$\times (\varphi_f(z) - \varphi_f(y)) \mathrm{d}z \mathrm{d}y \mathrm{d}x$$

$$= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f \rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x| \leq n^{-1/3}R\}} \mathbb{1}_{\{|z-x| \leq n^{-1/3}R\}} w^n_s(y)$$
$$\times \Big[D\varphi_f(y)(z-y) + \frac{1}{2}(z-y)H\varphi_f(y)(z-y) + \mathcal{O}(|z-y|^3) \mathbb{1}_{\{y \in S_f\}} \Big] \mathrm{d}z \mathrm{d}y \mathrm{d}x.$$

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

$$\frac{u_n n^{1+\frac{2d}{3}}}{V_R} F'(\langle \overline{w}_s^n, f \rangle) \int_{\mathbb{R}^d} w_s^n(y) \int_{\mathbb{R}^d} \text{Vol}(B_n(y) \cap B_n(z)) D\varphi_f(y)(z-y) \mathrm{d}z \mathrm{d}y, \qquad (5.18)$$

and since $\text{Vol}(B_n(y) \cap B_n(z))$ 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.

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

$$u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f \rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x| \le n^{-1/3}R\}} \mathbb{1}_{\{|z-x| \le n^{-1/3}R\}} w_s^n(y)$$

$$\times \frac{1}{2} \sum_{i=1}^d (z_i - y_i)^2 \frac{\partial^2}{\partial y_i^2} \varphi_f(y) \mathrm{d}y \mathrm{d}z \mathrm{d}x \qquad (5.19)$$

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

$$\frac{\partial^2}{\partial y_i^2} \varphi_f(y) = \varphi_{\frac{\partial^2 f}{\partial y_i^2}}(y) = \frac{\partial^2 f}{\partial y_i^2}(y) + \mathcal{O}(n^{-2/3}) \mathbb{1}_{\{B_n(y) \cap S_f \ne \emptyset\}} \qquad (5.20)$$

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

$$A_n(s) = u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f \rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x| \le n^{-1/3}R\}} \mathbb{1}_{\{|z-x| \le n^{-1/3}R\}} w_s^n(y)$$

$$\times \frac{1}{2} \sum_{i=1}^d (z_i - y_i)^2 \frac{\partial^2 f}{\partial y_i^2}(y) \mathrm{d}y \mathrm{d}z \mathrm{d}x$$

$$+ \mathcal{O}(n^{-2/3}) n^{\frac{2}{3}(1+d)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x| \le n^{-1/3}R\}} \mathbb{1}_{\{|z-x| \le n^{-1/3}R\}} |z-y|^2 \mathbb{1}_{\{B_n(y) \cap S_f \ne \emptyset\}}$$

$$= \frac{u n^{\frac{2}{3}(1+d)}}{2V_R} F'(\langle \overline{w}_s^n, f \rangle) \int_{\mathbb{R}^d} \int_{B_n(x)^2} w_s^n(y) \sum_{i=1}^d (z_i - y_i)^2 \frac{\partial^2 f}{\partial y_i^2}(y) \mathrm{d}y \mathrm{d}z \mathrm{d}x + \mathcal{O}(n^{-2/3})$$

$$= \frac{u \Gamma_R}{2} F'(\langle \overline{w}_s^n, f \rangle) \int_{\mathbb{R}^d} w_s^n(y) \Delta f(y) \mathrm{d}y + \mathcal{O}(n^{-2/3})$$

$$= \frac{u \Gamma_R}{2} F'(\langle \overline{w}_s^n, f \rangle) \langle \overline{w}_s^n, \Delta f \rangle + \mathcal{O}(n^{-2/3}), \qquad (5.21)$$

where

$$\Gamma_R = \frac{n^{\frac{2}{3}(1+d)}}{V_R} \int_{B_n(y)} \int_{B_n(x)} (z_1 - y_1)^2 \mathrm{d}z \mathrm{d}x = \frac{1}{V_R} \int_{B(0,R)} \int_{B(x,R)} (z_1)^2 \mathrm{d}z \mathrm{d}x \qquad (5.22)$$

was defined in (1.34), and the last equality uses another Taylor expansion to show that for any s,

$$\langle w_s^n, \Delta f \rangle = \langle \overline{w}_s^n, \Delta f \rangle + \mathcal{O}(n^{-2/3}) \qquad (5.23)$$

with an error term uniformly bounded in s. In particular, since $|\langle \overline{w}_s^n, f \rangle| \le \|f\| \text{Vol}(S_f)$, 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:

$$|B_n(s)| \le \frac{u_n^2 n^{1+\frac{d}{3}}}{2} |F''(\langle \overline{w}_s^n, f \rangle)| \int 2 \text{Vol}(B_n(x))^2 \mathbb{1}_{\{x \in S_f\}} \|f\|^2 \mathrm{d}x \le \mathcal{C}_B n^{\frac{1-d}{3}}, \qquad (5.24)$$

and, again by the same arguments,

$$|C_n(s)| \leq \mathcal{C}_C \, n^{-\frac{2d}{3}}, \quad |D_n(s)| \leq \mathcal{C}_D \quad \text{and} \quad |E_n(s)| \leq \mathcal{C}_E \, n^{-\frac{1+d}{3}}, \tag{5.25}$$

where the constants $\mathcal{C}_B, \mathcal{C}_C, \mathcal{C}_D, \mathcal{C}_E$ are all independent of n and s. Coming back to (5.15) and combining all the estimates we just obtained, for every $s < t$ we have

$$|\mathcal{A}_t^n - \mathcal{A}_s^n| \leq \left(\mathcal{C}_A + \mathcal{C}_B \, n^{\frac{1-d}{3}} + \mathcal{C}_C \, n^{-\frac{2d}{3}} + \mathcal{C}_D + \mathcal{C}_E \, n^{-\frac{1+d}{3}}\right)(t - s). \tag{5.26}$$

From there it is easy to deduce that for every $T > 0$, given a sequence of stopping times $(\tau_n)_{n \geq 1}$ bounded by T, for every $\varepsilon > 0$ there exists $\eta > 0$ such that

$$\limsup_{n \to \infty} \sup_{\theta \in [0, \eta]} \mathbb{P}\left[\left|\mathcal{A}_{\tau_n + \theta}^n - \mathcal{A}_{\tau_n}^n\right| > \varepsilon\right] = 0, \tag{5.27}$$

which corresponds to (5.1) and shows that the sequence of finite variation parts of $(\Psi_{F,f}(\overline{M}_t^n))_{t \geq 0}$ is tight.

Similarly, we obtain that

$$\left[F(\langle \Theta_{x, n^{-1/3}R, u_n}^{\pm}(w_s^n), \varphi_f \rangle) - F(\langle w_s^n, \varphi_f \rangle)\right]^2 \leq \mathcal{C}_F'' \|f\|^2 u_n^2 \mathrm{Vol}(B_n(x))^2 \mathbb{1}_{\{B_n(x) \cap S_f \neq \emptyset\}}. \tag{5.28}$$

Notice that this bound is independent of the value of w_s^n. Substituting into the definition of $\mathcal{Q}_t^n$ given in (5.13), we obtain that for every $s < t$,

$$|\mathcal{Q}_t^n - \mathcal{Q}_s^n| \leq \mathcal{C}_F \, n^{\frac{1-d}{3}}(t - s), \tag{5.29}$$

for a constant $\mathcal{C}_F$ independent of n (and s, t). Therefore, for every $T > 0$, every sequence of stopping times $(\tau_n)_{n \geq 1}$ bounded by T and every $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\limsup_{n \to \infty} \sup_{\theta \in [0, \eta]} \mathbb{P}\left[\left|\mathcal{Q}_{\tau_n + \theta}^n - \mathcal{Q}_{\tau_n}^n\right| > \varepsilon\right] = 0 \tag{5.30}$$

and the sequence of predictable quadratic variations of the martingale part of the process $(\Psi_{F,f}(\overline{M}_t^n))_{t \geq 0}$ is not only tight, but also when $d \geq 2$ it tends to 0 uniformly over compact time intervals. By the Aldous-Rebolledo criterion (see again (i)-item (c) in Section 2), we conclude that $(\overline{M}^n)_{n \geq 1}$ is tight in $D_{\mathcal{M}_\lambda}[0, \infty)$, as required.

Finally, coming back to (5.14), we see that every increment of $\langle \overline{M}^n, f \rangle = \langle M^n, \varphi_f \rangle$ is bounded by $\|f\| V_R u_n n^{-d/3}$. Consequently, for every $T > 0$ we have

$$\sup_{t \in [0, T]} \sup_{f \in C_c^\infty(\mathbb{R}^d) : \|f\| \leq 1} \left|\langle \overline{M}_t^n, f \rangle - \langle \overline{M}_{t-}^n, f \rangle\right| \leq V_R u n^{-(1+d)/3}, \tag{5.31}$$

and thus any potential limit for $(\overline{M}^n)_{n \geq 1}$ has continuous paths in $\mathcal{M}_\lambda$.

2) Identifying the limit.

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 $(\overline{M}^{n_k})_{k \geq 1}$ and for any $t \geq 0$, we write w_t^∞ for (some representative of) the density of M_t^∞.

In order to show that M^∞ satisfies the martingale problem stated in Theorem 1.11, we use the fact (established in the previous paragraph) that for every $f \in C_c^\infty(\mathbb{R}^d)$ and every $n \geq 1$,

$$\left(\Psi_{\mathrm{Id}, f}(\overline{M}_t^n) - \Psi_{\mathrm{Id}, f}(\overline{M}_0^n) - \int_0^t \mathcal{L}^n \Psi_{\mathrm{Id}, \varphi_f}(M_s^n) \mathrm{d}s\right)_{t \geq 0} \tag{5.32}$$

is a martingale with predictable quadratic variation (5.13) (with $F = \mathrm{Id}$), where $\mathcal{L}^n$ was defined in (5.8) and φ_f in (5.10). We first show that for every $t \geq 0$,

$$\lim_{k \to \infty} \mathbb{E}\left[\left|\int_0^t \mathcal{L}^{n_k} \Psi_{\mathrm{Id}, \varphi_f}(M_s^{n_k}) \mathrm{d}s - \int_0^t \left\{\frac{u \Gamma_R}{2} \langle w_s^\infty, \Delta f \rangle - 2Ru\sigma \langle w_s^\infty(1 - w_s^\infty), f \rangle\right\} \mathrm{d}s\right|\right] = 0, \tag{5.33}$$

so that we can then use the fact that the quantity in (5.32) is a martingale, the fact that $\Psi_{\mathrm{Id},f}$ is a bounded continuous function and the Dominated Convergence Theorem to conclude that for every $0 \leq t < t'$, $m \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_m \leq t$ and $h_1, \ldots, h_m \in C_b(\mathcal{M}_\lambda)$,

$$
\mathbb{E}\left[\left(\langle w_{t'}^\infty, f \rangle - \langle w_t^\infty, f \rangle - \int_t^{t'} \left\{ \frac{u\Gamma_R}{2} \langle w_s^\infty, \Delta f \rangle - 2Ru\sigma \langle w_s^\infty(1 - w_s^\infty), f \rangle \right\} \mathrm{d}s \right) \right.
$$
$$
\left. \times \left(\prod_{i=1}^m h_i\big(M_{t_i}^\infty\big) \right) \right] = 0 \quad (5.34)
$$

and consequently that $\mathcal{Z}^f$ is a martingale (with respect to the natural filtration of M^∞). In the case $d \geq 2$ this property will be sufficient to conclude, since we showed in (5.29) that the quadratic variation of the martingale (5.32) tended to 0 as $n \to \infty$, and therefore the limit $\mathcal{Z}^f$ is the constant process equal to 0. In one dimension, we shall still have to prove that the quadratic variation of $\mathcal{Z}^f$ is non-trivial and has the announced form. This is what we do in the last part of this point **2**).

Let us prove (5.33). Specialising the computation of $\mathcal{A}^n$ in (5.8) to the case $F = \mathrm{Id}$, we have by (5.21)

$$
\begin{aligned}
A_{n_k}(s) &= \frac{u\Gamma_R}{2} \langle \overline{w}_s^{n_k}, \Delta f \rangle + \mathcal{O}(n_k^{-2/3}) \\
&= \frac{u\Gamma_R}{2} \int_{\mathbb{R}^d \times \{0,1\}} \Delta f(x) \mathbb{1}_{\{0\}}(\kappa) \overline{M}_s^{n_k}(\mathrm{d}x, \mathrm{d}\kappa) + \mathcal{O}(n_k^{-2/3}) \\
&\to \frac{u\Gamma_R}{2} \langle w_s^\infty, \Delta f \rangle \qquad \text{as } k \to \infty.
\end{aligned}
\quad (5.35)
$$

These quantities being bounded by $(u\Gamma_R/2)\|\Delta f\|\mathrm{Vol}(S_f) + \mathcal{O}(n_k^{-2/3})$, independently of s, the convergence also happens in $\mathbb{L}^1$ norm. Next, Taylor-expanding f to write that for every $y \in B(x, Rn_k^{-1/3})$,

$$
\varphi_f(y) = f(x) + \mathcal{O}(|y - x|) = f(x) + \mathcal{O}\big(n_k^{-1/3}\big), \quad (5.36)
$$

we obtain that

$$
\begin{aligned}
D_{n_k}(s) &= \sigma u n_k^{d/3} \int_{\mathbb{R}^d} \left\{ \overline{w}_s^{n_k}(x)^2 \big\langle \mathbb{1}_{B_{n_k}(x)}(1 - w_s^{n_k}), f(x) + \mathcal{O}\big(n_k^{-1/3}\big) \big\rangle \right. \\
&\qquad\qquad \left. - (1 - \overline{w}_s^{n_k}(x)^2) \big\langle \mathbb{1}_{B_{n_k}(x)} w_s^{n_k}, f(x) + \mathcal{O}\big(n_k^{-1/3}\big) \big\rangle \right\} \mathrm{d}x \\
&= \sigma u V_R \int_{\mathbb{R}^d} \left\{ \overline{w}_s^{n_k}(x)^2(1 - \overline{w}_s^{n_k}(x)) - (1 - \overline{w}_s^{n_k}(x)^2)\overline{w}_s^{n_k}(x) \right\} f(x) \mathrm{d}x + \mathcal{O}\big(n_k^{-1/3}\big) \\
&= -\sigma u V_R \langle \overline{w}_s^{n_k}(1 - \overline{w}_s^{n_k}), f \rangle + \mathcal{O}\big(n_k^{-1/3}\big).
\end{aligned}
\quad (5.37)
$$

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

$$
-\sigma u V_R \langle w_s^\infty, f \rangle. \quad (5.38)
$$

We now would like to show that the "quadratic" part of $D_{n_k}(s)$ converges to

$$
\sigma u V_R \langle (w_s^\infty)^2, f \rangle. \quad (5.39)
$$

Note that this is not a simple consequence of the weak convergence of $\overline{M}^{n_k}$ to M^∞, as $\langle (\overline{w}_s^{n_k})^2, f \rangle$ cannot be written as an integral with respect to $\overline{M}_s^{n_k}$ or $(\overline{M}_s^{n_k})^{\otimes 2}$. Instead, we shall approximate this expression by an integral with respect to $(\overline{M}_s^{n_k})^{\otimes 2}$ and use the continuity estimates obtained in Proposition A.1 to bound the remaining terms. (The

statement and proof of this proposition are postponed until Appendix A 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 $k \geq 1$ and $s \geq 0$, we have

$$
\begin{aligned}
&\left| \langle (\overline{w}_s^{n_k})^2, f \rangle - \langle (w_s^\infty)^2, f \rangle \right| \\
&\leq \left| \int_{\mathbb{R}^d} f(x) \overline{w}_s^{n_k}(x)^2 \mathrm{d}x - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{w}_s^{n_k}(x) \overline{w}_s^{n_k}(y) p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x \right| \\
&\quad + \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{w}_s^{n_k}(x) \overline{w}_s^{n_k}(y) p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) w_s^\infty(x) w_s^\infty(y) p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x \right| \\
&\quad + \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) w_s^\infty(x) w_s^\infty(y) p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x - \int_{\mathbb{R}^d} f(x) w_s^\infty(x)^2 \mathrm{d}x \right|. \tag{5.40}
\end{aligned}
$$

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

$$
\begin{aligned}
&\int_{(\mathbb{R}^d \times \{0,1\})^2} f(x) p_\varepsilon(y - x) \mathbb{1}_{\{0\}}(\kappa) \mathbb{1}_{\{0\}}(\kappa') \overline{M}_s^{n_k}(\mathrm{d}y, \mathrm{d}\kappa') \overline{M}_s^{n_k}(\mathrm{d}x, \mathrm{d}\kappa) \\
&\to \int_{(\mathbb{R}^d \times \{0,1\})^2} f(x) p_\varepsilon(y - x) \mathbb{1}_{\{0\}}(\kappa) \mathbb{1}_{\{0\}}(\kappa') M_s^\infty(\mathrm{d}y, \mathrm{d}\kappa') M_s^\infty(\mathrm{d}x, \mathrm{d}\kappa) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) w_s^\infty(x) w_s^\infty(y) p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x \tag{5.41}
\end{aligned}
$$

as k 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 k (and ε, s), this convergence also happens in $\mathbb{L}^1$ norm. That is, the expectation of the second term in (5.40) tends to 0 as $k \to \infty$.

Concerning the first term on the r.h.s. of (5.40), because $\overline{w}_s^{n_k}$ takes its values in $[0, 1]$, we have, by Fubini's Theorem,

$$
\begin{aligned}
&\mathbb{E}\left[\left| \int_{\mathbb{R}^d} f(x) \overline{w}_s^{n_k}(x)^2 \mathrm{d}x - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{w}_s^{n_k}(x) \overline{w}_s^{n_k}(y) p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x \right| \right] \\
&\qquad \leq \|f\| \int_{S_f} \int_{B(x,\varepsilon)} \mathbb{E}\left[\left| \overline{w}_s^{n_k}(x) - \overline{w}_s^{n_k}(y) \right| \right] p_\varepsilon(y - x) \mathrm{d}y \mathrm{d}x. \tag{5.42}
\end{aligned}
$$

By Proposition A.1 applied with $\varepsilon = R n_k^{-1/3}$, there exists $a, v, \lambda, C > 0$ independent of k such that for every $x, y \in \mathbb{R}^d$ satisfying $|x - y| < 1$ and every $s \in [0, t]$, we have

$$
\begin{aligned}
\mathbb{E}\left[\left| \overline{w}_s^{n_k}(x) - \overline{w}_s^{n_k}(y) \right| \right] \leq C \Big\{ & n_k^{-a} + \tau_{n_k}(x, y) + \left(|x - y|^{1/4} + \tau_{n_k}(x, y)^{1/2} \right) e^{\lambda(|x| + R n_k^{-1/3})} \\
& + n_k^{(1-d)/6} \tau_{n_k}(x, y)^{(2-d)/4} \Big\}, \tag{5.43}
\end{aligned}
$$

where

$$
\tau_n(x, y) = n^{-v} \vee |x - y|^{2/(d+1)}. \tag{5.44}
$$

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 on the r.h.s. of (5.40) is bounded by

$$
C' \left(n_k^{-a} + \varepsilon^{2/(d+1)} + \varepsilon^{1/4} + \varepsilon^{1/(d+1)} + n_k^{(1-d)/6} \varepsilon^{1/(d+1)} \right). \tag{5.45}
$$

Likewise, by taking $n \to \infty$ in Proposition A.1 (along the converging subsequence), we obtain that the last term on the r.h.s. of (5.40) is bounded by

$$
C' \left(\varepsilon^{1/4} + \varepsilon^{2/(d+1)} + \varepsilon^{1/(d+1)} + \varepsilon^{1/(d+1)} \mathbf{1}_{\{d=1\}} \right). \tag{5.46}
$$

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

$$\limsup_{k \to \infty} \mathbb{E}\big[\big|\langle(\overline{w}_s^{n_k})^2, f\rangle - \langle(w_s^\infty)^2, f\rangle\big|\big] \leq C\big(\varepsilon^{1/4} + \varepsilon^{1/(d+1)}\big), \tag{5.47}$$

and letting ε tend to 0 we can conclude that the part of the expression (5.37) for $D_{n_k}(s)$ which is quadratic in $\overline{w}_s^{n_k}$ indeed converges in $\mathbb{L}^1$ towards

$$\sigma u V_R \langle(w_s^\infty)^2, f\rangle. \tag{5.48}$$

Combining (5.35), (5.38) and (5.48), and using the facts that $B_n(s) = 0$ since $F = \mathrm{Id}$, and that $C_n(s)$ and $E_n(s)$ tend to zero uniformly in all possible values of $\overline{M}^n$, we conclude that (5.33) is satisfied. As we explained above, this is sufficient to conclude in the case $d \geq 2$ since the quadratic variation of the martingale $\mathcal{Z}^f$ is then 0.

We now turn to the case $d = 1$. Defining

$$W_t^n(f) := \langle\overline{w}_t^n, f\rangle - \langle\overline{w}_0^n, f\rangle - \int_0^t \mathcal{L}^n \Psi_{\mathrm{Id},\varphi_f}(M_s^n)\mathrm{d}s \tag{5.49}$$

$$= \langle\overline{w}_t^n, f\rangle - \langle\overline{w}_0^n, f\rangle - \int_0^t \left\{\frac{u\Gamma_R}{2}\langle\overline{w}_s^n, \Delta f\rangle - \sigma u V_R \langle\overline{w}_s^n(1 - \overline{w}_s^n), f\rangle\right\}\mathrm{d}s + \mathcal{O}(n^{-1/3}),$$

we know that $W^n(f)$ is a zero-mean martingale with predictable quadratic variation

$$u_n^2 n^{4/3} \int_0^t \int_{\mathbb{R}^d} \Big\{\overline{w}_s^n(x)(1 + s_n\overline{w}_s^n(x))\langle\mathbb{1}_{B_n(x)}(1 - w_s^n), f + \mathcal{O}(n^{-1/3})\rangle^2$$

$$+ (1 - \overline{w}_s^n(x) + s_n(1 - \overline{w}_s^n(x)^2))\langle\mathbb{1}_{B_n(x)}w_s^n, f + \mathcal{O}(n^{-1/3})\rangle^2\Big\}\mathrm{d}x\mathrm{d}s$$

$$= u^2 V_R^2 \int_0^t \langle\overline{w}_s^n(1 - \overline{w}_s^n), f^2\rangle\mathrm{d}s + \mathcal{O}(n^{-1/3}), \tag{5.50}$$

where, more precisely, the remainder term is bounded by a constant times $n^{-1/3}t$. As a consequence, for every $n \geq 1$, $0 \leq t < t'$, $m \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_m \leq t$ and $h_1, \ldots, h_m \in C_b(\mathcal{M}_\lambda)$,

$$\mathbb{E}\left[\left(\big(W_{t'}^n(f)\big)^2 - \big(W_t^n(f)\big)^2 - u^2 V_R^2 \int_t^{t'} \langle\overline{w}_s^n(1 - \overline{w}_s^n), f^2\rangle\mathrm{d}s + \mathcal{O}(n^{-1/3})\right)\left(\prod_{i=1}^m h_i\big(\overline{M}_{t_i}^n\big)\right)\right] = 0 \tag{5.51}$$

Observe that for every $n \geq 1$ and every $t \geq 0$,

$$|W_t^n(f)| \leq \mathrm{Vol}(S_f)\left[2\|f\| + t\left(\frac{u\Gamma_R}{2}\|\Delta f\| + \sigma u V_R\|f\| + \mathcal{O}(n^{-1/3})\right)\right], \tag{5.52}$$

and so we can let $n \to \infty$ in (5.51) (along the converging subsequence) and use the Dominated Convergence Theorem, together with (5.35), (5.38) and (5.48), to conclude that

$$\mathbb{E}\left[\left(\big(\mathcal{Z}_{t'}^f\big)^2 - \big(\mathcal{Z}_t^f\big)^2 - u^2 V_R^2 \int_t^{t'} \langle w_s^\infty(1 - w_s^\infty), f^2\rangle\mathrm{d}s\right)\left(\prod_{i=1}^m h_i\big(M_{t_i}^\infty\big)\right)\right] = 0. \tag{5.53}$$

This allows us to identify the quadratic variation of the martingale $\mathcal{Z}^f$ as

$$[\mathcal{Z}^f]_t = u^2 V_R^2 \int_0^t \langle w_s^\infty(1 - w_s^\infty), f^2\rangle\mathrm{d}s, \qquad t \geq 0. \tag{5.54}$$

Since by (5.14) the jumps of $W^n(f)$ are all bounded by $Cn^{-2/3}$, $\mathcal{Z}^f$ is a continuous square-integrable martingale, starting at 0. By the Dubins-Schwarz Theorem (see Remark 5.1

below), $\mathcal{Z}^f$ is therefore a time-changed Brownian motion, solution to the stochastic differential equation

$$dW_t = uV_R\sqrt{\langle w_t^\infty(1-w_t^\infty), f^2\rangle}\,dB_t^f, \tag{5.55}$$

where B^f denotes standard Brownian motion. The bracket process between $\mathcal{Z}^f$ and $\mathcal{Z}^g$ is then obtained by the same kind of calculations, writing first the bracket process for a fixed n and then identifying the limit by letting $n_k \to \infty$.

Remark 5.1. We cannot *a priori* prove that $[\mathcal{Z}^f]_\infty = +\infty$ a.s., as required by the classical Dubins-Schwarz Theorem. Note however that this condition can be removed, at the expense of extending the probability space on which we work. Indeed, if we introduce a Brownian motion $(\beta_t^f)_{t\geq 0}$ independent of all other processes (possibly on some enlarged space) and set for every $t \geq 0$

$$B_t^f = \begin{cases} \mathcal{Z}_{\tau_t}^f & \text{if } t < [\mathcal{Z}^f]_\infty, \\ \mathcal{Z}_\infty^f + \beta_{t-[\mathcal{Z}^f]_\infty}^f & \text{if } t \geq [\mathcal{Z}^f]_\infty, \end{cases} \tag{5.56}$$

where

$$\tau_t := \inf\left\{s \geq 0 : [\mathcal{Z}^f]_s > t\right\}, \tag{5.57}$$

then by Theorem 1.7 in Chapter V of [44] we have that $(B_t^f)_{t\geq 0}$ is a standard Brownian motion and for every $t \geq 0$, $\mathcal{Z}_t^f = B_{[\mathcal{Z}^f]_t}^f$.

To summarise, we have shown that 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 $f \in C_c^\infty(\mathbb{R}^d)$,

$$\mathrm{d}\langle w_t^\infty, f\rangle = \left\{\frac{u\Gamma_R}{2}\langle w_t^\infty, \Delta f\rangle - \sigma uV_R\langle w_t^\infty(1-w_t^\infty), f\rangle\right\}\mathrm{d}t$$
$$+ \mathbb{1}_{\{d=1\}}uV_R\sqrt{\langle w_t^\infty(1-w_t^\infty), f^2\rangle}\mathrm{d}B_t^f, \tag{5.58}$$

with initial value $\langle w_0, f\rangle$, and in one dimension, by polarisation the covariation between $\langle w_\cdot^\infty, f\rangle$ and $\langle w_\cdot^\infty, g\rangle$ is as in the statement of Theorem 1.11(i).

3) Uniqueness of the limit.

Let us finally show that the system of equations (5.58) has at most one solution. We start with the case $d \geq 2$. Any test function of the form

$$M \mapsto \int_{(\mathbb{R}^d\times\{0,1\})^k} \psi(x_1,\ldots,x_k)\left\{\prod_{j=1}^k \mathbb{1}_{\{0\}}(\kappa_j)\right\} M(\mathrm{d}x_1,\mathrm{d}\kappa_1)\cdots M(\mathrm{d}x_k,\mathrm{d}\kappa_k)$$
$$= \int_{(\mathbb{R}^d)^k} \psi(x_1,\ldots,x_k)\left\{\prod_{i=1}^k w(x_i)\right\}\mathrm{d}x_1\cdots\mathrm{d}x_k, \tag{5.59}$$

with ψ continuous and integrable on $(\mathbb{R}^d)^k$ (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 (5.58) to this more general class of functions. Then in Chapter 7 of [35], it is proved that, when $\sigma = 0$, any solution to (5.58) is dual, through the set of functional relations (1.26), to a system of independent Brownian motions with variance parameter $u\Gamma_R$, in which particles never coalesce. This is easily modified to $\sigma > 0$, in which case particles branch into two at rate $u\sigma V_R$, independently of each other. Since the set of all test functions of the form (5.59) is separating by Lemma 2.1(c) in [49], we can proceed as in the proof of Proposition 4.4.7 in [21] to conclude that the system of equations (5.58) 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 1.11(ii).

When $d = 1$, we follow the same route and use Itô's Formula to extend (5.58) to functions of the product form $\prod_{i=1}^{k}\langle\cdot, f_i\rangle$ and then to the full class of functions (5.59) by the same density argument as before. Again in Chapter 7 of [35], it is proved that in one dimension and when $\sigma = 0$, any solution to these equations is dual, through the set of relations (1.26), 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 (independently of the other pairs). As earlier, this is easily modified to cover the case $\sigma > 0$, by imposing that particles 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 (5.58) 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 1.11(i) is proved.

Remark 5.2. Liang's notation is very different from ours. To see that his process (with selection added and the coalescence rate multiplied by $u^2 V_R^2$) 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 $\chi(\kappa) = \mathbb{1}_0(\kappa) = \rho(\kappa)$ and $\psi(x) = f(x), \phi(x) = g(x)$ in Proposition 7.2 in [35] indeed leads to

$$\mathrm{d}[\mathcal{Z}^f, \mathcal{Z}^g]_t = u^2 V_R^2 \int_{\mathbb{R}^d} w_t^\infty(x) f(x) g(x) \mathrm{d}x - u^2 V_R^2 \int_{\mathbb{R}^d} \left(w_t^\infty(x)\right)^2 f(x) g(x) \mathrm{d}x$$

$$= u^2 V_R^2 \int_{\mathbb{R}^d} w_t^\infty(x) \left(1 - w_t^\infty(x)\right) f(x) g(x) \mathrm{d}x. \tag{5.60}$$

5.2 Proof of Theorem 1.13.

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 $(\Xi^n_t)_{n\geq 1}$, is rather more involved and will be broken into a number of smaller steps.

Recall that Ξ^n is defined on the probability space $(\mathbf{\Omega}, \mathcal{F}', \mathbf{P})$ and takes its values in the set $\mathcal{M}_p(\mathbb{R}^d)$ of all finite point measures on $\mathbb{R}^d$, which we have endowed with the topology of weak convergence. The linear hull of the set of test functions (recall (1.13))

$$\Phi_{\exp, \ln f} : \Xi \mapsto \prod_{i=1}^{|\Xi|} f(\xi^i) = \exp\left\{\int_{\mathbb{R}^d} (\ln f(x)) \Xi(\mathrm{d}x)\right\}, \tag{5.61}$$

where $f \in C^1(\mathbb{R}^d)$ takes values in $[0, 1]$, is dense in $C_0(\mathcal{M}_p(\mathbb{R}^d))$ (the space of continuous functions on $\mathcal{M}_p(\mathbb{R}^d)$ tending to 0 at "infinity") for the topology of the uniform convergence (*cf.* Lemma 0.2 in [27], where the formalism is different but the result is equivalent to our claim). Consequently, the linear span of the set of functions (5.61) is dense in $C_b(\mathcal{M}_p(\mathbb{R}^d))$ for the topology of uniform convergence over compact sets, and functions of the above form will thus be sufficient to characterise the law of an $\mathcal{M}_p(\mathbb{R}^d)$-valued random variable. In this section, the atoms of the point measures considered will be viewed as particles evolving in $\mathbb{R}^d$.

We start with the following result.

Lemma 5.3. *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 1.13. In particular, the only possible limit point for the sequence $(\Xi^n)_{n\geq 1}$ is Ξ^∞.*

Proof of Lemma 5.3. 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 $\psi_i \in C_c(\mathbb{R}^d)$ being a probability density function on $\mathbb{R}^d$ for every i.

Let us write $(M_t^{(n)})_{t \geq 0}$ for the unscaled SLFVS with parameters s_n, u_n (in the fixed radius case), and $w_t^{(n)}$ for a representative of the density of $M_t^{(n)}$, for every $t \geq 0$. Recall the notation $(M_t^n)_{t \geq 0}$ for the scaled process whose density at time t is $w_{nt}^{(n)}(n^{1/3}\cdot)$. Let $w^0 \in C^1(\mathbb{R}^d)$, and suppose that $M_0^{(n)}$ is such that M_0^n has density w^0 for every $n \geq 1$. With this initial condition and Relation (5.23) (where Δf can be replaced by any function $f \in C_c^2(\mathbb{R}^d)$), it is easy to check that $\overline{M}_0^n$, as defined in Theorem 1.11, converges to the measure $M^0 \in \mathcal{M}_\lambda$ with density w^0 as $n \to \infty$. Hence, by Theorem 1.11 the sequence of processes $(\overline{M}^n)_{n \geq 1}$ converges weakly to M^∞ starting at M^0. Using the approximation (5.23) 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, together with Fubini's Theorem and the duality formula (1.26), we have

$$\mathbb{E}_{M_0^{(n)}}\left[\prod_{i=1}^k \left(\int_{\mathbb{R}^d \times \{0,1\}} \psi_i(x_i) \mathbb{1}_{\{0\}}(\kappa_i) \overline{M}_t^n(\mathrm{d}x_i, \mathrm{d}\kappa_i) \right) \right]$$

$$= \mathbb{E}_{M_0^{(n)}}\left[\int_{(\mathbb{R}^d)^k} \psi_1(x_1) \cdots \psi_k(x_k) \left\{ \prod_{i=1}^k \overline{w}_t^n(x_i) \right\} \mathrm{d}x_1 \ldots \mathrm{d}x_k \right]$$

$$= \mathbb{E}_{M_0^{(n)}}\left[\int_{(\mathbb{R}^d)^k} \psi_1(x_1) \cdots \psi_k(x_k) \left\{ \prod_{i=1}^k w_{nt}^{(n)}(n^{1/3}x_i) \right\} \mathrm{d}x_1 \ldots \mathrm{d}x_k \right] + \mathcal{O}(n^{-2/3})$$

$$= n^{-dk/3} \mathbb{E}_{M_0^{(n)}}\left[\int_{(\mathbb{R}^d)^k} \psi_1(n^{-1/3}x_1) \cdots \psi_k(n^{-1/3}x_k) \left\{ \prod_{i=1}^k w_{nt}^{(n)}(x_i) \right\} \mathrm{d}x_1 \ldots \mathrm{d}x_k \right]$$
$$+ \mathcal{O}(n^{-2/3})$$

$$= n^{-dk/3} \int_{(\mathbb{R}^d)^k} \psi_1(n^{-1/3}x_1) \cdots \psi_k(n^{-1/3}x_k) \mathbf{E}_{\Xi[x_1,\ldots,x_k]}\left[\prod_{j=1}^{N_{nt}} w_0^{(n)}(\xi_{nt}^j) \right] \mathrm{d}x_1 \ldots \mathrm{d}x_k$$
$$+ \mathcal{O}(n^{-2/3})$$

$$= \int_{(\mathbb{R}^d)^k} \psi_1(x_1) \cdots \psi_k(x_k) \mathbf{E}_{\Xi[n^{1/3}x_1,\ldots,n^{1/3}x_k]}\left[\prod_{j=1}^{N_{nt}} w_0^{(n)}(n^{1/3}(n^{-1/3}\xi_{nt}^j)) \right] \mathrm{d}x_1 \ldots \mathrm{d}x_k$$
$$+ \mathcal{O}(n^{-2/3})$$

$$= \int_{(\mathbb{R}^d)^k} \psi_1(x_1) \cdots \psi_k(x_k) \mathbf{E}_{\Xi[x_1,\ldots,x_k]}\left[\prod_{j=1}^{N_t^n} w_0^n(\xi_t^{n,j}) \right] \mathrm{d}x_1 \ldots \mathrm{d}x_k + \mathcal{O}(n^{-2/3})$$

$$= \mathbf{E}_{\Xi_0^n}\left[\prod_{j=1}^{N_t^n} w^0(\xi_t^{n,j}) \right] + \mathcal{O}(n^{-2/3}). \tag{5.62}$$

Now the expression on the l.h.s. of (5.62) converges to the corresponding expression for M^∞ as $n \to \infty$. Therefore, if Ξ is the limit of a subsequence of $(\Xi^n)_{n \geq 1}$, then for every $t \geq 0$

$$\mathbb{E}_{M^0}\left[\prod_{i=1}^k \left(\int_{\mathbb{R}^d \times \{0,1\}} \psi_i(x_i) \mathbb{1}_{\{0\}}(\kappa_i) M_t^\infty(\mathrm{d}x_i, \mathrm{d}\kappa_i) \right) \right] = \mathbf{E}_{\Xi_0}\left[\prod_{j=1}^{N_t^\infty} w^0(\xi_t^j) \right]. \tag{5.63}$$

On the other hand, as explained in Point **3)** of the proof of Theorem 1.11, the same equality (5.63) holds for any w^0 if we replace Ξ_t in the r.h.s. by the empirical distribution at time t, Ξ_t^∞, of the system of independent branching (and in dimension 1, coalescing) Brownian motions described in Theorem 1.13. As mentioned in the paragraph around (5.61), test functions of the form used in the r.h.s. of (5.63) are separating. 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 (1.26) 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

$$x \mapsto \psi_1(x_1)\cdots\psi_k(x_k), \tag{5.64}$$

with $\psi_i \in C_c(\mathbb{R}^d)$ a probability density function for every i, are dense (for the $\mathbb{L}^1$ norm) in the set of probability densities ψ on $(\mathbb{R}^d)^k$, an analogue of Relation (5.63) can be established for this more general class of initial densities ψ. The same chain of arguments is then sufficient to conclude the proof of Lemma 5.3. □

Tightness

We now show tightness of the sequence $(\Xi^n)_{n\geq 1}$. To ease the notation, we write $\mathbf{P}_\psi$ 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 first show that the compact containment condition holds if we see $(\Xi^n)_{n\geq 1}$ as a sequence of $\mathcal{M}_p(\widehat{\mathbb{R}^d})$-valued Markov processes, where $\widehat{\mathbb{R}^d}$ is the one-point compactification of $\mathbb{R}^d$. We can then use Theorem 3.9.1 in [21], together with the fact that the linear span of functions of the form (5.61) is dense in $C_b(\mathcal{M}_p(\widehat{\mathbb{R}^d}))$ for the topology of uniform convergence on compact sets to reduce the tightness of $(\Xi^n)_{n\geq 1}$ to that of $(\Phi_{\exp,\ln f}(\Xi^n))_{n\geq 1}$ for every $f \in C^\infty(\widehat{\mathbb{R}^d})$ with values in $[0,1]$. More precisely, we show that for every such f, every $T > 0$, every sequence of stopping times $(\tau_n)_{n\geq 1}$ bounded by T and every $\varepsilon > 0$, there exists $\delta = \delta(f, T, \psi, \varepsilon)$ such that

$$\limsup_{n\to\infty} \mathbf{P}_\psi\left[\sup_{0\leq t\leq\delta}\left|\prod_{i=1}^{N_{\tau_n+t}^n} f(\xi_{\tau_n+t}^{n,i}) - \prod_{i=1}^{N_{\tau_n}^n} f(\xi_{\tau_n}^{n,i})\right| > \varepsilon\right] \leq \varepsilon. \tag{5.65}$$

(This is actually stronger than the classical Aldous criterion based on stopping times [2], which considers the supremum over $t \in [0,\delta]$ of the probability that the increment between times τ_n and $\tau_n + t$ is larger than ε.) Finally, using Lemma 5.3 and Corollary 3.9.3 in [21], we shall be able to conclude that $(\Xi^n)_{n\geq 1}$ is tight in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0,\infty)$, as desired (and furthermore that Ξ^n converges weakly to Ξ^∞ in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0,\infty)$).

We shall proceed in a number of steps. First we control the maximum number of particles in Ξ_t^n up to time $T + 1$. Not only does this give us the compact containment condition, but conditional on this result, it is then easy to control the probability that there is a branch in an interval of length δ (by branch, we mean that a particle is replaced by two "parental" particles during a selective event). If we can also show that with high probability there is no coalescence (*i.e.*, no group of at least two particles is ever removed during the same event and replaced by one or two "parental" particles), 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 5.5, is showing that indeed there is no accumulation of coalescence events.

Let us replace $\mathbb{R}^d$ by its one-point compactification $\widehat{\mathbb{R}^d}$, so that the set of finite point measures with a total mass less than K is compact for every $K > 0$. Recall the notation $|\Xi| = \langle\Xi, 1\rangle$ for the total mass of the measure Ξ. The following lemma thus implies the compact containment condition.

Lemma 5.4. *Let $T > 0$. Given $\varepsilon > 0$, there exists $K > 0$ such that for every $n \geq 1$,*

$$\mathbf{P}_\psi\left[\sup_{0\leq t\leq T+1}|\Xi_t^n| > K\right] \leq \frac{\varepsilon}{4}. \tag{5.66}$$

Proof of Lemma 5.4. 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 $n s_n V_R u_n = u \sigma V_R$. 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 $(\Xi^n_t)_{t \geq 0}$ is stochastically bounded by the number of particles in a Yule process in which particles split (independently of one another) into two offspring at rate $u \sigma V_R$. Let $T > 0$. Since the initial value, Ξ^n_0, has $k < \infty$ particles, we conclude that there exists $K \in \mathbb{N}$ such that for every $n \geq 1$,

$$\mathbf{P}_\psi \left[\sup_{0 \leq t \leq T+1} |\Xi^n_t| > K \right] \leq \frac{\varepsilon}{4}, \tag{5.67}$$

as required. $\qquad\square$

From now on, all our calculations proceed conditional on the event

$$A_n = \left\{ \sup_{t \in [0, T+1]} |\Xi^n_t| \leq K \right\}. \tag{5.68}$$

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

$$K\, \mathbf{P}_\psi \left[\text{a given particle branches in } (t, t+\delta] \right] \leq K\left(1 - e^{-u\sigma V_R \delta}\right) \leq u\sigma K V_R\, \delta. \tag{5.69}$$

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

$$\mathbf{P}_\psi \left[\text{at least 1 particle created in } (\tau_n, \tau_n + \delta_1]\,;\, A_n \right] \leq \frac{\varepsilon}{4}. \tag{5.70}$$

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 5.5. *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*

$$\mathbf{P}_\psi \left[\text{at least 1 coalescence in } (\tau_n, \tau_n + \delta_2]\,;\, A_n, B^c_{\delta_2} \right] \leq \frac{\varepsilon}{4}. \tag{5.71}$$

Before proving Proposition 5.5, 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" (*i.e.*, is removed and replaced by another particle seen as its parent) at rate $n u_n V_R (1 + s_n) = n^{2/3} u V_R (1 + o(1))$, 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 $R n^{-1/3}$, and so the length of the jump is bounded by $2 R n^{-1/3}$. 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,

$$\mathbf{P}_\psi \left[\sup_{t \in [0, s]} \left| \xi^{n,1}_{T_n + t} - \xi^{n,1}_{T_n} \right| > \eta \right] \leq \frac{C_1}{\eta^2}\, s, \tag{5.72}$$

where we have used the strong Markov property of $\xi^{n,1}$ at time T_n. From this, we can draw two conclusions. The first one, which is not necessary for the rest of the proof but gives some nice insight on our sequence of processes, is that taking $s = T$ and $T_n = 0$, we can find a compact set $E \subset \mathbb{R}^d$ such that for every $n \geq 1$,

$$\mathbf{P}_\psi \left[\sup_{t \in [0, T]} \Xi^n_t(E^c) > 0\,;\, A_n \right] \leq \varepsilon. \tag{5.73}$$

Indeed, since ψ is integrable, there exists a compact set $\tilde{E}$ such that $\mathbf{P}_\psi[\Xi_0^n(\tilde{E}^c) > 0] < \varepsilon/2$. Conditionally on all the initial particles belonging to $\tilde{E}$, by (5.72) 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 $\varepsilon/2$.

Second, conditional on the number of individuals not changing during a time interval of length δ, we can index the particles of $\Xi_{\tau_n}^n$ and $\Xi_{\tau_n+\delta}^n$ by a common indexing set which we denote I_n, in such a way that a particle in $\Xi_{\tau_n+\delta}^n$ has the same label as a particle in $\Xi_{\tau_n}^n$ if and only if the position of the former can be seen as the result of a (potentially empty) series of jumps carried out by the latter during $(\tau_n, \tau_n+\delta]$. Under this assumption, for any $f \in C^\infty(\widehat{\mathbb{R}^d})$ with values in $[0, 1]$, a Taylor expansion yields

$$\left| \prod_{i \in I_n} f(\xi_{\tau_n+t}^{n,i}) - \prod_{i \in I_n} f(\xi_{\tau_n}^{n,i}) \right| \leq C \|\nabla f\| \sum_{i \in I_n} \left| \xi_{\tau_n+t}^{n,i} - \xi_{\tau_n}^{n,i} \right|, \tag{5.74}$$

for some $C > 1$, where the sup norm of ∇f is finite since $\widehat{\mathbb{R}^d}$ is compact. Together with (5.72) and the choice $s = \delta$, $T_n = \tau_n$ and $\eta = \varepsilon/(KC\|\nabla f\|)$, this shows that there exists $\delta_3 \in (0, \delta_2]$ such that for n large enough, writing C_δ^c for the event that there is no coalescence in $(\tau_n, \tau_n + \delta]$,

$$\mathbf{P}_\psi \left[\sup_{t \in [0, \delta_3]} \left| \prod_{i \in I_n} f(\xi_{\tau_n+t}^{n,i}) - \prod_{i \in I_n} f(\xi_{\tau_n}^{n,i}) \right| > \varepsilon \,;\, A_n, B_{\delta_3}^c, C_{\delta_3}^c \right] \leq \frac{KC_1}{\eta^2} \delta_3 \leq \frac{\varepsilon}{4}. \tag{5.75}$$

Combining Lemma 5.4, (5.70), Proposition 5.5 and (5.75), we obtain (5.65) with $\delta = \delta_3$.

It remains to prove Proposition 5.5. Let us remark that it is not enough to consider particles 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 5.5. 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 $\varepsilon/(2K(K-1))$, irrespective of their initial separation. Once this bound has been established, we can write

$$\mathbf{P}_\psi \left[\text{at least 1 coalescence in } (\tau_n, \tau_n + \delta_2) \,;\, A_n, B_{\delta_2}^c \right] \leq \frac{K(K-1)}{2} \frac{\varepsilon}{2K(K-1)} = \frac{\varepsilon}{4}, \tag{5.76}$$

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 particle 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 $\tilde{\xi}_t^{n,i} = \xi_{n^{1/3}t}^i$, $i \in \{1, 2\}$.

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 particle 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 particle 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 infinitesimal generator of

$((\tilde{\xi}_t^{n,1}, \tilde{\xi}_t^{n,2}))_{t\geq 0}$, applied to any function $\phi \in C_0((\widehat{\mathbb{R}^d})^2)$ (the space of continuous functions on $(\widehat{\mathbb{R}^d})^2$ vanishing at infinity) takes the form

$$u(1+s_n)\int_{B(\tilde{\xi}^1,R)\setminus B(\tilde{\xi}^2,R)}\int_{B(x,R)}\frac{1}{V_R}\left(\phi(z,\tilde{\xi}^2)-\phi(\tilde{\xi}^1,\tilde{\xi}^2)\right)\mathrm{d}z\mathrm{d}x$$

$$+u(1+s_n)\int_{B(\tilde{\xi}^2,R)\setminus B(\tilde{\xi}^1,R)}\int_{B(x,R)}\frac{1}{V_R}\left(\phi(\tilde{\xi}^1,z)-\phi(\tilde{\xi}^1,\tilde{\xi}^2)\right)\mathrm{d}z\mathrm{d}x$$

$$+u(1-un^{-1/3})(1+s_n)\int_{B(\tilde{\xi}^1,R)\cap B(\tilde{\xi}^2,R)}\int_{B(x,R)}\frac{1}{V_R}\left(\phi(z,\tilde{\xi}^2)+\phi(\tilde{\xi}^1,z)-2\phi(\tilde{\xi}^1,\tilde{\xi}^2)\right)\mathrm{d}z\mathrm{d}x$$

$$+u^2n^{-1/3}(1+s_n)\int_{B(\tilde{\xi}^1,R)\cap B(\tilde{\xi}^2,R)}\int_{B(x,R)}\frac{1}{V_R}\left(\phi(z,z)-\phi(\tilde{\xi}^1,\tilde{\xi}^2)\right)\mathrm{d}z\mathrm{d}x. \tag{5.77}$$

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

$$\mathbf{P}_\psi\left[\tilde{T} > \delta n^{2/3}\right] = \mathbf{E}_\psi\left[\exp\left\{-\frac{u^2(1+s_n)}{n^{1/3}}\int_0^{\delta n^{2/3}} V_R(0,\hat{\xi}_s^{n,1}-\hat{\xi}_s^{n,2})\,\mathrm{d}s\right\}\right], \tag{5.78}$$

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

$$\sigma_1^n = \inf\{t \geq 0 : |\hat{\xi}_t^{n,1}-\hat{\xi}_t^{n,2}| \leq 2R\}, \qquad \tau_1^n = \inf\{t \geq \sigma_1^n : |\hat{\xi}_t^{n,1}-\hat{\xi}_t^{n,2}| > 2R\}, \tag{5.79}$$

and for every $k \geq 1$,

$$\sigma_k^n = \inf\{t \geq \tau_{k-1}^n : |\hat{\xi}_t^{n,1}-\hat{\xi}_t^{n,2}| \leq 2R\}, \qquad \tau_k^n = \inf\{t \geq \sigma_k^n : |\hat{\xi}_t^{n,1}-\hat{\xi}_t^{n,2}| > 2R\}. \tag{5.80}$$

Now, we have the following result.

Lemma 5.6. *There exists $\mathcal{C} > 0$ such that for every $n, k \geq 1$,*

$$\mathbf{E}_\psi\left[\tau_k^n - \sigma_k^n\right] \leq \mathcal{C}. \tag{5.81}$$

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 5.6 is similar to that of Lemma 6.6 in [7] (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 5.6 and the fact that $V_R(0, \cdot)$ is bounded, this shows that for every $\eta > 0$,

$$\lim_{n \to \infty} \mathbf{P}_\psi \left[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2}) \, \mathrm{d}s > \eta \frac{n^{1/3}}{u^2(1 + s_n)} \right] = 0 \qquad (5.82)$$

As a consequence, coming back to (5.78), observing that

$$\mathbf{P}_\psi \left[\tilde{T} \le \delta n^{2/3} \right] = \mathbf{P}_\psi \left[\mathrm{Exp}(1) \le \frac{u^2(1 + s_n)}{n^{1/3}} \int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2}) \, \mathrm{d}s \right] \qquad (5.83)$$

(where $\mathrm{Exp}(1)$ denotes an exponential random variable with parameter 1) and choosing η small enough that $\mathbf{P}[\mathrm{Exp}(1) \le \eta] \le \varepsilon/(2K(K-1))$, we can conclude that for any $\delta > 0$,

$$\limsup_{n \to \infty} \mathbf{P}_\psi \left[\tilde{T} \le \delta n^{2/3} \right] \le \frac{\varepsilon}{2K(K-1)}. \qquad (5.84)$$

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

$$\mathbf{P}_{\{x_1, x_2\}} \left[\sigma_1^n > \delta n^{2/3} \right] \ge \frac{\mathcal{C}'}{\log(\delta n^{2/3})}, \qquad (5.85)$$

where we have written $\mathbf{P}_{\{x_1, x_2\}}$ 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 [9], and so we only sketch the main ideas. We can a.s. embed the trajectories of the difference process $\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}$ into the trajectories of a two-dimensional Brownian motion, in the same spirit as Skorokhod's embedding in one dimension (see *e.g.* [11]). Now, since the jumps of the difference process (when outside $B(0, 2R)$) are rotationally invariant, we have

$$\inf_{|x_1 - x_2| > 2R} \mathbf{P}_{\{x_1, x_2\}} \left[\hat{\xi}^{n,1} - \hat{\xi}^{n,2} \text{ leaves } B(0, 4R) \text{ before entering } B(0, 2R) \right] > 0, \quad (5.86)$$

and the result then follows from that for Brownian motion, namely Theorem 2 in [45] applied with $a = 2R$ and $r \ge 4R$. As a consequence, the number N_E^n 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

$$\mathbf{P}_\psi \left[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2}) \, \mathrm{d}s > \eta \frac{n^{1/3}}{u^2(1 + s_n)} \right]$$

$$\le \mathbf{P}_\psi \left[N_E^n > C_E^n \log(\delta n^{2/3}) \right] + \mathbf{P}_\psi \left[\sum_{k=1}^{\lceil C_E^n \log(\delta n^{2/3}) \rceil} (\tau_k^n - \sigma_k^n) > \eta \frac{n^{1/3}}{u^2(1 + s_n) V_R} \right]$$

$$\le e^{-C_E^n \mathcal{C}'} + \frac{u^2(1 + s_n) V_R}{\eta n^{1/3}} C_E^n \log(\delta n^{2/3}) \mathcal{C}, \qquad (5.87)$$

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

$$\lim_{n \to \infty} \mathbf{P}_\psi \left[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2}) \, \mathrm{d}s > \eta \frac{n^{1/3}}{u^2(1 + s_n)} \right] = 0, \qquad (5.88)$$

and we conclude as in (5.84).

d = 1: Finally, when $d = 1$ it is shown in [41] that there exists $\mathcal{C}' > 0$ such that for every x_1, x_2 such that $|x_1 - x_2| > 2R$,

$$\mathbf{P}_{\{x_1,x_2\}}[\sigma_1^n > \delta n^{2/3}] \geq \frac{\mathcal{C}'}{\sqrt{\delta}\, n^{1/3}}. \tag{5.89}$$

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

$$\mathbf{P}_\psi\left[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2})\,\mathrm{d}s > \eta\,\frac{n^{1/3}}{u^2(1+s_n)}\right]$$

$$\leq \mathbf{P}_\psi\left[N_E^n > C_E^n \sqrt{\delta} n^{1/3}\right] + \mathbf{P}_\psi\left[\sum_{k=1}^{\lceil C_E^n \sqrt{\delta} n^{1/3}\rceil} (\tau_k^n - \sigma_k^n) > \eta\,\frac{n^{1/3}}{u^2(1+s_n)\|V_R\|}\right]$$

$$\leq e^{-C_E^n \mathcal{C}'} + \frac{u^2(1+s_n)\|V_R\|}{\eta n^{1/3}}\, C_E^n \sqrt{\delta} n^{1/3} \mathcal{C}. \tag{5.90}$$

Choosing C_E^n to be a constant large enough for the first term to be less than $\varepsilon/(6K(K-1))$, and then $\delta_3 > 0$ small enough for the second term to be less than $\varepsilon/(6K(K-1))$, and finally taking η small enough, we obtain that for any $\delta \leq \delta_3$,

$$\mathbf{P}_\psi\left[\tilde{T} \leq \delta n^{2/3}\right] \leq \mathbf{P}_\psi\left[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2})\,\mathrm{d}s > \eta\,\frac{n^{1/3}}{u^2(1+s_n)}\right] + \mathbf{P}[\mathrm{Exp}(1) \leq \eta]$$

$$\leq \frac{\varepsilon}{6K(K-1)} + \frac{\varepsilon}{6K(K-1)} + \frac{\varepsilon}{6K(K-1)} = \frac{\varepsilon}{2K(K-1)}. \tag{5.91}$$

We have now proved the desired bound for the probability of a coalescence in any dimension and the proof of Proposition 5.5 is complete. $\qquad\square$

6 Convergence of the rescaled SLFVS and its dual – the stable radius case

We proceed exactly as for the case of fixed radius.

6.1 Proof of Theorem 1.14

As in the proof of Theorem 1.11, we first show that the sequence $(\overline{M}^n)_{n\geq 1}$ is tight in $D_{\mathcal{M}_\lambda}[0, \infty)$, then we show that any limit point of a subsequence satisfies the martingale problem stated in Theorem 1.14, and finally we prove that there exists at most one solution in $D_{\mathcal{M}_\lambda}[0, \infty)$ to this martingale problem to conclude that $(\overline{M}^n)_{n\geq 1}$ indeed converges to it.

1) Tightness.

We use the same method as in the proof of Theorem 1.11, but the computations required are different. Note that for every $n \geq 1$ the process $\overline{M}^n$ has sample paths in $D_{\mathcal{M}_\lambda}[0, \infty)$, since the unscaled process from which it is constructed has a.s. càdlàg paths by Theorem 1.2. Using again Theorem 3.9.1 in [21] and the compactness of $\mathcal{M}_\lambda$, we reduce the proof of tightness of $(\overline{M}^n)_{n\geq 1}$ to the proof of tightness of $(\Psi_{F,f}(\overline{M}^n))_{n\geq 1}$ for every $F \in C^3(\mathbb{R})$ and $f \in C_c^\infty(\mathbb{R}^d)$.

Hence, let us now fix F and f as above. Since $\Psi_{F,f}$ is a bounded function on $\mathcal{M}_\lambda$ and consequently $(\Psi_{F,f}(\overline{M}_t^n))_{n\geq 1}$ is a tight sequence for every $t \geq 0$, by the Aldous-Rebolledo criterion we only have to prove the equivalent of (5.1) and (5.2) after finding an expression for the predictable finite variation $\mathcal{A}^n$ of $(\Psi_{F,f}(\overline{M}_t^n))_{t\geq 0}$ and for its predictable quadratic variation $\mathcal{Q}^n$. As earlier, we first replace f by any function $\varphi \in C_c^\infty(\mathbb{R}^d)$ and then specialise the formulae we derive to a suitably chosen function φ_f to conclude.

For any given $n \geq 1$, the extended generator of the unscaled process with parameters satisfying (1.27), (1.28), (1.29) and (1.31), acting on functions of the form $\Psi_{F,\varphi}$, is given by

$$\mathcal{L}\Psi_{F,\varphi}(M) = \int_{\mathbb{R}^d} \int_1^\infty \int_{B(x,r)^2} \frac{1}{V_r^2} \left\{ w(y)(1 + s_n w(z)) \big[F(\langle \Theta_{x,r,u_n}^+(w), \varphi \rangle) - F(\langle w, \varphi \rangle) \big] \right. \tag{6.1}$$

$$\left. + (1 - w(y) + s_n(1 - w(y)w(z))) \big[F(\langle \Theta_{x,r,u_n}^-(w), \varphi \rangle) - F(\langle w, \varphi \rangle) \big] \right\} \mathrm{d}y\mathrm{d}z\mu(\mathrm{d}r)\mathrm{d}x,$$

where, as usual now, w is a representative of the density of M. Arguing as in the part on tightness of the proof of Theorem 1.11 and using (6.1) with F and F^2, we obtain that the predictable finite variation part of $(\Psi_{F,\varphi}(M_t))_{t\geq 0}$ is given at any time $t \geq 0$ by

$$\mathcal{A}_t = \int_0^t \mathcal{L}\Psi_{F,\varphi}(M_s)\,\mathrm{d}s, \tag{6.2}$$

and its predictable quadratic variation by

$$\mathcal{Q}_t = \int_0^t \int_{\mathbb{R}^d} \int_1^\infty \int_{B(x,r)^2} \frac{1}{V_r^2} \left\{ w_s(y)(1 + s_n w_s(z)) \big[F(\langle \Theta_{x,r,u_n}^+(w_s), \varphi \rangle) - F(\langle w_s, \varphi \rangle) \big]^2 \right.$$

$$+ (1 - w_s(y) + s_n(1 - w_s(y)w_s(z))) \big[F(\langle \Theta_{x,r,u_n}^-(w_s), \varphi \rangle)$$

$$\left. - F(\langle w_s, \varphi \rangle) \big]^2 \right\} \mathrm{d}y\mathrm{d}z\mu(\mathrm{d}r)\mathrm{d}x\mathrm{d}s. \tag{6.3}$$

To make the expressions easier to read, below we retain the notation β, γ and δ from (1.33). Let us now consider the process $(M_t^n)_{t\geq 0}$ whose density at time t is $w_t^n(\cdot) := w_{nt}(n^\beta \cdot)$. Writing explicitly the martingale problem satisfied by M^n and performing a change in the time and space variables, we obtain that the extended generator of this Markov process is given by

$$\mathcal{L}^n \Psi_{F,\varphi}(M) \tag{6.4}$$

$$= n \int_{\mathbb{R}^d} \int_1^\infty \int_{B(x,r)^2} \frac{1}{V_r^2} \left\{ w(n^{-\beta}y)\big(1 + s_n w(n^{-\beta}z)\big) \right.$$

$$\times \big[F(\langle \Theta_{n^{-\beta}x,n^{-\beta}r,u_n}^+(w), \varphi \rangle) - F(\langle w, \varphi \rangle) \big] \big)$$

$$\left. + (1 - w(n^{-\beta}y) + s_n(1 - w(n^{-\beta}y)w(n^{-\beta}z))) \big[F(\langle \Theta_{n^{-\beta}x,n^{-\beta}r,u_n}^-(w), \varphi \rangle) - F(\langle w, \varphi \rangle) \big] \right\}$$

$$\mathrm{d}y\mathrm{d}z\mu(\mathrm{d}r)\mathrm{d}x$$

$$= n^{1-\beta\alpha} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r^2} \left\{ w(y)(1 + s_n w(z)) \big[F(\langle \Theta_{x,r,u_n}^+(w), \varphi \rangle) - F(\langle w, \varphi \rangle) \big] \right.$$

$$\left. + \big(1 - w(y) + s_n(1 - w(y)w(z))\big) \big[F(\langle \Theta_{x,r,u_n}^-(w), \varphi \rangle) - F(\langle w, \varphi \rangle) \big] \right\} \mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x,$$

which allows us to write as in the fixed radius case that the predictable finite variation part of $\Psi_{F,\varphi}(M^n)$ is equal to $(\int_0^t \mathcal{L}^n \Psi_{F,\varphi}(M_s^n)\mathrm{d}s)_{t\geq 0}$, while its predictable quadratic variation is given by the integral with respect to time of the function in (6.4) (applied to M_s^n) in which the increments $[F(\langle \Theta_{x,r,u_n}^\pm(w_s^n), \varphi \rangle) - F(\langle w_s^n, \varphi \rangle)]$ are squared. It remains to apply these results to φ_f defined by

$$\varphi_f(x) = \frac{n^{d\beta}}{V_1} \int_{B(x,n^{-\beta})} f(y)\,\mathrm{d}y, \tag{6.5}$$

and to use the fact that for every $t \geq 0$,

$$\Psi_{F,\varphi_f}(M_t^n) = F(\langle w_t^n, \varphi_f \rangle) = F(\langle \overline{w}_t^n, f \rangle) = \Psi_{F,f}(\overline{M}_t^n) \tag{6.6}$$

(where $\overline{w}_t^n$ is the density of $\overline{M}_t^n$ defined in (1.32)), to identify $\mathcal{A}^n$ as

$$\mathcal{A}_t^n = \int_0^t \mathcal{L}^n \Psi_{F,\varphi_f}(M_s^n)\,\mathrm{d}s, \tag{6.7}$$

and $\mathcal{Q}^n$ as

$$\mathcal{Q}_t^n = n^{1-\beta\alpha} \int_0^t \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r^2} \Big\{ w_s^n(y)(1 + s_n w_s^n(z)) \big[F(\langle \Theta_{x,r,u_n}^+(w_s^n), \varphi_f\rangle)$$
$$- F(\langle \overline{w}_s^n, f\rangle) \big]^2$$
$$+ \big(1 - w_s^n(y) + s_n(1 - w_s^n(y)w_s^n(z))\big) \big[F(\langle \Theta_{x,r,u_n}^-(w_s^n), \varphi_f\rangle)$$
$$- F(\langle \overline{w}_s^n, f\rangle) \big]^2 \Big\} \mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\mathrm{d}s. \tag{6.8}$$

Now that we have an expression for $\mathcal{A}^n$ and $\mathcal{Q}^n$, let us bound their increments to complete the proof of tightness of $(\Psi_{F,f}(\overline{M}^n))_{n\geq 1}$. We start with $\mathcal{A}^n$. As before, it is convenient to split $\mathcal{L}^n \Psi_{F,\varphi_f}(M_s^n)$ into its neutral and selective components. Using a Taylor expansion of the function F, we obtain that the neutral part is equal to

$$n^{1-\beta\alpha} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} \frac{1}{V_r} \Big[w_s^n(y)\langle \Theta_{x,r,u_n}^+(w_s^n) - w_s^n, \varphi_f\rangle$$
$$+ (1 - w_s^n(y))\langle \Theta_{x,r,u_n}^-(w_s^n) - w_s^n, \varphi_f\rangle \Big] \mathrm{d}y\mathrm{d}r\mathrm{d}x$$
$$+ n^{1-\beta\alpha} \frac{F''(\langle \overline{w}_s^n, f\rangle)}{2} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} \frac{1}{V_r} \Big[w_s^n(y)\langle \Theta_{x,r,u_n}^+(w_s^n) - w_s^n, \varphi_f\rangle^2$$
$$+ (1 - w_s^n(y))\langle \Theta_{x,r,u_n}^-(w_s^n) - w_s^n, \varphi_f\rangle^2 \Big] \mathrm{d}y\mathrm{d}r\mathrm{d}x + \varepsilon_n$$
$$= n^{1-\beta\alpha-\gamma} u F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r} w_s^n(y)(\varphi_f(z) - \varphi_f(y))\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x$$
$$+ n^{1-\beta\alpha-2\gamma} \frac{u^2}{2} F''(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} \frac{1}{V_r} \Big\{ w_s^n(y)\langle \mathbb{1}_{B(x,r)}(1 - w_s^n), \varphi_f\rangle^2$$
$$+ (1 - w_s^n(y))\langle \mathbb{1}_{B(x,r)} w_s^n, \varphi_f\rangle^2 \Big\} \mathrm{d}y\mathrm{d}r\mathrm{d}x + \varepsilon_n, \tag{6.9}$$

with

$$|\varepsilon_n| \leq n^{1-\alpha\beta-3\gamma} \frac{u^3 C_F}{3!} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} \frac{2}{V_r} \langle \mathbb{1}_{B(x,r)}, \varphi_f\rangle^3 \mathrm{d}y\mathrm{d}r\mathrm{d}x, \tag{6.10}$$

where the constant C_F is the supremum of $F^{(3)}$ over the bounded set in which its argument takes its values (recall that $\varphi_f \in C_c^\infty(\mathbb{R}^d)$). Consider the first term on the right hand side of (6.9). 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 φ_f and a symmetry argument to cancel the integral of $(z - y)\mathrm{d}z$, we obtain that

$$\left| u F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \int_{n^{-\beta}}^1 \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r} w_s^n(y)(\varphi_f(z) - \varphi_f(y))\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x \right| \tag{6.11}$$
$$\leq C \left| \int_{\mathbb{R}^d} \int_{n^{-\beta}}^1 \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r} |z - y|^2 \mathbb{1}_{\{B(x,r)\cap S_{\varphi_f}\neq\emptyset\}}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x \right|$$
$$\leq C'\mathrm{Vol}(S_{\varphi_f} + B(0,1)) \int_{n^{-\beta}}^1 \frac{1}{r^{d+1+\alpha}} r^{d+2}\mathrm{d}r = C''(1 - n^{-\beta(2-\alpha)})$$

for some constants $C, C', C'' > 0$ (where S_{φ_f} denotes the compact support of φ_f). To control the integral over radii in $[1, \infty)$, the cruder bound $|\varphi_f(y) - \varphi_f(z)| \leq 2\|f\|$ suffices

and, using the fact that

$$\text{Vol}\{x : S_{\varphi_f} \cap B(x,r) \neq \emptyset\} \leq C_2(r^d \vee 1), \tag{6.12}$$

we have

$$\left| uF'(\langle \overline{w}_s^n, f \rangle) \int_{\mathbb{R}^d} \int_1^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r} w_s^n(y)(\varphi_f(z) - \varphi_f(y)) \mathrm{d}y \mathrm{d}z \mathrm{d}r \mathrm{d}x \right|$$

$$\leq C \int_{\mathbb{R}^d} \int_1^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r} \left(\mathbb{1}_{\{y \in S_{\varphi_f}\}} + \mathbb{1}_{\{z \in S_{\varphi_f}\}} \right) \mathrm{d}y \mathrm{d}z \mathrm{d}r \mathrm{d}x$$

$$\leq C' \int_1^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \mathbb{1}_{\{B(x,r) \cap S_{\varphi_f} \neq \emptyset\}} \text{Vol}(S_{\varphi_f}) \mathrm{d}x \mathrm{d}r$$

$$\leq C'' \int_1^\infty \frac{1}{r^{d+1+\alpha}} r^d \mathrm{d}r \leq C''', \tag{6.13}$$

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

To control the second term on the right hand side of (6.9), we use (6.12) together with the inequality

$$|\langle \mathbb{1}_{B(x,r)} w_s^n, \varphi_f \rangle| \leq \|f\| \text{Vol}(S_{\varphi_f} \cap B(x,r)) \leq C_1 \|f\| (r^d \wedge 1), \tag{6.14}$$

to see that it is bounded by

$$n^{-\gamma} \frac{u^2 C_F}{2} \times 2C_1^2 \|f\|^2 \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} (r^d \wedge 1)^2 \mathbb{1}_{\{S_{\varphi_f} \cap B(x,r) \neq \emptyset\}} \mathrm{d}r \mathrm{d}x$$

$$= C_3 n^{-\gamma} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} (r^d \wedge 1)^2 (r^d \vee 1) \mathrm{d}r$$

$$= C_3 n^{-\gamma} \int_{n^{-\beta}}^1 \frac{r^{2d}}{r^{d+1+\alpha}} \mathrm{d}r + C_3 n^{-\gamma} \int_1^\infty \frac{r^d}{r^{d+1+\alpha}} \mathrm{d}r = C_4 n^{-\gamma} \left(1 - n^{-\beta(d-\alpha)} \right). \tag{6.15}$$

When $d \geq 2$, $d - \alpha > 0$ 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_s^n, as $n \to \infty$. As a consequence, in any dimension the absolute value of the neutral term of $\mathcal{L}^n \Psi_{F,\varphi_f}(M)$ 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 (i.e., that involving s_n) of $\mathcal{L}^n \Psi_{F,\varphi_f}(M_s^n)$ is bounded by (recall that $1 - \alpha\beta - \gamma = 0$)

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

$$\leq Cn^{-\delta} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \left(1 \wedge r^d \right) \mathbb{1}_{\{B(x,r) \cap S_{\varphi_f} \neq \emptyset\}} \mathrm{d}x \mathrm{d}r$$

$$\leq C'n^{-\delta} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \left(1 \wedge r^d \right) \left(1 \vee r^d \right) \mathrm{d}r \leq C'' n^{-\delta+\alpha\beta} = C'', \tag{6.16}$$

since $\alpha\beta - \delta = 0$. Combining (6.9), (6.11), (6.13), (6.15) and (6.16), we obtain that there exists a constant C independent of n such that for every $0 \leq t_1 \leq t_2$,

$$\int_{t_1}^{t_2} \left| \mathcal{L}^n \Psi_{F,\varphi_f}(M_s^n) \right| \mathrm{d}s \leq C(t_2 - t_1), \tag{6.17}$$

and therefore for every $T > 0$, every sequence of stopping times $(\tau_n)_{n \geq 1}$ bounded by T, and every $\varepsilon > 0$, we can choose $\eta > 0$ small enough so that

$$\limsup_{n \to \infty} \sup_{\theta \in [0,\eta]} \mathbb{P} \left[\left| \mathcal{A}_{\tau_n+\theta}^n - \mathcal{A}_{\tau_n}^n \right| > \varepsilon \right] = 0, \tag{6.18}$$

which corresponds to the first part of the Aldous-Rebolledo criterion.

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

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

$$\leq C'n^{-\gamma} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \left(1 \wedge r^d\right)^2 \mathbb{1}_{\{B(x,r)\cap S_{\varphi_f} \neq \emptyset\}} \mathrm{d}r\mathrm{d}x$$

$$\leq C''n^{-\gamma} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \left(1 \wedge r^d\right)^2 \left(1 \vee r^d\right)\mathrm{d}r \leq C'''n^{-\gamma}\left(1 + n^{-\beta(d-\alpha)}\right), \quad (6.19)$$

which is bounded by a constant independent of n. As before, we conclude that the equivalent of (6.18) with $\mathcal{A}^n$ replaced by $\mathcal{Q}^n$ is satisfied for $\eta > 0$ small enough. The Aldous-Rebolledo criterion allows us to conclude that the sequence of real-valued processes $(\Psi_{F,f}(\overline{M}^n))_{n\geq 1}$ is tight, and since this is true for every $F \in C^3(\mathbb{R})$ and $f \in C_c^\infty(\mathbb{R}^d)$, we obtain the tightness of $(\overline{M}^n)_{n\geq 1}$ in $D_{\mathcal{M}_\lambda}[0,\infty)$.

2) Identifying the limit.

Suppose $M^\infty \in D_{\mathcal{M}_\lambda}[0,\infty)$ is the weak limit of a subsequence $(\overline{M}^{n_k})_{k\geq 1}$, and for every $t \geq 0$, write w_t^∞ for a representative of the density of M_t^∞. We know from the previous paragraph that for every $f \in C_c^\infty(\mathbb{R}^d)$ and every $n \geq 1$,

$$\left(\Psi_{\mathrm{Id},f}(\overline{M}_t^n) - \Psi_{\mathrm{Id},f}(\overline{M}_0^n) - \int_0^t \mathcal{L}^n \Psi_{\mathrm{Id},\varphi_f}(M_s^n)\mathrm{d}s\right)_{t\geq 0} \quad (6.20)$$

is a martingale with predictable quadratic variation (6.8) (with $F = \mathrm{Id}$), where $\mathcal{L}^n$ was defined in (6.4) and φ_f in (6.5). As in the fixed radius case, we first show that for every $t \geq 0$,

$$\lim_{k\to\infty} \mathbb{E}\left[\left|\int_0^t \mathcal{L}^{n_k}\Psi_{\mathrm{Id},\varphi_f}(M_s^{n_k})\mathrm{d}s - \int_0^t \left\{\langle w_s^\infty, \mathcal{D}^\alpha f\rangle - \frac{2u\sigma}{\alpha}\langle w_s^\infty(1-w_s^\infty), f\rangle\right\}\mathrm{d}s\right|\right] = 0, \quad (6.21)$$

so that we can then use the fact that the quantity in (6.20) is a martingale, the fact that $\Psi_{\mathrm{Id},f}$ is a bounded continuous function and the Dominated Convergence Theorem to conclude that for every $0 \leq t < t'$, $m \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_m \leq t$ and $h_1, \ldots, h_m \in C_b(\mathcal{M}_\lambda)$,

$$\mathbb{E}\left[\left(\langle w_{t'}^\infty, f\rangle - \langle w_t^\infty, f\rangle - \int_t^{t'}\left\{\langle w_s^\infty, \mathcal{D}^\alpha f\rangle - \frac{2u\sigma}{\alpha}\langle w_s^\infty(1-w_s^\infty), f\rangle\right\}\mathrm{d}s\right)\right.$$
$$\left.\times\left(\prod_{i=1}^m h_i(M_{t_i}^\infty)\right)\right] = 0 \quad (6.22)$$

and consequently that $\mathcal{Z}^f$ is a martingale (with respect to the natural filtration of M^∞). In the case $d \geq 2$ this property is again sufficient to conclude, since we showed in (6.19) that the quadratic variation of the martingale (6.20) tended to 0 as $n \to \infty$, and therefore the limit $\mathcal{Z}^f$ is the constant process equal to 0. We shall thus end this point **2)** by showing that in one dimension, the quadratic variation of $\mathcal{Z}^f$ and the bracket process between $\mathcal{Z}^f$ and $\mathcal{Z}^g$ have the required form.

Let us fix $f \in C_c^\infty(\mathbb{R}^d)$ and show (6.21). Let us first analyse the part of $\mathcal{L}^n\Psi_{\mathrm{Id},\varphi_f}(M_s^n)$ corresponding to neutral events. By (6.9) with $F = \mathrm{Id}$, since $1 - \alpha\beta - \gamma = 0$ this neutral part takes to form

$$un^{1-\alpha\beta-\gamma} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^2} \frac{1}{V_r} w_s^n(y)(\varphi_f(z) - \varphi_f(y))\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x$$

$$= u\int_{\mathbb{R}^d} w_s^n(y)\int_{\mathbb{R}^d}\left(\int_{n^{-\beta}\vee\frac{|z-y|}{2}}^{\infty} \frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\,\mathrm{d}r\right)(\varphi_f(z) - \varphi_f(y))\,\mathrm{d}z\mathrm{d}y, \quad (6.23)$$

where $V_r(y, z)$ is again the volume of the intersection $B(y, r) \cap B(z, r)$. Now, a simple Taylor expansion to the second order gives us that

$$\varphi_f(z) - \varphi_f(y) = f(z) - f(y) + \mathcal{O}(n^{-2\beta})\big(\mathbb{1}_{\{B_n(z) \cap S_f \neq \emptyset\}} + \mathbb{1}_{\{B_n(y) \cap S_f \neq \emptyset\}}\big), \qquad (6.24)$$

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

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

$$\leq C n^{-2\beta} \int_{S_f + B_n(0)} \int_{\mathbb{R}^d} \left(n^{-\beta} \vee \frac{|z-y|}{2} \right)^{-d-\alpha} \mathrm{d}z \mathrm{d}y \leq C' n^{-\beta(2-\alpha)} \to 0 \qquad (6.25)$$

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

$$u \int_{\mathbb{R}^d} w_s^n(y) \int_{\mathbb{R}^d} \left(\int_{n^{-\beta} \vee \frac{|z-y|}{2}}^{\infty} \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,z)}{V_r} \, \mathrm{d}r \right) (f(z) - f(y)) \, \mathrm{d}z \mathrm{d}y. \qquad (6.26)$$

Now, our computations (6.11) and (6.13) in the proof of tightness imply that the function

$$a_n(y) : y \mapsto \int_{\mathbb{R}^d} \left(\int_{n^{-\beta} \vee \frac{|z-y|}{2}}^{\infty} \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,r)}{V_r} \, \mathrm{d}r \right) (f(z) - f(y)) \, \mathrm{d}z \qquad (6.27)$$

is a continuous function, uniformly bounded in y and n. Hence, up to a vanishing error term we can first replace w_s^n by $\overline{w}_s^n$ in (6.26) and, second, use dominated convergence to pass to the limit as $n \to \infty$ in (6.26), along the converging subsequence. Doing so, and using the fact that all the error terms go to 0 uniformly in s, we obtain that the limit in $\mathbb{L}^1$ norm of the neutral term in $\int_0^t \mathcal{L}^{n_k} \Psi_{\mathrm{Id}, \varphi_f}(M_s^{n_k}) \mathrm{d}s$ is equal to

$$u \int_0^t \int_{\mathbb{R}^d} w_s^\infty(y) \int_{\mathbb{R}^d} \Phi(|z-y|)(f(z) - f(y)) \mathrm{d}z \mathrm{d}y \mathrm{d}s, \qquad (6.28)$$

where, as in (1.36),

$$\Phi(|z-y|) := \int_{\frac{|z-y|}{2}}^{\infty} \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,z)}{V_r} \, \mathrm{d}r. \qquad (6.29)$$

In passing, let us show the following property of the operator we obtain in the limit.

Lemma 6.1. *For $f \in C_c^\infty(\mathbb{R}^d)$ write*

$$\mathcal{D}^\alpha f(y) = u \int_{\mathbb{R}^d} \Phi(|z-y|)(f(z) - f(y)) \mathrm{d}z. \qquad (6.30)$$

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

Proof of Lemma 6.1. It is reassuring to first check that this is the generator of a well-defined Lévy process:

$$\int_{\mathbb{R}^d} (1 \wedge |y|^2) \int_0^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0,y)}{V_r} \, \mathrm{d}r \, \mathrm{d}y$$

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

$$\leq C'' \int_0^1 \frac{r^{d+2}}{r^{d+1+\alpha}} \mathrm{d}r + C' \int_1^\infty \frac{1}{r^{d+1+\alpha}} \mathrm{d}r < \infty, \qquad (6.31)$$

since $\alpha \in (1,2)$. Therefore the measure

$$\nu_\alpha(\mathrm{d}y) = \int_0^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0,y)}{V_r} \, \mathrm{d}r \, \mathrm{d}y \tag{6.32}$$

is a Lévy measure and there exists a unique Lévy process with values in $\mathbb{R}^d$ whose Lévy triplet is $(0,0,\nu_\alpha)$. By Theorem 6.8 in [30], the operator $\mathcal{D}^\alpha$ is its infinitesimal generator.

To verify that the associated Lévy process is a symmetric stable process, we check the scaling property (the symmetry property is obvious from the form of ν_α). Let $b > 0$. The generator of $(b^{-1/\alpha}\zeta_{bt})_{t\geq 0}$ is given by

$$\begin{aligned}
\mathcal{D}_b^\alpha f(y) &= bu \int_{\mathbb{R}^d} \Phi(|z - b^{1/\alpha}y|)(f(b^{-1/\alpha}z) - f(y)) \, \mathrm{d}z \\
&= ub^{1+d/\alpha} \int_{\mathbb{R}^d} \Phi(|b^{1/\alpha}z - b^{1/\alpha}y|)(f(z) - f(y)) \, \mathrm{d}z.
\end{aligned} \tag{6.33}$$

But a simple change of variables gives us that

$$\begin{aligned}
\Phi(|b^{1/\alpha}z - b^{1/\alpha}y|) &= \int_{\frac{b^{1/\alpha}|z-y|}{2}}^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(b^{1/\alpha}y, b^{1/\alpha}z)}{V_r} \, \mathrm{d}r \\
&= b^{-1-d/\alpha} \int_{\frac{|z-y|}{2}}^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,z)}{V_r} \, \mathrm{d}r,
\end{aligned} \tag{6.34}$$

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

Having identified the neutral part of the limit, we now turn to the part of $\mathcal{L}^n \Psi_{\mathrm{Id},\varphi_f}(M_s^n)$ corresponding to the selective events. It is given by

$$u\sigma n^{1-\beta\alpha-\gamma-\delta} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^3} \frac{1}{V_r^2} (w_s^n(y)w_s^n(z) - w_s^n(z'))\varphi_f(z') \, \mathrm{d}y\mathrm{d}z\mathrm{d}z'\mathrm{d}r\mathrm{d}x. \tag{6.35}$$

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

$$\begin{aligned}
u\sigma n^{-\delta} &\int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} w_s^n(z')\varphi_f(z') \, \mathrm{d}z'\mathrm{d}r\mathrm{d}x \\
&= u\sigma n^{-\delta} \int_{\mathbb{R}^d} w_s^n(z')\varphi_f(z') \left(\int_{n^{-\beta}}^\infty \frac{V_1 r^d}{r^{d+1+\alpha}} \, \mathrm{d}r \right) \mathrm{d}z' = \frac{u\sigma V_1}{\alpha} \langle \overline{w}_s^n, f \rangle,
\end{aligned} \tag{6.36}$$

where the last equality uses the fact that $\alpha\beta - \delta = 0$. It is then straightforward to obtain that

$$\lim_{k\to\infty} \frac{u\sigma V_1}{\alpha} \mathbb{E}\left[\left| \int_0^t \langle \overline{w}_s^{n_k}, f \rangle \mathrm{d}s - \int_0^t \langle w_s^\infty, f \rangle \mathrm{d}s \right| \right] = 0. \tag{6.37}$$

Similar calculations show that the "quadratic" term in (6.35) is equal to

$$u\sigma n^{-\delta} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^{n^{-\beta}\log n} \frac{1}{r^{d+1+\alpha}} \left(\int_{B(x,r)} \frac{1}{V_r} w_s^n(y)\mathrm{d}y \right)^2 \int_{B(x,r)} \varphi_f(z)\mathrm{d}z\mathrm{d}r\mathrm{d}x + \mathcal{O}\big((\log n)^{-\alpha}\big). \tag{6.38}$$

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 for every $t \geq 0$, $(\overline{w}_t^{n_k})^2$ converges to $(w_t^\infty)^2$ in the appropriate sense.

Concerning the first point, we have

$$\left(\int_{B(x,r)} \frac{1}{V_r} w_s^n(y) \mathrm{d}y \right)^2$$

$$= \left(\int_{B(x,r)} \frac{1}{V_r} w_s^n(y) \mathrm{d}y + \overline{w}_s^n(x) \right) \left(\int_{B(x,r)} \frac{1}{V_r} w_s^n(y) \mathrm{d}y - \overline{w}_s^n(x) \right) + \overline{w}_s^n(x)^2. \qquad (6.39)$$

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

Lemma 6.2. *Under the conditions of Theorem 1.14, for every $r \in [n^{-\beta}, n^{-\beta} \log n]$,*

$$\lim_{n \to \infty} \mathbb{E}\left[\left| \int_{B(x,r)} \frac{w_s^n(y)}{V_r} \, \mathrm{d}y - \overline{w}_s^n(x) \right| \right] = 0 \qquad (6.40)$$

uniformly in $x \in \mathbb{R}^d$ and uniformly in s over compact time intervals $[0, t]$.

From this result, we can conclude from a dominated convergence argument and a Taylor expansion of φ_f that the "quadratic" part of (6.35) is equal to

$$u\sigma n^{-\delta} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^{n^{-\beta} \log n} \frac{V_r}{r^{d+1+\alpha}} \overline{w}_s^n(x)^2 f(x) \, \mathrm{d}r\mathrm{d}x + \epsilon_n = \frac{u\sigma V_1}{\alpha} \langle (\overline{w}_s^n)^2, f \rangle + \epsilon_n, \qquad (6.41)$$

where ϵ_n tends to zero as $n \to \infty$ uniformly in s over compact intervals of time.

As concerns the second point, we proceed as in (5.40) and below. Using Proposition B.1(ii) in Appendix B, the facts that the support of f is bounded, and that p_ε is supported in $B(0, \varepsilon)$ (so that τ_2 in (B.4) 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 (5.40) of $\langle (\overline{w}_s^n)^2, f \rangle$ is bounded by a constant (independent of n, ε) times

$$n^{-a} + \varepsilon^{\alpha/(d+1)} + \varepsilon^{1/4} + \varepsilon^{\alpha/(2d+2)} + n^{-\beta(d-1)}\varepsilon^{(\alpha-d)/(2d+2)}. \qquad (6.42)$$

Letting n tend to infinity in the above expression, we can write that the third term in the decomposition (5.40) is bounded by a constant times

$$\varepsilon^{\alpha/(d+1)} + \varepsilon^{1/4} + \varepsilon^{\alpha/(2d+2)} + \varepsilon^{(\alpha-1)/4}\mathbf{1}_{\{d=1\}}. \qquad (6.43)$$

Finally, the second term in the decomposition (5.40) tends to 0 by the assumption that $\overline{M}^{n_k}$ converges to M^∞. As in the fixed radius case, we can therefore conclude from (6.41) that

$$\lim_{k \to \infty} \mathbb{E}\left[\left| \int_0^t \left\{ \frac{u\sigma V_1}{\alpha} \langle (\overline{w}_s^{n_k})^2, f \rangle + \epsilon_n \right\} \mathrm{d}s - \frac{u\sigma V_1}{\alpha} \int_0^t \langle (w_s^\infty)^2, f \rangle \mathrm{d}s \right| \right] = 0. \qquad (6.44)$$

(Note that this convergence is independent of the representatives of the different densities that we choose.)

Combining (6.28), (6.37) and (6.44), we obtain (6.21) and we can therefore conclude that $\mathcal{Z}^f$ is a martingale with respect to the natural filtration of M^∞. As we already mentioned, when $d \geq 2$ this is sufficient to conclude that M^∞ satisfies the equations stated in Theorem 1.14(ii).

To identify the quadratic variation of $\mathcal{Z}^f$ and the bracket process between $\mathcal{Z}^f$ and $\mathcal{Z}^g$ when $d = 1$, we proceed exactly as in the fixed radius case and therefore we do not provide all the details. Setting

$$W_t^n(f) := \langle \overline{w}_t^n, f \rangle - \langle \overline{w}_0^n, f \rangle - \int_0^t \mathcal{L}^n \Psi_{\mathrm{Id}, \varphi_f}(M_s^n)\mathrm{d}s, \qquad t \geq 0, \qquad (6.45)$$

we know from the paragraph **1)** on tightness that for every $n \geq 1$, $W^n(f)$ is a zero-mean martingale with predictable quadratic variation $\mathcal{Q}^n$ given in (6.8) (with $F = \mathrm{Id}$). As a consequence, for every $n \geq 1$, $0 \leq t < t'$, $m \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_m \leq t$ and $h_1, \ldots, h_m \in C_b(\mathcal{M}_\lambda)$,

$$\mathbb{E}\left[\left(\left(W^n_{t'}(f) \right)^2 - \left(W^n_t(f) \right)^2 - \mathcal{Q}^n_{t'} + \mathcal{Q}^n_t \right) \left(\prod_{i=1}^m h_i\big(\overline{M}^n_{t_i}\big) \right) \right] = 0. \qquad (6.46)$$

Recall from our calculations in **1)** that all summands in the expression for $W^n_t(f)$ are bounded uniformly in $n \geq 1$ and t in a compact time interval. Furthermore, the same calculations as those we performed to obtain the limit of the selection part of $\mathcal{L}^n \Psi_{\mathrm{Id},\varphi_f}(M^n_s)$ (see in particular (6.37) and (6.44)) show that

$$\lim_{k \to \infty} \mathbb{E}\left[\left| \mathcal{Q}^{n_k}_t - \frac{4u^2}{\alpha - 1} \int_0^t \langle w_s^\infty (1 - w_s^\infty), f^2 \rangle \mathrm{d}s \right| \right] = 0, \qquad (6.47)$$

uniformly over compact intervals of time. Letting $n \to \infty$ in (6.46) along the converging subsequence, we arrive at

$$\mathbb{E}\left[\left(\left(\mathcal{Z}^f_{t'} \right)^2 - \left(\mathcal{Z}^f_t \right)^2 - \frac{4u^2}{\alpha - 1} \int_t^{t'} \langle w_s^\infty (1 - w_s^\infty), f^2 \rangle \mathrm{d}s \right) \left(\prod_{i=1}^m h_i\big(M^\infty_{t_i}\big) \right) \right] = 0. \qquad (6.48)$$

This allows us to identify the predictable quadratic variation of the martingale $\mathcal{Z}^f$ as

$$[\mathcal{Z}^f]_t = \frac{4u^2}{\alpha - 1} \int_0^t \langle w_s^\infty (1 - w_s^\infty), f^2 \rangle \mathrm{d}s, \qquad t \geq 0. \qquad (6.49)$$

By the analogue of (5.14) (with $n^{-1/3}$ replaced by $n^{-\beta}$), every jump of $W^n(f)$ is bounded by $u_n \mathrm{Vol}(S_{\varphi_f})$ independently of the size of the radius of the event, where we recall that $u_n = u n^{-\gamma}$. Consequently, $\mathcal{Z}^f$ has a.s. continuous trajectories. Since $\mathcal{Z}^f_0 = 0$, we can use the Dubins-Schwarz Theorem (or rather its extension since we do not know whether $[\mathcal{Z}^f]_\infty = +\infty$, see Remark 5.1) to conclude that $\mathcal{Z}^f$ is a time-changed Brownian motion, solution to the stochastic differential equation

$$dW_t = \frac{2u}{\sqrt{\alpha - 1}} \sqrt{\langle w_t^\infty (1 - w_t^\infty), f^2 \rangle}\, dB^f_t, \qquad (6.50)$$

where B^f denotes standard Brownian motion. The bracket process between $\mathcal{Z}^f$ and $\mathcal{Z}^g$ is then obtained by the same kind of calculations, writing first the bracket process for a fixed n and then identifying the limit by letting $n_k \to \infty$. We thus obtain that any limit of a subsequence of $(\overline{M}^n)_{n \geq 1}$ satisfies the set of equations stated in Theorem 1.14(i).

3) Uniqueness of the limit.

The argument is exactly the same as in the corresponding part of the proof of Theorem 1.11. Indeed, by another modification of the results of Chapter 7 in [35] (replacing Brownian motion by the symmetric α-stable process $(\zeta_t)_{t \geq 0}$ generated by $\mathcal{D}^\alpha$ – see Lemma 6.1), we obtain that any solution to the limiting system of equations stated in Theorem 1.14 is dual through the set of relations (1.26) to a system of particles following independent symmetric α-stable processes (with the same law as ζ), and branching independently at rate $u\sigma V_1/\alpha$ into two particles starting at the location of their parent. In one dimension, each pair of particles also coalesces at a rate $4u^2/(\alpha - 1)$ times the local time at zero of their separation, independently of the other pairs. Since the set of all test functions of the form (5.59) is separating, we can again conclude that there is at most one solution to the system of equations of Theorem 1.14. Hence, this solution exists and the full sequence $(\overline{M}^n)_{n \geq 0}$ converges to it in $D_{\mathcal{M}_\lambda}[0, \infty)$.

6.2 Proof of Theorem 1.16

Most of the proof is identical to that of Theorem 1.13. 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 [35], in which the only change is that Brownian motion is replaced by the stable process generated by $\mathcal{D}^\alpha$ (see (6.30)) and we have added natural selection/branching of particles. This gives us the analogue of Lemma 5.3 in the case of stable radii, whose proof is exactly the same as that of Lemma 5.3.

Lemma 6.3. *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 α-stable motions Ξ^∞, described in the statement of Theorem 1.16. In particular, the only possible limit point for the sequence $(\Xi^n)_{n \geq 1}$ is Ξ^∞.*

Next, we have to show that the sequence $(\Xi^n)_{n \geq 1}$ is tight. Let again $\mathbf{P}_\psi$ denote 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 1.13, after showing that the compact containment condition holds if we replace $\mathbb{R}^d$ by its one-point compactification $\widehat{\mathbb{R}^d}$ and consider each Ξ^n as taking its values in $\mathcal{M}_p(\widehat{\mathbb{R}^d})$, we shall use Theorem 3.9.1 in [21] to deduce the tightness of $(\Xi^n)_{n \geq 1}$ in $D_{\mathcal{M}_p(\widehat{\mathbb{R}^d})}[0, \infty)$ from the tightness of $(\Phi_{\exp, \ln f}(\Xi^n))_{n \geq 1}$ in $D_{[0,1]}[0, \infty)$ for every $f \in C^\infty(\widehat{\mathbb{R}^d})$ with values in $[0, 1]$. More precisely, we show that for any such function f, every $T > 0$, every sequence of stopping times $(\tau_n)_{n \geq 1}$ bounded by T and every $\varepsilon > 0$, there exists $\delta = \delta(f, T, \psi, \varepsilon) > 0$ such that

$$\limsup_{n \to \infty} \mathbf{P}_\psi \left[\sup_{0 \leq t \leq \delta} \left| \prod_{i=1}^{N_{\tau_n+t}^n} f\big(\xi_{\tau_n+t}^{n,i}\big) - \prod_{i=1}^{N_{\tau_n}^n} f\big(\xi_{\tau_n}^{n,i}\big) \right| > \varepsilon \right] \leq \varepsilon. \tag{6.51}$$

Once these properties have been shown, we can use Lemma 6.3 and Corollary 3.9.3 in [21] to conclude that $(\Xi^n)_{n \geq 1}$ is tight in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0, \infty)$ and converges to Ξ^∞.

Again, we proceed in four steps. First, by exactly the same arguments as in the proof of Theorem 1.13, for every $T > 0$ and every $\varepsilon > 0$, there exists $K > 0$ such that for every $n \in \mathbb{N}$ we have

$$\mathbf{P}_\psi[A_n] := \mathbf{P}_\psi \left[\sup_{0 \leq s \leq T+1} |\Xi_s^n| \leq K \right] \geq 1 - \frac{\varepsilon}{4}, \tag{6.52}$$

which, in particular, grants us the compact containment condition since the set of all point measures on the compact space $\widehat{\mathbb{R}^d}$ with total mass less than K is compact. Furthermore, there exists $\delta_1 \in (0, 1)$, independent of the subinterval of $[0, T]$ considered, such that

$$\mathbf{P}_\psi \big[\text{at least 1 particle created in } (\tau_n, \tau_n + \delta_1] ; A_n \big] \leq \frac{\varepsilon}{4}. \tag{6.53}$$

As before, the difficulty will be to control the coalescence (*i.e.*, the events in which two or more particles are removed and replaced by one or two "parental" particles), 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 $\Xi_{\tau_n}^n$. Then, exactly as before, we can write

$$\left| \prod_{i \in I_n} f\big(\xi_{\tau_n+t}^{n,i}\big) - \prod_{i \in I_n} f\big(\xi_{\tau_n}^{n,i}\big) \right| \leq C \|\nabla f\| \sum_{i \in I_n} \big| \xi_{\tau_n+t}^{n,i} - \xi_{\tau_n}^{n,i} \big|, \tag{6.54}$$

and it suffices to consider the motion of a single particle to control the evolution of the whole set of particles. 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 infinitesimal generator

$$D^n\phi(x) := u(1+s_n)\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{V_r(x,y)}{V_r} \left(\phi(y) - \phi(x)\right) \mathrm{d}y\,\mathrm{d}r, \qquad (6.55)$$

for every $\phi \in C_0(\widehat{\mathbb{R}^d})$ and $x \in \widehat{\mathbb{R}^d}$. 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 6.1 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.* [42], we have for any $\eta, \delta > 0$, and any stopping time T_n

$$\mathbf{P}_\psi\left[\sup_{t\in[0,\delta]} \left|X_{T_n+t} - X_{T_n}\right| > \eta\right] < C\frac{\delta}{\eta^\alpha} \qquad (6.56)$$

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

Since

$$\mathbf{P}_\psi\left[\sup_{t\in[0,\delta]} \left|\xi_{T_n+t}^n - \xi_{T_n}^n\right| > \eta\right]$$

$$\leq \mathbf{P}_\psi\left[\sup_{t\in[0,\delta]} \left|X_{T_n+t} - X_{T_n}\right| > \eta\right] + \mathbf{P}_\psi\left[\sup_{t\in[0,\delta]} \left|Z_{T_n+t}^n - Z_{T_n}^n\right| > \eta\right], \qquad (6.57)$$

it remains to show that

$$\mathbf{P}_\psi\left[\sup_{t\in[0,\delta]} \left|Z_{T_n+t}^n - Z_{T_n}^n\right| > \eta\right] \to 0, \qquad \text{as } n \to \infty. \qquad (6.58)$$

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

$$(1+s_n)u\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{V_r(x,y)}{V_r} \left(f(y) - f(x)\right)^2 \mathrm{d}y\mathrm{d}r$$

$$= (1+s_n)u\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{V_r(x,y)}{V_r} \left((y-x).\nabla f(x) + \mathcal{O}(|y-x|^2)\right)^2 \mathrm{d}y\mathrm{d}r$$

$$\leq C\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{B(x,2r)} |y-x|^2 \,\mathrm{d}y\mathrm{d}r = C'n^{-\beta(2-\alpha)}, \qquad (6.59)$$

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 on the right hand side vanishes by rotational symmetry. Hence, we can conclude that for any η, δ,

$$\lim_{n\to\infty} \mathbf{P}_\psi\left[\sup_{t\in[0,\delta]} \left|Z_{T_n+t}^n - Z_{T_n}^n\right| > \eta\right] = 0. \qquad (6.60)$$

Coming back to (6.57), and taking $T_n = \tau_n$ and η fixed, we can conclude that there exists $\delta_3 \in (0, \delta_2]$ such that for n large enough,

$$\mathbf{P}_\psi\left[\sup_{t\in[0,\delta_3]} \left|\xi_{\tau_n+t}^n - \xi_{\tau_n}^n\right| > \eta\right] \leq \frac{\varepsilon}{4K}. \qquad (6.61)$$

Choosing $\eta = \varepsilon/(KC\|\nabla f\|)$ and recalling (6.54), we obtain that for all sufficiently large n,

$$\mathbf{P}_\psi\left[\sup_{t\in[0,\delta_3]} \left|\prod_{i\in I_n} f\left(\xi_{\tau_n+t}^{n,i}\right) - \prod_{i\in I_n} f\left(\xi_{\tau_n}^{n,i}\right)\right| > \varepsilon\,;\, A_n, B_{\delta_3}^c, C_{\delta_3}^c\right] \leq \frac{\varepsilon}{4}, \qquad (6.62)$$

where as in the fixed radius case, B_δ^c is the event that there is no branching event in $(\tau^n, \tau^n + \delta]$ and C_δ^c is the event that there is no coalescence in $(\tau^n, \tau^n + \delta]$.

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 particles coalesce. The result will be based on the following lemma, in which we use again the interpretation of the replacement of a particle by its "parent" as a jump by this particle (or ancestral lineage – when there are two parents, we choose one of them uniformly at random).

Lemma 6.4. *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 jump process obtained by following the (unscaled) position of one particle 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

$$\limsup_{n \to \infty} \mathbf{E}_\psi \left[\frac{1}{n^\gamma} \int_0^{n^{1-\gamma} t} \frac{1}{2^\alpha \vee |\zeta^n_s|^\alpha} \, \mathrm{d}s \right] \leq C(t). \tag{6.63}$$

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

(ii) When $d \geq 2$,

$$\lim_{n \to \infty} \mathbf{E}_\psi \left[\frac{1}{n^\gamma} \int_0^{n^{1-\gamma} t} \frac{1}{2^\alpha \vee |\zeta^n_s|^\alpha} \, \mathrm{d}s \right] = 0. \tag{6.64}$$

We defer the proof of Lemma 6.4 until after the end of the proof of Theorem 1.16.

Suppose that we start with a sample of two (non independent) particles at some (unscaled) separation $z_0 \in \mathbb{R}^d$. As before, we work on the timescale n^γ so that a single particle jumps at rate $\mathcal{O}(1)$ and we suppose the two particles ξ^1 and ξ^2 are currently at locations 0 and z (in fact, only their separation matters). Then, the infinitesimal generator Γ of the difference walk $(\xi^2_{n^\gamma t} - \xi^1_{n^\gamma t})_{t \geq 0}$ (until it reaches a cemetery state Δ, say the point "infinity" in $(\widehat{\mathbb{R}^d})^2$, corresponding to the two walks having coalesced) is equal, for every given $\phi \in C_0((\widehat{\mathbb{R}^d})^2)$, to

$\Gamma \phi(z)$

$$= 2u(1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \mathbb{1}_{\{0 \notin B(x,r)\}} \mathbb{1}_{\{z \in B(x,r)\}} \frac{\mathbb{1}_{\{y \in B(x,r)\}}}{V_r} \, \mathrm{d}x \mathrm{d}r \right.$$

$$+ (1 - u_n) \int_1^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \mathbb{1}_{\{0 \in B(x,r)\}} \mathbb{1}_{\{z \in B(x,r)\}} \frac{\mathbb{1}_{\{y \in B(x,r)\}}}{V_r} \, \mathrm{d}x \mathrm{d}r \left. \right\} (\phi(y) - \phi(z)) \mathrm{d}y$$

$$+ u^2 n^{-\gamma} (1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \mathbb{1}_{\{0 \in B(x,r)\}} \mathbb{1}_{\{z \in B(x,r)\}} \frac{\mathbb{1}_{\{y \in B(x,r)\}}}{V_r} \, \mathrm{d}x \mathrm{d}r \right\}$$

$$\times (\phi(\Delta) - \phi(z)) \, \mathrm{d}y,$$

$$= 2u(1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,z)}{V_r} \, \mathrm{d}r \right\} (\phi(y) - \phi(z)) \, \mathrm{d}y$$

$$- 2u^2 n^{-\gamma} (1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0,y,z)}{V_r} \, \mathrm{d}r \right\} (\phi(y) - \phi(z)) \, \mathrm{d}y$$

$$+ u^2 n^{-\gamma} (1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0,y,z)}{V_r} \, \mathrm{d}r \right\} (\phi(\Delta) - \phi(z)) \, \mathrm{d}y, \tag{6.65}$$

where $V_r(0, y, z)$ denotes the volume of the intersection $B(0, r) \cap B(y, r) \cap B(z, r)$.

From the first two terms on the r.h.s. of (6.65), we see that 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 walk but with

each jump $z \mapsto y$ "cancelled" with probability

$$\Delta_n(z, y) = \frac{2u^2 n^{-\gamma}(1 + s_n) \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0, y, z)}{V_r} \, dr}{2u(1 + s_n) \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(y, z)}{V_r} \, dr}. \tag{6.66}$$

(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 (compare the second and third terms on the r.h.s. of (6.65)), 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 $\varepsilon/(4K(K-1))$.

Now, according to the expression on the right hand side of (6.65), when the two particles lie at separation $z \in \mathbb{R}^d$, a cancelled event occurs at instantaneous rate

$$2u^2 n^{-\gamma}(1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0, y, z)}{V_r} \, dr \right\} dy$$

$$\leq 2u^2 n^{-\gamma}(1 + s_n) \int_{\mathbb{R}^d} \int_{1 \vee \frac{|z|}{2} \vee \frac{|y|}{2}} \frac{1}{r^{d+1+\alpha}} \, dr dy = C_1 n^{-\gamma} \left(2 \vee |z| \right)^{-\alpha}. \tag{6.67}$$

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

$$\mathbf{E}_\psi \left[\exp \left\{ -\int_0^{n^{1-\gamma}t} 2u^2 n^{-\gamma}(1 + s_n) \int_{\mathbb{R}^d} \left\{ \int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0, y, \zeta_s^n)}{V_r} \, dr \right\} dy ds \right\} \right] \tag{6.68}$$

$$\geq \mathbf{E}_\psi \left[\exp \left\{ -C_1 n^{-\gamma} \int_0^{n^{1-\gamma}t} \left(2 \vee |\zeta_s^n| \right)^{-\alpha} ds \right\} \right] \geq 1 - C_1 \mathbf{E}_\psi \left[n^{-\gamma} \int_0^{n^{1-\gamma}t} \frac{ds}{\left(2 \vee |\zeta_s^n| \right)^\alpha} \right].$$

But Lemma 6.4 shows that we can indeed find $\delta_2 > 0$ such that

$$\limsup_{n \to \infty} \mathbf{E}_\psi \left[n^{-\gamma} \int_0^{n^{1-\gamma}\delta_2} \frac{ds}{\left(2 \vee |\zeta_s^n| \right)^\alpha} \right] \leq \frac{\varepsilon}{2C_1 K(K-1)}. \tag{6.69}$$

Consequently,

$$\mathbf{P}_\psi[\text{at least 1 coalescence in } (\tau_n, \tau_n + \delta_2); A_n, B_{\delta_2}^c] \leq \frac{K(K-1)}{2} \frac{\varepsilon}{2K(K-1)} = \frac{\varepsilon}{4}, \tag{6.70}$$

which was the last result we needed to complete the proof of tightness and therefore of Theorem 1.16.

Proof of Lemma 6.4. 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

$$2(1 + s_n) \left(\int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0, y)}{V_r} \, dr \right) dy. \tag{6.71}$$

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

$$2(1 + s_n) \left(\int_0^1 \frac{1}{r^{d+1+\alpha}} \frac{V_r(0, y)}{V_r} \, dr \right) dy, \tag{6.72}$$

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 (6.30), 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 $s \geq 1$,

$$\mathbf{P}_\psi\left[\frac{|Z_s^n|}{\sqrt{s}} > (\log n)^2\right] \leq Ce^{-(\log n)^2/d}. \tag{6.73}$$

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

$$\mathbf{P}_\psi\left[\frac{|Z_s^n|}{\sqrt{s}} > (\log n)^2\right] \leq d\,\mathbf{P}_\psi\left[\frac{|Z_s^{n(1)}|}{\sqrt{s}} > \frac{(\log n)^2}{d}\right] = 2d\,\mathbf{P}_\psi\left[\frac{Z_s^{n(1)}}{\sqrt{s}} > \frac{(\log n)^2}{d}\right], \tag{6.74}$$

where $Z_s^{n(1)}$ denotes the first coordinate of Z_s^n. Now, $(Z_s^{n(1)})_{s\geq 0}$ is again a symmetric Lévy process with jumps bounded by 2, and so Theorem 25.3 in [46] shows that for every $s, q \geq 0$, $\mathbb{E}[\exp(qZ_s^{n(1)})] < \infty$. In this case, it is known that the characteristic exponent Ψ^n of $(Z_s^{n(1)})_{s\geq 0}$, given here by a formula of the form

$$\Psi^n(q) = \int_{[-2,2]} \left(1 - e^{iqx} + iqx\mathbb{1}_{\{|x|<1\}}\right)\mathfrak{m}^n(\mathrm{d}x), \tag{6.75}$$

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

$$\mathbf{E}_\psi\left[e^{qZ_s^{n(1)}}\right] = e^{s\psi^n(q)}, \qquad \text{with} \quad \psi^n(q) = -\Psi^n(-iq). \tag{6.76}$$

As a consequence, the Markov inequality gives us that

$$\mathbf{P}_\psi\left[\frac{Z_s^{n(1)}}{\sqrt{s}} > \frac{(\log n)^2}{d}\right] \leq e^{-(\log n)^2/d + s\psi^n(1/\sqrt{s})}. \tag{6.77}$$

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

$$\psi^n(q) = -\int_{[-2,2]} \left(1 - [1 + qx + q^2x^2/2 + \mathcal{O}(q^3x^3)] + qx\mathbb{1}_{\{|x|<1\}}\right)\mathfrak{m}^n(\mathrm{d}x)$$

$$= q\int_{[-2,2]} x\mathbb{1}_{\{|x|\geq 1\}}\mathfrak{m}^n(\mathrm{d}x) + \frac{q^2}{2}\int_{[-2,2]} x^2\mathfrak{m}^n(\mathrm{d}x) + \mathcal{O}(q^3), \tag{6.78}$$

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 (6.74) and (6.77), this gives us (6.73).

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 $2^\alpha \wedge |\zeta_s^n|^\alpha$ by $1 \wedge |\zeta_s^n|^\alpha$ 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

$$\mathbf{E}_\psi\left[\frac{1}{n^\gamma}\int_0^{n^{1-\gamma}t} \frac{1}{1 \vee |\zeta_s^n|^\alpha}\,\mathrm{d}s\right] = \mathcal{O}(n^{a-\gamma}) + \frac{1}{n^\gamma}\int_{n^a}^{n^{1-\gamma}t} \mathbf{E}_\psi\left[\frac{1}{1 \vee |X_s - Z_s^n|^\alpha}\right]\mathrm{d}s$$

$$\leq Cn^{a-\gamma} + n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t} \mathbf{P}_\psi\left[\frac{|Z_s^n|}{\sqrt{s}} > (\log n)^2\right]\mathrm{d}s + n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t} \mathbf{E}_\psi\left[\frac{\mathbb{1}_{\{|Z_s^n|\leq(\log n)^2\sqrt{s}\}}}{1 \vee |X_s - Z_s^n|^\alpha}\right]\mathrm{d}s$$

$$\leq Cn^{a-\gamma} + Cn^{1-2\gamma}te^{-(\log n)^2/d} + n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t} \mathbf{E}_\psi\left[\frac{\mathbb{1}_{\{|Z_s^n|\leq(\log n)^2\sqrt{s}\}}}{1 \vee |X_s - Z_s^n|^\alpha}\right]\mathrm{d}s. \tag{6.79}$$

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

$$p_s^\alpha(0, x) =: p_s^\alpha(x) = s^{-d/\alpha} p_1^\alpha(xs^{-1/\alpha}) \tag{6.80}$$

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

$$0 \leq p_1^\alpha(x) \leq C_{d,\alpha}\big(1 + |x|^{d+\alpha}\big)^{-1}. \tag{6.81}$$

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

$$\mathbf{E}_\psi\left[\frac{1}{1 \vee |X_s - z|^\alpha}\right] \leq s^{-d/\alpha} \int_{\mathbb{R}^d} \frac{1}{(1 \vee |x - z|^\alpha)(1 + |xs^{-1/\alpha}|^{d+\alpha})} \, \mathrm{d}x \tag{6.82}$$

$$\leq s^{-d/\alpha} \int_{B(z,1)} \frac{1}{1 + |xs^{-1/\alpha}|^{d+\alpha}} \, \mathrm{d}x + s^{-d/\alpha} \int_{B(z,1)^c} \frac{1}{|x - z|^\alpha(1 + |xs^{-1/\alpha}|^{d+\alpha})} \, \mathrm{d}x$$

$$\leq Cs^{-d/\alpha} + C's^{-d/\alpha} \int_{B(0,s^{1/\alpha}) \setminus B(z,1)} \frac{\mathrm{d}x}{|x - z|^\alpha} + C''s^{-d/\alpha} \int_{B(0,s^{1/\alpha})^c} \frac{\mathrm{d}x}{|x - z|^\alpha |xs^{-1/\alpha}|^{d+\alpha}}.$$

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

$$|z|s^{-1/\alpha} \leq (\log n)^2 s^{\frac{1}{2} - \frac{1}{\alpha}} \leq (\log n)^2 n^{-a(2-\alpha)/(2\alpha)} \to 0, \tag{6.83}$$

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

$$C's^{-d/\alpha} \int_1^{s^{1/\alpha}} \rho^{d-1-\alpha} d\rho = C's^{-1}, \tag{6.84}$$

while the third term is bounded by

$$C''s^{-d/\alpha}s^{1+d/\alpha} \int_{s^{1/\alpha}}^\infty \rho^{d-1-2\alpha-d} d\rho = C''s^{-1}. \tag{6.85}$$

Since all the constants depend on neither z (in the range considered) nor s, we deduce that the right hand side of (6.79) is bounded by

$$C'n^{a-\gamma} + Cn^{1-2\gamma}te^{-(\log n)^2/d} + C''n^{-\gamma}\big(n^{-a(d-\alpha)/d} + \log n + \log t\big) \to 0 \qquad \text{as } n \to \infty, \tag{6.86}$$

which proves (ii).

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

$$n^{-\gamma} \int_{n^a}^{n^{1-\gamma}t} s^{-1/\alpha} \mathrm{d}s \leq Cn^{-\gamma}n^{(1-\frac{1}{\alpha})(1-\gamma)}t^{1-\frac{1}{\alpha}}. \tag{6.87}$$

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 $t \to 0$. $\qquad \square$

References

[1] Abramowitz, M. and Stegun, I. A., eds.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. *New York: Dover Publications*, 1972. MR-1225604

[2] Aldous, D.: Stopping times and tightness. *Ann. Probab.* **6**, (1978), 335–340. MR-0474446

[3] Bah, B. and Pardoux, E.: Lambda-lookdown model with selection. *Stoch. Process. Appl.* **125**, (2015), 1089–1126. MR-3303969

[4] Brunet, E. and Derrida, B.: Effect of microscopic noise on front propagation. *J. Stat. Phys.* **103**, (2001), 269–282. MR-1828730

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

[6] Biswas, N., Etheridge, A. M. and Klimek, A.: The spatial Lambda-Fleming-Viot process with fluctuating selection, arXiv:1802.08188.

[7] Barton, N. H., Etheridge, A. M. and Véber, A.: A new model for evolution in a spatial continuum. *Electron. J. Probab.* **15**, (2010), 162–216. MR-2594876

[8] Barton, N. H., Etheridge, A. M. and Véber, A.: Modelling evolution in a spatial continuum. *JSTAT*, (2013), P01002. MR-3036210

[9] Berestycki, N., Etheridge, A. M. and Véber, A.: Large-scale behaviour of the spatial Λ-Fleming-Viot process. *Ann. Inst. H. Poincaré Probab. Statist.* **49**, (2013), 374–401. MR-3088374

[10] Biler, P. and Woyczynski, W. A.: Global and exploding solutions for nonlocal quadratic evolution problems. *SIAM J. Appl. Math.* **59**, (1998), 845–869. MR-1661243

[11] Billingsley, P.: Probability and Measure. *Wiley*, New York, 1995. MR-1324786

[12] Burkholder, D. L.: Distribution function inequalities for martingales. *Ann. Probab.* **1**, (1973), 19–42. MR-0365692

[13] Conlon, J. G. and Doering, C. R.: On travelling waves for the stochastic Fisher-Kolmogorov-Petrovsky-Piscunov equation. *J. Stat. Phys.* **120**, (2005), 421–477. MR-2182316

[14] Cabré, X. and Roquejoffre, J.-M.: The influence of fractional diffusion in Fisher-KPP equations. *Communications in Mathematical Physics* **320**, (2013), 679–722. MR-3057187

[15] Etheridge, A. M.: Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ.* **80**, (2008), 121–144. MR-2433141

[16] Etheridge, A. M.: Some Mathematical Models from Population Genetics: École d'été de Probabilités de Saint-Flour XXXIX-2009. *Springer*, 2011. MR-2759587

[17] Etheridge, A. M., Freeman, N. and Penington, S.: Branching Brownian motion, mean curvature flow and the motion of hybrid zones. *Electron. J. Probab.* **22**, (2017), 103. MR-3733661

[18] Etheridge, A. M., Freeman, N., Penington, S. and Straulino, D.: Branching Brownian motion and selection in the spatial Lambda-Fleming-Viot process. *Ann. Applied Probab.* **27**, (2017), 2605–2645. MR-3719942

[19] Etheridge, A. M., Freeman, N. and Straulino, D.: The Brownian net and selection in the spatial Lambda-Fleming-Viot process. *Electron. J. Probab.* **22**, (2017), 39. MR-3646065

[20] Etheridge, A. M. and Kurtz, T. G.: Genealogical constructions of population models. *Ann. Probab.* **47**, (2019), 1827–1910. MR-3980910

[21] Ethier, S. N. and Kurtz, T. G.: Markov processes: characterization and convergence. *Wiley*, 1986. MR-0838085

[22] Evans, S. N.: Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Instit. H. Poincaré Probab. Statist.* **33**, (1997), 339–358. MR-1457055

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

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

[25] Forien, R. and Penington, S.: A central limit theorem for the spatial Lambda-Fleming-Viot process with selection. *Electron. J. Probab.* **22**, (2017), 5. MR-3613698

[26] Foucart, C.: The impact of selection in the Λ-Wright-Fisher model. *Electron. Commun. Probab.* **18**, (2013), 1–10. Erratum: *Electron. Commun. Probab.* **15**, (2014), 1–3. MR-3101637

[27] Ikeda, N., Nagasawa, M. and Watanabe, S.: Branching Markov processes I. *J. Math. Kyoto Univ.* **8**, (1968), 233–278. MR-0232439

[28] Kallenberg, O.: Random measures. *Academic Press*, London-New York-San Francisco, 1976. MR-0431373

[29] Kallenberg, O.: Foundations of Modern Probability, 2nd edition. *Springer*, New York, 2002. MR-1876169

[30] Khoshnevisan, D. and Schilling R. L.: From Lévy-Type Processes to Parabolic SPDEs. *Advanced Courses in Mathematics CRM Barcelona, Springer*, 2016. MR-3587832

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

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

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

[34] Lawler, G. F. and Limic, V.: Random Walk: a modern introduction. *Cambridge University Press*, 2010. MR-2677157

[35] Liang, R. H.: Two continuum-sites stepping stone models in population genetics with delayed coalescence. *PhD Thesis*, University of California, Berkeley, 2009. MR-2713924

[36] Miller, L. R.: Evolution of highly fecund organisms. *PhD Thesis*, University of Oxford, 2015.

[37] Mueller, C., Mytnik, L. and Quastel, J.: Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Invent. Math.* **184**, (2011), 405–453. MR-2793860

[38] Mueller, C. and Sowers, R.: Random travelling waves for the KPP equation with noise. *J. Functional Analysis* **128**, (1995), 439–498. MR-1319963

[39] Mueller, C. and Tribe, R.: Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Relat. Fields* **102**, (1995), 519–545. MR-1346264

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

[41] Port, S. C. and Stone, C. J.: Infinitely divisible processes and their potential theory. II. *Ann. Instit. Fourier* **21**, (1971), 179–265. MR-0346919

[42] Pruitt, W. E.: The growth of random walks and Lévy processes. *Ann. Probab.* **9**, (1981), 948–956. MR-0632968

[43] Rebolledo, R.: Sur l'existence de solutions à certains problèmes de semimartingales. *C.R. Acad. Sci. Paris*, (1980), 290. MR-0579985

[44] Revuz, D. and Yor, M.: Continuous martingales and Brownian motion, 3rd edition. *Springer*, 1999. MR-1725357

[45] Ridler-Rowe, C. J.: On first hitting times of some recurrent two-dimensional random walks. *Z. Wahr. verw. Geb.* **5**, (1966), 187–201. MR-0199901

[46] Sato, K.: Lévy Processes and Infinitely Divisible Distributions. *Cambridge University Press*, 1999. MR-1739520

[47] Shiga, T. and Shimizu, A.: Infinite dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* **20**, (1980), 395–416. MR-0591802

[48] Stein, E. and Weiss, G.: Introduction to Fourier Analysis on Euclidean Spaces. *Princeton University Press*, 1971. MR-0304972

[49] Véber, A. and Wakolbinger, A.: The spatial Λ-Fleming-Viot process: An event-based construction and a lookdown representation. *Ann. Instit. H. Poincaré Probab. Statist.* **51**, (2015), 570–598. MR-3335017

[50] Yule, G. U.: A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis. *Philos. Trans. Roy. Soc. London Ser. B* **213**, (1924), 21–87.

A Continuity estimates in the fixed radius case

In this section, we state the continuity estimates for the scaled measures M_T^n required in the proof of Theorem 1.11. Because their proof is an adaptation of the (long and slightly more involved) proof of Proposition B.1(ii), we do not give it here and instead refer to Appendix B. These estimates have the same flavour as the one dimensional estimates derived in [39] for the convergence of the local densities of 1's in the long range voter or contact process.

Proposition A.1. *Under the conditions of Theorem 1.11, for every $T > 0$ there exist $a, \lambda, v, C > 0$ such that for every $n \geq 1$, $z_1, z_2 \in \mathbb{R}^d$ such that $|z_1 - z_2| < 1$ and $\epsilon \in (0, 1)$,*

$$\mathbb{E}\left[\left|\frac{1}{V_\epsilon}\int_{\mathbb{R}^d} w_T^n(x)(\mathbb{1}_{\{|x-z_1|<\epsilon\}} - \mathbb{1}_{\{|x-z_2|<\epsilon\}})\mathrm{d}x\right|\right]$$
$$\leq Cn^{-a} + C\tau + C\big(|z_1 - z_2|^{1/4} + \tau^{1/2}\big)e^{\lambda(|z_1|+\epsilon)} + Cn^{-(d-1)/6}\tau^{(2-d)/4}, \quad \text{(A.1)}$$

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 $\epsilon_n \leq 1$).

B Continuity estimates in the stable radius case

Our aim in this section is to obtain some continuity estimates for the measure M_T^n (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 B.1 below is more complete than Proposition A.1. Lemma 6.2 will then follow as a corollary of item (i).

Proposition B.1. *Suppose the conditions of Theorem 1.14 are satisfied. Fix $T > 0$. Then,*

(i) *There exist $a, C > 0$ (dependent on T) such that for every $z \in \mathbb{R}^d$, $t \in [0, T]$, $n \geq 1$ and $n^{-\beta} \leq \epsilon_n < \epsilon_n' \leq 1$,*

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

where

$$\tau_1 = \tau_1(n) = n^{-\beta(2-\alpha)/(2(d+1))}. \quad \text{(B.2)}$$

(ii) *There exist $a, \lambda, C > 0$ (dependent on T) such that for every $|z_1 - z_2| < 1$, $t \in [0, T]$, $n \geq 1$ and $\epsilon \in (0, 1)$,*

$$\mathbb{E}\left[\left|\frac{1}{V_\epsilon}\int_{\mathbb{R}^d} w_t^n(x)(\mathbb{1}_{\{|x-z_1|<\epsilon\}} - \mathbb{1}_{\{|x-z_2|<\epsilon\}})\mathrm{d}x\right|\right]$$
$$\leq Cn^{-a} + C\tau_2 + C\big(|z_1 - z_2|^{1/4} + (\tau_2)^{1/2}\big)e^{\lambda(|z_1|+\epsilon)} + C\big(n^{-\beta(d-1)}\tau_2^{1-d/\alpha}\big)^{1/2}, \quad \text{(B.3)}$$

where

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

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

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

Corollary B.2. *Suppose the conditions of Theorem 1.14 are satisfied and fix $T > 0$. Then for every $|z_1 - z_2| < 1$, $t \in [0, T]$ and $\epsilon \in (0, 1)$,*

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

where C depends on T.

Before proving Proposition B.1, let us show how it implies Lemma 6.2.

Proof of Lemma 6.2. Set $\epsilon_n = n^{-\beta}$ and $\epsilon_n' \in [n^{-\beta}, n^{-\beta}\log n]$ in (i). Then

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

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

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

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 (B.1) tends to zero and the lemma follows. $\square$

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

Proof of Proposition B.1. We define for $x \in \mathbb{R}^d$,

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

$\sqcap_r^{*k}$ 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 (6.4) for the extended 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 (see the beginning of the proof of Theorem 1.14, where we can replace $\varphi \in C_c(\mathbb{R}^d)$ by $\varphi \in \mathbb{L}^1(\mathbb{R}^d)$ by a density argument): for any representative w_t^n of the density of each M_t^n, we have

$$\langle w_T^n, \varphi\rangle = \langle w_0^n, \varphi\rangle + \mathcal{M}_T^{n,\varphi} + u_n n^{1-\beta\alpha}\int_0^T \int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r^2} \quad \text{(B.5)}$$
$$\times \Big\{ w_t^n(y)(1 + s_n w_t^n(z))\langle\mathbb{1}_{B(x,r)}(1 - w_t^n), \varphi\rangle$$
$$- \big(1 - w_t^n(y) + s_n(1 - w_t^n(y)w_t^n(z))\big)\langle\mathbb{1}_{B(x,r)}w_t^n, \varphi\rangle\Big\}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\mathrm{d}t$$
$$= \langle w_0^n, \varphi\rangle + \mathcal{M}_T^{n,\varphi} + u\int_0^T\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r^2}\big\{w_t^n(y)\langle\mathbb{1}_{B(x,r)}, \varphi\rangle$$
$$- \langle\mathbb{1}_{B(x,r)}w_t^n, \varphi\rangle + s_n(w_t^n(y)w_t^n(z)\langle\mathbb{1}_{B(x,r)}, \varphi\rangle - \langle\mathbb{1}_{B(x,r)}w_t^n, \varphi\rangle)\big\}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\mathrm{d}t$$

(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 (B.5) is equal to:

$$\int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} \frac{1}{V_r} \{w_t^n(y)\langle \mathbb{1}_{B(x,r)}, \varphi\rangle - \langle \mathbb{1}_{B(x,r)} w_t^n, \varphi\rangle\} \mathrm{d}y \mathrm{d}r \mathrm{d}x$$

$$= \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \frac{1}{V_r} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|x-y|\leq r\}} \mathbb{1}_{\{|x-z|\leq r\}} \{w_t^n(y)\varphi(z) - w_t^n(z)\varphi(z)\} \mathrm{d}z \mathrm{d}y \mathrm{d}r \mathrm{d}x$$

$$= \int_{n^{-\beta}}^\infty \int_{\mathbb{R}^d} \frac{V_r}{r^{d+1+\alpha}} \{(\sqcap_r^{*2} * w_t^n)(z)\varphi(z) - w_t^n(z)\varphi(z)\} \mathrm{d}z \mathrm{d}r$$

$$= \int_{\mathbb{R}^d} w_t^n(z) \int_{n^{-\beta}}^\infty \frac{V_r}{r^{d+1+\alpha}} \{(\sqcap_r^{*2} * \varphi)(z) - \varphi(z)\} \mathrm{d}r \mathrm{d}z. \tag{B.6}$$

The second term in the integrand in (B.5) is equal to

$$s_n \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} (\tilde{w}_t^n(x;r)^2 \langle \mathbb{1}_{B(x,r)}, \varphi\rangle - \langle \mathbb{1}_{B(x,r)} w_t^n, \varphi\rangle) \mathrm{d}r \mathrm{d}x$$

$$= s_n \int_{(\mathbb{R}^d)^2} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \mathbb{1}_{\{|x-y|<r\}} (\tilde{w}_t^n(x;r)^2 - w_t^n(y))\varphi(y) \mathrm{d}y \mathrm{d}x \mathrm{d}r.$$

Since $u_n^2 n^{1-\beta\alpha} = u^2 n^{-\gamma} = u^2 n^{-(\alpha-1)/(2\alpha-1)}$, the martingale term in (B.5) has predictable quadratic variation

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

$$+ \big(1 - \tilde{w}_t^n(x;r) + s_n(1 - \tilde{w}_t^n(x;r)^2)\big)\langle \mathbb{1}_{B(x,r)} w_t^n, \varphi\rangle^2 \Big\} \mathrm{d}r \mathrm{d}x \mathrm{d}t.$$

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

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

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

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

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 $\zeta_0^n(\cdot; z, \epsilon) = \frac{1}{V_\epsilon}\mathbb{1}_{\{|\cdot - z|<\epsilon\}}$. The parameter ϵ can be taken to depend on n. We observe that ν^n is radially symmetric. Let

$$a^n(x; r) = r^{-d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|x-y|<r\}} (\tilde{w}^n(y; r)^2 - w^n(x)) \mathrm{d}y$$

$$b^n(x; r) = \tilde{w}^n(x; r)(1 + s_n \tilde{w}^n(x; r))$$

$$c^n(x; r) = 1 - \tilde{w}^n(x; r) + s_n(1 - \tilde{w}^n(x; r)^2).$$

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 B.3),

then substituting in the martingale problem in the usual way, we obtain

$$\langle w_T^n, \zeta_0^n(\cdot; z, \epsilon)\rangle = \langle w_0^n, \zeta_T^n(\cdot; z, \epsilon)\rangle + \mathcal{M}_T^{n, \zeta_0^n(\cdot; z, \epsilon)}$$

$$+ us_n \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} a_t^n(x; r)\zeta_{T-t}^n(x; z, \epsilon)\mathrm{d}x\mathrm{d}r\mathrm{d}t, \qquad (B.7)$$

$$\left[\mathcal{M}^{n, \zeta_0^n(\cdot; z, \epsilon)}\right]_T = u^2 n^{-\gamma} \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \Big\{ b_t^n(x; r)\langle \mathbb{1}_{B(x,r)}(1 - w_t^n), \zeta_{T-t}^n(\cdot; z, \epsilon)\rangle^2$$

$$+ c_t^n(x; r)\langle \mathbb{1}_{B(x,r)}w_t^n, \zeta_{T-t}^n(\cdot; z, \epsilon)\rangle^2 \Big\}\mathrm{d}x\mathrm{d}r\mathrm{d}t$$

$$= u^2 n^{-\gamma} \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \Big\{ b_t^n(x; r)\left(\int_{B(x,r)}(1 - w_t^n(y))\zeta_{T-t}^n(y; z, \epsilon)\mathrm{d}y\right)^2$$

$$+ c_t^n(x; r)\left(\int_{B(x,r)} w_t^n(y)\zeta_{T-t}^n(y; z, \epsilon)\mathrm{d}y\right)^2 \Big\}\mathrm{d}x\mathrm{d}r\mathrm{d}t. \qquad (B.8)$$

In order to control the different terms appearing in (B.7) and (B.8), 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 $\frac{V_1}{A}r^{-(1+\alpha)}\mathbb{1}_{\{r>n^{-\beta}\}}\mathrm{d}r$, $Z_{1,k}$ 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

$$X_t^n = X_0^n + \sum_{k=1}^{K_t} Y_k, \qquad (B.9)$$

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,

$$q_t^{n,\{k\}}(x) = f_Y^{*k}(x)\mathbb{P}[K_t = k] = e^{-At}\frac{(At)^k}{k!}f_Y^{*k}(x)$$

$$q_t^n(x) = \sum_{k=1}^\infty q_t^{n,\{k\}}(x). \qquad (B.10)$$

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 B.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 6.1.)

Lemma B.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

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

The quantity τ_1 (*resp.*, τ_2) will be used in the bounds needed to prove Proposition B.1(*i*) (*resp.*, (*ii*)). Observe that for $t \geq \tau_2$ and $|z_1 - z_2| < 1$, the estimate on the right hand side of Lemma B.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 B.1(*i*) and (*ii*) in parallel. In both cases, we set

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

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

$$(i) \quad \langle w_T^n, \zeta_0^n(\cdot; z, \epsilon_n) - \zeta_0^n(\cdot; z, \epsilon_n') \rangle,$$
$$(ii) \quad \langle w_T^n, \zeta_0^n(\cdot; z_1, \epsilon) - \zeta_0^n(\cdot; z_2, \epsilon) \rangle$$

for the range of parameters stated in Proposition B.1, using (B.7) and (B.8).

B.1 Drift terms

Let us split the different terms into the cases in which 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):

$$\left| us_n \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} a_t^n(x;r) \big(\zeta_{T-t}^{n,[0,L)}(x;z,\epsilon_n) - \zeta_{T-t}^{n,[0,L)}(x;z,\epsilon_n') \big) \mathrm{d}x \mathrm{d}r \mathrm{d}t \right|$$

$$\leq C u s_n \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} \big(\zeta_{T-t}^{n,[0,L)}(x;z,\epsilon_n) + \zeta_{T-t}^{n,[0,L)}(x;z,\epsilon_n') \big) \mathrm{d}x \mathrm{d}r \mathrm{d}t$$

$$\leq C u s_n n^{\alpha\beta} \int_0^T \mathbb{P}[K_t < L] \, \mathrm{d}t \leq C n^{-(1-c_1)\alpha\beta} \tag{B.11}$$

by Lemma B.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

$$\zeta_{T-t}^{n,[L,\infty)}(x;z,\epsilon_n) - \zeta_{T-t}^{n,[L,\infty)}(x;z,\epsilon_n') = \int_{\mathbb{R}^d} \big(\zeta_0^n(x';z,\epsilon_n) - \zeta_0^n(x';z,\epsilon_n') \big) q_{T-t}^{n,[L,\infty)}(x-x') \mathrm{d}x'$$

$$= \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n) \big(q_{T-t}^{n,[L,\infty)}(x-x') - q_{T-t}^{n,[L,\infty)}(x-z) \big) \mathrm{d}x'$$

$$- \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n') \big(q_{T-t}^{n,[L,\infty)}(x-x') - q_{T-t}^{n,[L,\infty)}(x-z) \big) \mathrm{d}x'.$$

(The extra terms cancel since $\int_{\mathbb{R}^d} \zeta_0^n(x',z',\epsilon_n) \mathrm{d}x' = 1$ 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 B.3(iii) and (iv) (recalling also that the support of $\zeta_0^n(\cdot;z,\epsilon)$ is contained in $B(z,\epsilon)$):

$$C u s_n \int_0^{T-\tau_1} \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n) \big| q_{T-t}^{n,[L,\infty)}(x-x') - q_{T-t}^{n,[L,\infty)}(x-z) \big| \mathrm{d}x' \mathrm{d}x \mathrm{d}r \mathrm{d}t$$

$$\leq C s_n n^{\alpha\beta} \int_{\tau_1}^T \int_{B(z,\log n)} \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n) \big(t^{-(d+1)/\alpha} |z-x'|$$

$$+ C_{d,T} n^{-\beta(2-\alpha)d/(2(d+1))} \big) \mathrm{d}x' \mathrm{d}x \mathrm{d}t$$

$$+ C' s_n n^{\alpha\beta} \int_{\tau_1}^T \int_{B(z,\log n)^c} \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n) e^{-|x-z|} \mathrm{d}x' \mathrm{d}x \mathrm{d}t$$

$$\leq C \epsilon_n (\log n)^d \int_{\tau_1}^T t^{-(d+1)/\alpha} \mathrm{d}t + C_{d,T} T (\log n)^d n^{-\beta(2-\alpha)d/(2(d+1))} \big) + C' T \frac{(\log n)^{d-1}}{n}$$

$$\leq C \Big(\epsilon_n (\log n)^d \tau_1^{1-\frac{d+1}{\alpha}} + (\log n)^d n^{-\beta(2-\alpha)d/(2(d+1))} + \frac{(\log n)^{d-1}}{n} \Big). \tag{B.12}$$

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

$$\left| us_n \int_0^{T-\tau_2} \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} a_t^n(x;r) \big[\zeta_{T-t}^{n,[L,\infty)}(x;z_1,\epsilon) - \zeta_{T-t}^{n,[L,\infty)}(x;z_2,\epsilon) \big] \mathrm{d}x \mathrm{d}r \mathrm{d}t \right|$$

$$\leq C u s_n n^{\alpha\beta} \sup_{t \in [0,T-\tau_2]} \big\| \zeta_{T-t}^{n,[L,\infty)}(x;z_1,\epsilon) - \zeta_{T-t}^{n,[L,\infty)}(x;z_2,\epsilon) \big\|_\lambda \int_0^{T-\tau_2} \int_{\mathbb{R}^d} e^{-\lambda|x|} \mathrm{d}x \mathrm{d}t$$

$$\leq C_{\lambda,d,T} \big(|z_1 - z_2|^{1/2} + \tau_2 \big) e^{\lambda\epsilon} e^{\lambda|z_1|}. \tag{B.13}$$

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

$$\left| us_n \int_{T-\tau_1}^{T} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} a_t^n(x; r) \left[\zeta_{T-t}^{n,[L,\infty)}(x; z, \epsilon_n) - \zeta_{T-t}^{n,[L,\infty)}(x; z, \epsilon_n') \right] \mathrm{d}x \mathrm{d}r \mathrm{d}t \right|$$

$$\leq C s_n \int_{T-\tau_1}^{T} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} \left(\zeta_{T-t}^{n,[L,\infty)}(x; z, \epsilon_n) + \zeta_{T-t}^{n,[L,\infty)}(x; z, \epsilon_n') \right) \mathrm{d}x \mathrm{d}r \mathrm{d}t$$

$$\leq C s_n n^{\alpha\beta} \tau_1 = C \tau_1. \tag{B.14}$$

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

For the terms involving the initial condition w_0^n, similar arguments using Lemma B.3(i) and (v), and Lemma B.6 lead to

$$\left| \langle w_0^n, \zeta_T^n(\cdot; z, \epsilon_n) - \zeta_T^n(\cdot; z, \epsilon_n') \rangle \right| \leq C e^{-n^{c_1 \alpha\beta/2}} + C n^{-\beta(2-\alpha)d/(2(d+1))},$$

and

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

B.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

$$\left| u^2 n^{-\gamma} \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} b_t^n(x; r) \langle \mathbb{1}_{B(x,r)}(1 - w_t^n), \zeta_{T-t}^{n,[0,L)}(\cdot; z, \epsilon_n) \rangle^2 \mathrm{d}z \mathrm{d}r \mathrm{d}t \right|$$

$$\leq C n^{-\gamma} \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int \zeta_t^{n,[0,L)}(y; z, \epsilon_n) \int \mathbb{1}_{\{|y-x|<r\}}$$

$$\times \int \mathbb{1}_{\{|y'-x|<r\}} \zeta_t^{n,[0,L)}(y'; z, \epsilon_n) \mathrm{d}y' \mathrm{d}x \mathrm{d}y \mathrm{d}r \mathrm{d}t$$

$$\leq C n^{-\gamma} \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int \zeta_t^{n,[0,L)}(y; z, \epsilon) \int \mathbb{1}_{\{|x-y|<r\}} \mathrm{d}x \mathrm{d}y \mathrm{d}r \mathrm{d}t$$

$$\leq C n^{-\gamma} \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{1+\alpha}} \int \zeta_t^{n,[0,L)}(y; z, \epsilon) \mathrm{d}y \mathrm{d}r \mathrm{d}t$$

$$\leq C n^{-\gamma} n^{\alpha\beta} \int_0^T \mathbb{P}[K_t < L] \mathrm{d}t \leq C n^{\beta} n^{-(1-(\alpha-1)/(2\alpha))\alpha\beta} = C n^{-(\alpha-1)\beta/2} \tag{B.15}$$

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

Now we turn to

$$\left| u^2 n^{-\gamma} \int_0^{T-\tau_1} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} b_t^n(x; r) \langle \mathbb{1}_{B(x,r)}(1 - w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot; z, \epsilon_n) \right.$$

$$\left. - \zeta_{T-t}^{n,[L,\infty)}(\cdot; z, \epsilon_n') \rangle^2 \mathrm{d}x \mathrm{d}r \mathrm{d}t \right|.$$

Once again we write

$$\zeta_t^{n,[L,\infty)}(y; z, \epsilon_n) - \zeta_t^{n,[L,\infty)}(y; z, \epsilon_n')$$

$$= \int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) \left(q_t^{n,[L,\infty)}(y - x') - q_t^{n,[L,\infty)}(y - z) \right) \mathrm{d}x'$$

$$- \int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n') \left(q_t^{n,[L,\infty)}(y - x') - q_t^{n,[L,\infty)}(y - z) \right) \mathrm{d}x'.$$

This gives us

$$\left| \int_{\mathbb{R}^d} b_{T-t}^n(x;r) \langle \mathbb{1}_{B(x,r)}(1 - w_{T-t}^n), \zeta_t^{n,[L,\infty)}(\cdot; z, \epsilon_n) - \zeta_t^{n,[L,\infty)}(\cdot; z, \epsilon_n') \rangle^2 \mathrm{d}x \right|$$

$$\leq C \int_{(\mathbb{R}^d)^3} \mathbb{1}_{\{|x-y| \leq r\}} \mathbb{1}_{\{|x-y'| \leq r\}} \left[\left(\int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) \big| q_t^{n,[L,\infty)}(y - x') \right.\right.$$

$$\left. - q_t^{n,[L,\infty)}(y - z) \big| \mathrm{d}x' \right)$$

$$\times \left. \left(\int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) \big| q_t^{n,[L,\infty)}(y' - x') - q_t^{n,[L,\infty)}(y' - z) \big| \mathrm{d}x' \right) + S_t^n \right] \mathrm{d}y' \mathrm{d}y \mathrm{d}x,$$

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 B.3(iii):

$$\int_{B(z,\log n)^2} V_r(y, y') \left(\int_{\mathbb{R}^d} \zeta_0^n(x; z, \epsilon_n) \left(t^{-\frac{d+1}{\alpha}} |z - x| + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right) \mathrm{d}x \right)$$

$$\times \left(\int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) \left(t^{-\frac{d+1}{\alpha}} |z - x'| + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right) \mathrm{d}x' \right) \mathrm{d}y' \mathrm{d}y$$

$$\leq r^d \int_{B(z,\log n)^2} \mathbb{1}_{\{|y-y'| \leq 2r\}} \left(\epsilon_n t^{-\frac{d+1}{\alpha}} + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right)^2 \mathrm{d}y' \mathrm{d}y$$

$$\leq r^d (r \wedge \log n)^d (\log n)^d \left(\epsilon_n t^{-\frac{d+1}{\alpha}} + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right)^2.$$

Integrating over t and r, we obtain

$$n^{-\gamma} \int_{\tau_1}^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} r^d (r \wedge \log n)^d (\log n)^d \left(\epsilon_n t^{-\frac{d+1}{\alpha}} + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right)^2 \mathrm{d}r \mathrm{d}t$$

$$\leq C n^{-\gamma} (\log n)^d \left(\int_{n^{-\beta}}^{\log n} r^{d-1-\alpha} \mathrm{d}r + (\log n)^d \int_{\log n}^{\infty} r^{-1-\alpha} \mathrm{d}r \right)$$

$$\times \left(\epsilon_n^2 \int_{\tau_1}^T t^{-\frac{2(d+1)}{\alpha}} \mathrm{d}t + 2\epsilon_n n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \int_{\tau_1}^T t^{-\frac{d+1}{\alpha}} \mathrm{d}t + T n^{-\frac{\beta(2-\alpha)d}{(d+1)}} \right)$$

$$\leq C n^{-\gamma} (\log n)^d \left((\log n)^{d-\alpha} + n^{\beta(\alpha-d)} \right) \left[\epsilon_n^2 \tau_1^{1-\frac{2(d+1)}{\alpha}} + 2\epsilon_n n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \tau_1^{1-\frac{d+1}{\alpha}} + n^{-\frac{\beta(2-\alpha)d}{(d+1)}} \right].$$

$$\text{(B.16)}$$

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

$$\int_{B(z,\log n)} \int_{B(z,\log n)^c} V_r(y, y') \left(\int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) \left(t^{-\frac{d+1}{\alpha}} |z - x| + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right) \mathrm{d}x \right)$$

$$\times \left(\int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) e^{-|z-y'|} \mathrm{d}x' \right) \mathrm{d}y' \mathrm{d}y$$

$$\leq C \left(t^{-\frac{d+1}{\alpha}} \epsilon_n + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right) \int_{B(z,\log n)^c} \int_{B(z,\log n) \cap B(y',2r)} r^d e^{-|z-y'|} \mathrm{d}y \mathrm{d}y'$$

$$\leq C \left(t^{-\frac{d+1}{\alpha}} \epsilon_n + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \right) r^d (r \wedge \log n)^d \frac{(\log n)^{d-1}}{n}.$$

Integrating over t and r as well, we obtain

$$n^{-\gamma} \int_{\tau_1}^{T} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \left(t^{-\frac{d+1}{\alpha}} \epsilon_n + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\right) r^d (r \wedge \log n)^d \frac{(\log n)^{d-1}}{n} \, \mathrm{d}r \mathrm{d}t$$

$$\leq n^{-\gamma} \frac{(\log n)^{d-1}}{n} \left[\epsilon_n \tau_1^{1-\frac{d+1}{\alpha}} + C_{d,T} n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\right] \left[(\log n)^{d-\alpha} + n^{\beta(\alpha-d)}\right]. \qquad \text{(B.17)}$$

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 B.3(iv) gives us the bound

$$\int_{(B(z,\log n)^c)^2} V_r(y, y') \left(\int_{\mathbb{R}^d} \zeta_0^n(x; z, \epsilon_n) e^{-|z-y|} \mathrm{d}x\right) \left(\int_{\mathbb{R}^d} \zeta_0^n(x'; z, \epsilon_n) e^{-|z-y'|} \mathrm{d}x'\right) \mathrm{d}y' \mathrm{d}y$$

$$\leq C r^d \int_{B(z,\log n)^c} \int_{B(z,\log n)^c \cap B(y,2r)} e^{-|z-y|} e^{-|z-y'|} \mathrm{d}y' \mathrm{d}y$$

$$\leq C r^d (1 \wedge r^d) \frac{(\log n)^{d-1}}{n}.$$

Integrating over t and r gives the bound

$$n^{-\gamma} \int_{\tau_1}^{T} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} r^d (1 \wedge r^d) \frac{(\log n)^{d-1}}{n} \, \mathrm{d}r \mathrm{d}t \leq C T n^{-\gamma} \frac{(\log n)^{d-1}}{n} (n^{\beta(\alpha-d)} + 1). \text{ (B.18)}$$

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

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x|<r\}} \mathbb{1}_{\{|z-x|<r\}} b_t^n(x; r)(1 - w_t^n(y))(1 - w_t^n(z)) \right.$$

$$\left. (\zeta_{T-t}^{n,[L,\infty)}(y; z_1, \epsilon) - \zeta_{T-t}^{n,[L,\infty)}(y; z_2, \epsilon))(\zeta_{T-t}^{n,[L,\infty)}(z; z_1, \epsilon) - \zeta_{T-t}^{n,[L,\infty)}(z; z_2, \epsilon)) \mathrm{d}z \mathrm{d}y \mathrm{d}x \right|$$

$$\leq \int_{\mathbb{R}^d} (\zeta_{T-t}^{n,[L,\infty)}(y; z_1, \epsilon) - \zeta_{T-t}^{n,[L,\infty)}(y; z_2, \epsilon)) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x|<r\}} \mathbb{1}_{\{|z-x|<r\}}$$

$$|\zeta_{T-t}^{n,[L,\infty)}(z; z_1, \epsilon) - \zeta_{T-t}^{n,[L,\infty)}(z; z_2, \epsilon)| \mathrm{d}z \mathrm{d}x \mathrm{d}y$$

$$\leq 2 \sup_{t \in [0, T-\tau_2]} \|\zeta_{T-t}^{n,[L,\infty)}(\cdot; z_1, \epsilon) - \zeta_{T-t}^{n,[L,\infty)}(\cdot; z_2, \epsilon)\|_\lambda$$

$$\int_{\mathbb{R}^d} \zeta_{T-t}^{n,[L,\infty)}(y; z_1, \epsilon) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y-x|<r\}} \mathbb{1}_{\{|z-x|<r\}} e^{-\lambda|z|} \mathrm{d}z \mathrm{d}x \mathrm{d}y$$

$$\leq C \sup_{t \in [0, T-\tau_2]} \|\zeta_{T-t}^{n,[L,\infty)}(\cdot; z_1, \epsilon) - \zeta_{T-t}^{n,[L,\infty)}(\cdot; z_2, \epsilon)\|_\lambda \int_{\mathbb{R}^d} \zeta_{T-t}^{n,[L,\infty)}(y; z_1, \epsilon)(r^{2d} \wedge r^d) \mathrm{d}y$$

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

which yields

$$\left| u^2 n^{-\gamma} \int_0^{T-\tau_2} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} b_t^n(x; r) \langle \mathbb{1}_{B(x,r)}(1 - w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot; z_1, \epsilon) \right.$$

$$\left. - \zeta_{T-t}^{n,[L,\infty)}(\cdot; z_2, \epsilon) \rangle^2 \mathrm{d}x \mathrm{d}r \mathrm{d}t \right|$$

$$\leq C_{\lambda,d,T} n^{-\gamma} \int_0^{T-\tau_2} \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} (r^{2d} \wedge r^d)(|z_1 - z_2|^{1/2} + \tau_2) e^{\lambda(|z_1|+\epsilon)} \mathrm{d}r \mathrm{d}t$$

$$\leq C_{\lambda,d,T} n^{-\gamma} (|z_1 - z_2|^{1/2} + \tau_2) e^{\lambda(|z_1|+\epsilon)} \left(\int_{n^{-\beta}}^{1} r^{d-1-\alpha} \mathrm{d}r + \int_1^{\infty} r^{-1-\alpha} \mathrm{d}r\right)$$

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

$$\leq C_{\lambda,d,T} n^{-(d-1)\beta} (|z_1 - z_2|^{1/2} + \tau_2) e^{\lambda(|z_1|+\epsilon)} \qquad \text{(B.19)}$$

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

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

$$\left| \int_{\mathbb{R}^d} b_t^n(x; r) \langle \mathbb{1}_{B(x,r)}(1 - w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot; z, \epsilon_n) \rangle^2 \mathrm{d}x \right|$$

$$\leq C \int \zeta_{T-t}^{n,[L,\infty)}(y; z, \epsilon_n) \int \int \mathbb{1}_{\{|y-x|<r\}} \mathbb{1}_{\{|y'-x|<r\}} \zeta_{T-t}^{n,[L,\infty)}(y'; z, \epsilon_n) \mathrm{d}y' \mathrm{d}x \mathrm{d}y$$

$$\leq C_d \int \zeta_{T-t}^{n,[L,\infty)}(y; z, \epsilon_n) \int \mathbb{1}_{\{|y-x|<r\}}(1 \wedge (((T-t)^{-d/\alpha} + e^{-n^{c_5}})r^d)) \mathrm{d}x \mathrm{d}y$$

$$\leq C_d(r^d \wedge (((T-t)^{-d/\alpha} + e^{-n^{c_5}})r^{2d})),$$

which implies that

$$\left| u^2 n^{-\gamma} \int_{T-\tau_1}^T \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} b_t^n(x; r)(\langle \mathbb{1}_{B(x,r)}(1 - w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot; z, \epsilon_n) \rangle^2 \mathrm{d}x \mathrm{d}r \mathrm{d}t \right|$$

$$\leq Cn^{-\gamma} \int_0^{\tau_1} \int_{n^{-\beta}}^\infty \left[r^{-1-\alpha} \wedge (t^{-d/\alpha} + e^{-n^{c_5}})r^{d-1-\alpha} \right] \mathrm{d}r \mathrm{d}t$$

$$= Cn^{-\gamma} \int_{n^{-\beta}}^\infty \int_0^{r^\alpha \wedge \tau_1} r^{-1-\alpha} \mathrm{d}t \mathrm{d}r + Cn^{-\gamma} \int_{n^{-\beta}}^\infty \int_{r^\alpha \wedge \tau_1}^{\tau_1} (t^{-d/\alpha} + e^{-n^{c_5}})r^{d-1-\alpha} \mathrm{d}t \mathrm{d}r$$

$$= Cn^{-\gamma} \int_{n^{-\beta}}^{\tau_1^{1/\alpha}} r^{-1} \mathrm{d}r + C\tau_1 n^{-\gamma} \int_{\tau_1^{1/\alpha}}^\infty r^{-1-\alpha} \mathrm{d}r + Cn^{-\gamma} \int_{n^{-\beta}}^{\tau_1^{1/\alpha}} \int_{r^\alpha}^{\tau_1} t^{-d/\alpha} r^{d-1-\alpha} \mathrm{d}t \mathrm{d}r$$

$$\leq Cn^{-\gamma} \log n + Cn^{-\gamma+\beta(\alpha-d)} \tau_1^{1-d/\alpha}. \tag{B.20}$$

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

Combining (B.15), (B.16), (B.17), (B.18) and (B.20) yields (recall that $\epsilon_n \leq \epsilon_n'$)

$$\left[\mathcal{M}^{n, \zeta_0^n(\cdot; z, \epsilon_n) - \zeta_0^n(\cdot; z, \epsilon_n')} \right]_T \leq Cn^{-(\alpha-1)\beta/2} + n^{-\gamma+\beta(\alpha-d)} \tau_1^{1-d/\alpha}$$

$$+ n^{-\gamma}(\log n)^d \left((\log n)^{d-\alpha} + n^{\beta(\alpha-d)} \right) \left[\epsilon_n'^2 \tau_1^{1-\frac{2(d+1)}{\alpha}} \right.$$

$$+ 2\epsilon_n' n^{-\frac{\beta(2-\alpha)d}{2(d+1)}} \tau_1^{1-\frac{d+1}{\alpha}} + n^{-\frac{\beta(2-\alpha)d}{(d+1)}} \right],$$

while combining (B.15), (B.19) and (B.20) gives us

$$\left[\mathcal{M}^{n, \zeta_0^n(\cdot; z_1, \epsilon) - \zeta_0^n(\cdot; z_2, \epsilon)} \right]_T \leq Cn^{-(\alpha-1)\beta/2} + C_{\lambda,d,T}(|z_1 - z_2|^{1/2} + \tau_2)e^{\lambda(|z_1|+\epsilon)}$$

$$+ C_d n^{-\gamma+\beta(\alpha-d)} \tau_2^{(\alpha-d)/\alpha}.$$

Now, by the Burkholder-Davis-Gundy inequality ([12]),

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

Combining this and the estimate for the drift term yields the desired result. $\square$

B.3 Lemmas

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

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

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 (B.9). As X^n has independent and stationary

increments, the Lévy-Khintchine Formula (see *e.g.* Theorems 2.7.10 and 2.8.1 of [46])
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

$$\psi^n(\theta) = \int_{\mathbb{R}^d}(e^{i\theta\cdot x} - 1)\nu^n(\mathrm{d}x). \tag{B.21}$$

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,

$$\psi(\theta) = \int_{\mathbb{R}^d}(e^{i\theta\cdot x} - 1)\nu(\mathrm{d}x). \tag{B.22}$$

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

Lemma B.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, (B.21) implies

$$\psi^n(\theta) = \frac{1}{2}\int_{\mathbb{R}^d}(e^{i\theta\cdot x} - 2 + e^{-i\theta\cdot x})\nu^n(\mathrm{d}x) = \frac{1}{2}\int_{\mathbb{R}^d}(e^{i\theta\cdot x/2} - e^{-i\theta\cdot x/2})^2\nu^n(\mathrm{d}x)$$

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

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

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

$$= \int_0^{n^{-\beta}} \frac{V_r}{r^{d+1+\alpha}}\int_{\mathbb{R}^d}\frac{\sin^2(\theta\cdot x/2)}{V_r^2}\int \mathbb{1}_{\{|y|<r\}}\mathbb{1}_{\{|x-y|<r\}}\mathrm{d}y\mathrm{d}x\mathrm{d}r$$

$$\le \int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d}\sin^2(\theta\cdot x/2)\int \mathbb{1}_{\{|x|<2r\}}\mathrm{d}x\mathrm{d}r$$

$$= \int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}}\int_{|x|<2r}\sin^2(\theta\cdot x/2)\mathrm{d}x\mathrm{d}r.$$

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

$$\frac{1}{2}|\psi^n(\theta) - \psi(\theta)| \le \frac{1}{4}\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}}\int_{|x|<2r}\left(\sum_{i=1}^d \theta_i x_i\right)^2 \mathrm{d}x\mathrm{d}r$$

$$\le \frac{1}{4}\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}}\int_{-2r}^{2r}\cdots\int_{-2r}^{2r}\left(\sum_{i=1}^d \theta_i x_i\right)^2 \mathrm{d}x_1\ldots\mathrm{d}x_d\mathrm{d}r$$

$$\le \frac{1}{4}\sum_{i=1}^d \theta_i^2\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}}\int_{-2r}^{2r}\cdots\int_{-2r}^{2r}x_i^2\mathrm{d}x_1\ldots\mathrm{d}x_d\mathrm{d}r.$$

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

$$\int_0^{n^{-\beta}} \frac{(4r)^{d-1}}{r^{d+1+\alpha}} \int_{-2r}^{2r} x_1^2 dx_1 dr = \int_0^{n^{-\beta}} \frac{4^{d-1}}{r^{2+\alpha}} \frac{2}{3}(2r)^3 dr = \frac{4^{d+1}}{3}\int_0^{n^{-\beta}} r^{1-\alpha} dr = \frac{4^{d+1}}{3} n^{-\beta(2-\alpha)}.$$

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

$$-\frac{1}{2}\psi^n(\theta) = \int_{\mathbb{R}^d} \sin^2(\theta \cdot x/2) \int_{n^{-\beta}}^{\infty} \frac{V_r}{r^{d+1+\alpha}} \sqcap_r^{*2}(x) dr dx$$

$$= \int_{\mathbb{R}^d} \sin^2(\theta \cdot x/2) \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha} V_r} \int_{\mathbb{R}^d} \mathbb{1}_{\{|y|<r\}} \mathbb{1}_{\{|x-y|<r\}} dy dr dx$$

$$\geq c_0 \int_{\mathbb{R}^d} \sin^2(\theta \cdot x/2) \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \mathbb{1}_{\{|x|<r\}} dr dx,$$

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

$$-\frac{1}{2}\psi^n(\theta_1) \geq 2c_0 \int_{n^{-\beta}}^{\infty} \int_0^r \sin^2(\theta_1 x/2) \frac{1}{r^{2+\alpha}} dx dr$$

$$= 2c_0 \int_0^{\infty} \int_{n^{-\beta} \vee x}^{\infty} \sin^2(\theta_1 x/2) \frac{1}{r^{2+\alpha}} dr dx$$

$$= \frac{2c_0}{1+\alpha} \int_0^{\infty} \sin^2(\theta_1 x/2)(n^{-\beta} \vee x)^{-(1+\alpha)} dx$$

$$\geq \frac{2c_0}{1+\alpha} \int_{n^{-\beta}}^{\infty} \sin^2(\theta_1 x/2) x^{-(1+\alpha)} dx$$

$$= \frac{2c_0}{1+\alpha} \theta_1^{\alpha} \int_{\theta_1 n^{-\beta}}^{\infty} \sin^2(y/2) y^{-(1+\alpha)} dy.$$

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 $|\theta| \leq n^\beta$,

$$-\psi^n(\theta) \geq c|\theta|^\alpha \tag{B.23}$$

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 $\theta_1 > 0$:

$$-\frac{1}{2}\psi^n(\theta) \geq c_0 \int_{n^{-\beta}}^{\infty} \int_0^r \int_{|x|=\rho} \sin^2(\theta_1 x_1/2) \frac{1}{r^{d+1+\alpha}} dx d\rho dr$$

$$= c_0 \int_0^{\infty} \int_{n^{-\beta} \vee \rho}^{\infty} \int_{|x|=\rho} \sin^2(\theta_1 x_1/2) \frac{1}{r^{d+1+\alpha}} dx dr d\rho$$

$$= \frac{c_0}{d+\alpha} \int_0^{\infty} \int_{|x|=\rho} \sin^2(\theta_1 x_1/2)(n^{-\beta} \vee \rho)^{-(d+\alpha)} dx d\rho$$

$$\geq \frac{c_0}{d+\alpha} \int_{n^{-\beta}}^{\infty} \int_{|x|=\rho} \sin^2(\theta_1 x_1/2) \rho^{-(d+\alpha)} dx d\rho$$

$$= \frac{c_0}{d+\alpha} \int_{n^{-\beta}}^{\infty} \int_{|y|=1} \sin^2(\rho \theta_1 y_1/2) \rho^{-(1+\alpha)} dy d\rho$$

$$= \frac{c_0}{d+\alpha} \theta_1^{\alpha} \int_{\theta_1 n^{-\beta}}^{\infty} \int_{|y|=1} \sin^2(r y_1/2) r^{-(1+\alpha)} dy dr.$$

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 (B.23) and we have proved (ii). $\qquad\square$

Lemma B.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*

$$\mathbb{E}[e^{i\theta \cdot Z_r}] = \frac{2^{d/2}\Gamma(d/2 + 1)}{|r\theta|^{d/2}} J_{d/2}(|r\theta|),$$

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 $t > 0$,*

$$\int_{|\theta| \geq n^\beta} |\tilde{q}_t^{n,[M,\infty)}(\theta)| d\theta \leq C_d n^{\beta d} a^{M-1}.$$

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 B.4$(i)$ implies for $t \leq T$ and sufficiently large n,

$$|e^{t(\psi^n(\theta) - \psi(\theta))} - 1| \leq Ct|\psi^n(\theta) - \psi(\theta)| \leq C_{d,T}\epsilon.$$

Hence

$$\int_{|\theta| \leq n^\beta} |(e^{t(\psi^n(\theta) - \psi(\theta)} - 1)e^{t\psi(\theta)}| d\theta$$

$$\leq \int_{|\theta| \leq \sqrt{\epsilon n^{\beta(2-\alpha)}}} |(e^{t(\psi^n(\theta) - \psi(\theta)} - 1)e^{t\psi(\theta)}| d\theta + \int_{\sqrt{\epsilon n^{\beta(2-\alpha)}} < |\theta| \leq n^\beta} (e^{t\psi^n(\theta)} + e^{t\psi(\theta)}) d\theta$$

$$\leq C_{d,T}(\epsilon n^{\beta(2-\alpha)})^{d/2}\epsilon + C_d \int_{\sqrt{\epsilon n^{\beta(2-\alpha)}}}^{n^\beta} r^{d-1} e^{-ctr^\alpha} dr$$

by Lemma B.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 $C n^{\beta d} e^{-cn^b}$ (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 [48], which states that the Fourier transform of the indicator function on the unit ball in d dimensions is

$$\int_{\mathbb{R}^d} \mathbb{1}_{[0,1]}(|x|)e^{i\theta \cdot x} dx = \left|\frac{\theta}{2\pi}\right|^{-d/2} J_{d/2}(|\theta|).$$

Hence, dividing by the volume of the unit ball in d dimensions, which is $\pi^{d/2}/\Gamma(d/2 + 1)$, yields

$$\mathbb{E}[e^{i\theta \cdot Z_1}] = \frac{2^{d/2}\Gamma(d/2 + 1)}{|\theta|^{d/2}} J_{d/2}(|\theta|). \tag{B.24}$$

Scaling Z_1 by a factor of r gives us the desired result.

For (iii), we recall from (B.9) the representation of X^n using random walks with step size Y_k. Let R be an $\mathbb{R}$-valued r.v. distributed according to $\frac{V_1}{A} r^{-(1+\alpha)} \mathbb{1}_{\{r > n^{-\beta}\}} dr$, 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 (B.24) is real and

$$\tilde{q}_t^{n,[M,\infty)}(\theta) = \mathbb{E}_{K_t}\big[(\mathbb{E}_R[\tilde{\rho}(R\theta)^2])^{K_t}\mathbb{1}_{\{K_t \geq M\}}\big],$$

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 $\tilde{\rho}(v) = \mathbb{E}[\cos(v \cdot Z)] = \mathbb{E}[\cos(v_1 Z^{(1)})]$, 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 $|v_1 Z^{(1)} - n\pi| \geq \delta_1$ for all $n \in \mathbb{Z}$, then $|\cos(v_1 Z^{(1)})| \leq \cos\delta_1 < 1$. Let $I_n = ((n\pi - \delta_1)/v_1, (n\pi + \delta_1)/v_1)$, then

$$\mathbb{P}\big[|v_1 Z^{(1)} - n\pi| < \delta_1 \text{ for some } n \in \mathbb{Z}\big] = \sum_{n=-\infty}^{\infty} \mathbb{P}\big[Z^{(1)} \in I_n\big].$$

Since $-1 \leq Z^{(1)} \leq 1$, the intervals I_n for which the probabilities on 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

$$\mathbb{P}\big[|v_1 Z^{(1)} - n\pi| < \delta_1 \text{ for some } n \in \mathbb{Z}\big] \leq 2\mathbb{P}\big[Z^{(1)} \geq [1 - \delta_1, 1]\big] \leq 2\delta_2$$

for some $\delta_2 \in (0, 1/4)$ if we pick a sufficiently small δ_1. This implies

$$\mathbb{E}\big[\cos(v_1 Z^{(1)})\big] = \mathbb{E}\big[\cos(v_1 Z^{(1)})\mathbb{1}_{|v_1 Z^{(1)} - n\pi| \geq \delta_1}\big] + \mathbb{E}\big[\cos(v_1 Z^{(1)})\mathbb{1}_{|v_1 Z^{(1)} - n\pi| < \delta_1}\big]$$
$$\leq (\cos\delta_1)\mathbb{P}\big[|v_1 Z^{(1)} - n\pi| \geq \delta_1\big] + \mathbb{P}\big[|v_1 Z^{(1)} - n\pi| < \delta_1\big] \leq a$$

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

$$\mathbb{E}_R[\tilde{\rho}(Rn^{\beta}\xi)^2] = n^{-\alpha\beta}\int_{n^{-\beta}}^{\infty} r^{-(1+\alpha)}\tilde{\rho}(rn^{\beta}\xi)^2 \mathrm{d}r = \int_1^{\infty} x^{-(1+\alpha)}\tilde{\rho}(x\xi)^2 \mathrm{d}x$$
$$= 2^d\Gamma(d/2+1)^2\int_1^{\infty} x^{-(1+\alpha)}\frac{J_{d/2}(|x\xi|)^2}{|x\xi|^d}\mathrm{d}x \leq C_d\int_1^{\infty} x^{-(1+\alpha)}|x\xi|^{-(d+1)}\mathrm{d}x$$
$$\leq C_d|\xi|^{-(d+1)},$$

where we use the fact $|J_\nu(z)| < Cz^{-1/2}$ for $\nu > 0$ ([1], 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

$$\int_{|\theta|\geq n^\beta}|\tilde{q}_t^{n,[M,\infty)}(\theta)|d\theta$$

$$= n^{\beta d}\int_{|\xi|\geq 1}|\tilde{q}_t^{n,[M,\infty)}(n^\beta\xi)|d\xi$$

$$\leq n^{\beta d}\int_{|\xi|\geq 1}(a\wedge C_d|\xi|^{-(d+1)})^M d\xi$$

$$\leq n^{\beta d}\left(\int_{1\leq|\xi|\leq(C_d/a)^{1/(d+1)}}a^M d\xi+\int_{|\xi|>(C_d/a)^{1/(d+1)}}(C_d|\xi|^{-(d+1)})^M d\xi\right)$$

$$\leq n^{\beta d}\left(C_d a^M(C_d/a)^{d/(d+1)}+\int_{(C_d/a)^{1/(d+1)}}^\infty(C_d r^{-(d+1)})^M r^{d-1}\mathrm{d}r\right).$$

We take $\rho=r/C_d^{1/(d+1)}$ (hence $C_d r^{-(d+1)}=\rho^{-(d+1)}$) to obtain

$$\int_{|\theta|\geq n^\beta}|\tilde{q}_t^{n,[M,\infty)}(\theta)|d\theta\leq C_d'n^{\beta d}\left(a^{M-d/(d+1)}+\int_{1/a^{1/(d+1)}}^\infty(\rho^{-(d+1)})^M(\rho C_d^{1/(d+1)})^{d-1}d\rho\right)$$

$$\leq C_d'n^{\beta d}\left(a^{M-1}+\int_{1/a^{1/(d+1)}}^\infty\rho^{-(d+1)(M-1)-2}d\rho\right)$$

$$\leq C_d'n^{\beta d}\left(a^{M-1}+(1/a^{1/(d+1)})^{-(d+1)(M-1)-1}\right)$$

$$\leq C_d'n^{\beta d}a^{M-1}$$

if $M\geq 2$. Hence we have established (iii). $\qquad\square$

Lemma B.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

$$\mathbb{P}[K_t<M]\leq e^{-V_1 n^{\alpha\beta}t}\left(\frac{eV_1 n^{\alpha\beta}t}{M}\right)^M$$

$$=\exp(-V_1 n^{\alpha\beta}t+M(1+\log V_1+\log(n^{\alpha\beta}t)-\log M)).$$

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

$$\mathbb{P}[K_t<M]\leq Ce^{-n^{c_3\alpha\beta/2}}.$$

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 $[n^{-(1-c_3)\alpha\beta},T]$. $\qquad\square$

Finally we turn to the proof of our key lemma.

Proof of Lemma B.3. Recall from (B.9) the representation of X^n using random walks with step size Y_k: conditioned on R_k, which has density $n^{-\alpha\beta}r^{-(1+\alpha)}\mathbb{1}_{\{r>n^{-\beta}\}}\mathrm{d}r$, $Y_k|R_k=r$ has density $\sqcap_r^{*2}(x)$. Recall also the definition of q_t^n given in (B.10) and let q be the density of the limiting α-stable process with Laplace exponent ψ defined in (B.22). We

write

$$2\pi|q_t^{n,[M,\infty)}(x) - q_t(x)| = \left| \int_{\mathbb{R}^d} (e^{t\psi^n(\theta)} - \mathbb{E}[e^{i\theta \cdot (X_t^n - X_0^n)}\mathbb{1}_{\{K_t < M\}}] - e^{t\psi(\theta)})e^{-i\theta \cdot x} d\theta \right|$$

$$\leq \left| \int_{|\theta| < n^\beta} (e^{t(\psi^n(\theta) - \psi(\theta))} - 1)e^{t\psi(\theta)}e^{-i\theta \cdot x} d\theta \right| + \left| \int_{|\theta| < n^\beta} \mathbb{P}[K_t < M]d\theta \right|$$

$$+ \left| \int_{|\theta| \geq n^\beta} e^{t\psi(\theta)} d\theta \right| + \left| \int_{|\theta| \geq n^\beta} \tilde{q}_t^{n,[M,\infty)}(\theta) d\theta \right|.$$

Lemma B.5 implies that the first and fourth terms are bounded above by

$$C_{d,T} n^{-\beta(2-\alpha)d/(2(d+1))}, \quad C_d n^{\beta d} a^{M-1},$$

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 B.4(ii) implies that the third term is bounded by

$$\int_{|\theta| \geq n^\beta} e^{t\psi(\theta)} d\theta \leq \int_{|\theta| \geq n^\beta} e^{-ct|\theta|^\alpha} d\theta \leq C_d \int_{n^\beta}^\infty r^{d-1} \exp\big(- cn^{-c_2\beta(2-\alpha)/(2(d+1))} r \big) dr$$

$$\leq C_d \int_{n^\beta}^\infty \exp\big(- cn^{-c_2\beta(2-\alpha)/(2(d+1))} r + (d-1)\log r \big) dr$$

$$\leq C_d \int_{n^\beta}^\infty \exp\big(- cn^{-c_2\beta(2-\alpha)/(2(d+1))} r/2 \big) dr \leq C_d \exp\Big(- \frac{c}{2} n^{\beta - c_2\beta(2-\alpha)/(2(d+1))} \Big).$$

Combining the estimates for these four terms yields the desired result in (i) for the case $M \geq 2$. Using Lemma B.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 in Lemma 6.1 that the process η_t with generator (6.30) 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 [46], 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 derivative of f_{η_t} is uniformly bounded, therefore f_{η_1} is uniformly continuous. This means that

$$|f_{\eta_t}(x) - f_{\eta_t}(y)| = t^{-d/\alpha}|f_{\eta_1}(t^{-1/\alpha}x) - f_{\eta_1}(t^{-1/\alpha}y)| \leq Ct^{-(d+1)/\alpha}|x - y|.$$

Hence

$$|q_t(x) - q_t(y)| \leq Ct^{-(d+1)/\alpha}|x - y|,$$

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 $\lambda > 0$,

$$f_{X_k}(x_1, 0, \ldots, 0)) \leq \mathbb{P}[X_{k,1} \geq x_1 - 1]$$

$$\leq e^{-\lambda(x_1-1)}\mathbb{E}\big[e^{(\lambda,0,\ldots,0) \cdot X_k}\big] = e^{-\lambda(x_1-1)}\mathbb{E}\big[e^{(\lambda,0,\ldots,0) \cdot Y_1}\big]^k.$$

We would like to estimate $\mathbb{E}[e^{(\lambda,0,\ldots,0) \cdot Y_1}]$, for which we calculate, using Lemma B.5(ii),

$$\mathbb{E}\big[e^{(\lambda,0,\ldots,0) \cdot Y_1}\big] - 1 = n^{-\alpha\beta} \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \left(\frac{2^d \Gamma(d/2+1)^2}{(r\lambda)^d} J_{d/2}(r\lambda)^2 - 1 \right) dr$$

$$= n^{-\alpha\beta}\lambda^\alpha \int_{\lambda n^{-\beta}}^\infty \frac{1}{\rho^{1+\alpha}} \left(\frac{2^d \Gamma(d/2+1)^2}{\rho^d} J_{d/2}(\rho)^2 - 1 \right) d\rho.$$

From [1], p. 362, 9.1.69, Bessel functions are related to generalised hypergeometric functions in the following way

$$\Gamma\left(\frac{d}{2}+1\right)J_{d/2}(x)(x/2)^{-d/2} = {}_0F_1\left(\frac{d}{2}+1;-x^2/4\right) := 1 + \sum_{n=1}^{\infty}\frac{1}{(\frac{d}{2}+1)\dots(\frac{d}{2}+n)}\frac{(-x^2/4)^n}{n!}.$$

Hence

$$\left|\left(\Gamma\left(\frac{d}{2}+1\right)J_{d/2}(\rho)(\rho/2)^{-d/2}\right)^2 - 1\right| \le C_d\rho^2$$

for $\rho \in [0,1]$. This implies

$$\mathbb{E}\left[e^{(\lambda,0,\dots,0)\cdot Y_1}\right] - 1 \le n^{-\alpha\beta}\lambda^\alpha\left(\int_0^1 C_d\rho^{1-\alpha}d\rho + 2\int_1^\infty \rho^{-(1+\alpha)}d\rho\right)$$

$$\le C_\lambda n^{-\alpha\beta},$$

where we also use $\left|\frac{2^{d/2}\Gamma(d/2+1)}{\rho^{d/2}}J_{d/2}(\rho)\right| = |\mathbb{E}[e^{i\rho\cdot Z_1}]| \le 1$ in the first inequality. Hence

$$\mathbb{E}\left[e^{(\lambda,0,\dots,0)\cdot Y_1}\right] \le 1 + C_\lambda n^{-\alpha\beta} \le e^{C_\lambda n^{-\alpha\beta}},$$

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

$$q_t^{n,[1,\infty)}(x) \le \mathbb{E}_{K_t}\left[e^{-\lambda(x_1-1)}e^{C_\lambda n^{-\alpha\beta}K_t}\mathbb{1}_{\{K_t\ge 1\}}\right] \le e^{-\lambda(x_1-1)}\exp\left(V_1 n^{\alpha\beta}t(e^{C_\lambda n^{-\alpha\beta}}-1)\right)$$

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,

$$\zeta_t^{n,[L,\infty)}(x;y,\epsilon) - \zeta_t^{n,[L,\infty)}(x;z,\epsilon)$$

$$= \int(\zeta_0^n(x';y,\epsilon) - \zeta_0^n(x';z,\epsilon))q_t^{n,[L,\infty)}(x-x')\mathrm{d}x'$$

$$= \int(\zeta_0^n(x'-y;0,\epsilon) - \zeta_0^n(x'-z;0,\epsilon))q_t^{n,[L,\infty)}(x-x')\mathrm{d}x'$$

$$= \int \zeta_0^n(x';0,\epsilon)(q_t^{n,[L,\infty)}(x-y-x') - q_t^{n,[L,\infty)}(x-z-x'))\mathrm{d}x'. \qquad (B.25)$$

For $t \in [n^{-c_2\beta(2-\alpha)/(2(d+1))}, T]$ and $|y-z| \le 1$, we have

$$\sup_x |q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)|e^{\lambda|x|}$$

$$\le \sup_{x:|x-z|<2}\left|q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)\right|e^{\lambda|x|}$$

$$+ \sup_{x:|x-z|\ge 2}\left|q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)\right|e^{\lambda|x|}$$

$$\le C_{\lambda,d,T}\big[(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})e^{\lambda|z|}$$

$$+ \sup_{x:|x-z|\ge 2}\min(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))}, e^{-2\lambda|x-y|} + e^{-2\lambda|x-z|})e^{\lambda|x|}\big],$$

where we use (iii) for the first term, and (iii) and (iv) (applied with 2λ) for the second. Hence,

$$\sup_x \left| q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)\right| e^{\lambda|x|}$$

$$\leq C_{\lambda,d,T}\Big[(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})e^{\lambda|z|}$$
$$+ \sup_x(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})^{1/2}(e^{-2\lambda|x-y|} + e^{-2\lambda|x-z|})^{1/2}e^{\lambda|x|}\Big]$$

$$\leq C_{\lambda,d,T}\Big[(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))} + (t^{-(d+1)/\alpha}|y-z|$$
$$+ n^{-\beta(2-\alpha)d/(2(d+1))})^{1/2}\Big]e^{\lambda|z|}$$

$$\leq C_{\lambda,d,T}\big(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))}\big)e^{\lambda|z|}.$$

Plugging this estimate into (B.25) yields

$$\sup_x \left| \zeta_t^{n,[L,\infty)}(x;y,\epsilon) - \zeta_t^{n,[L,\infty)}(x;z,\epsilon)\right| e^{\lambda|x|}$$

$$\leq \sup_x \int \zeta_0^n(x';0,\epsilon) \left| q_t^{n,[L,\infty)}(x-x'-y) - q_t^{n,[L,\infty)}(x-x'-z)\right| e^{\lambda|x-x'|}e^{\lambda(|x|-|x-x'|)}\mathrm{d}x'$$

$$\leq C_{\lambda,d,T}\big(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))}\big)e^{\lambda|z|}\int \zeta_0^n(x';0,\epsilon)e^{\lambda|x'|}\mathrm{d}x'$$

$$\leq C_{\lambda,d,T}e^{\lambda\epsilon}\big(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))}\big)e^{\lambda|z|},$$

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. $\qquad\square$

Lemma B.7. *There exists $c_5 > 0$ such that for all $t > 0$,*

$$\sup_x \zeta_t^{n,[L,\infty)}(x;z,\epsilon) \leq C_d(t^{-d/\alpha} + e^{-n^{c_5}}),$$

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 $\tilde{q}_t^{n,[L,\infty)}(\theta) = \mathbb{E}[e^{i\theta\cdot(X_t^n-X_0^n)}\mathbb{1}_{\{K_t\geq L\}}]$. Then

$$\zeta_t^{n,[L,\infty)}(x;z,\epsilon)$$
$$= \frac{1}{2\pi}\int_{\mathbb{R}^d} \tilde{\zeta}_t^{n,[L,\infty)}(\theta)e^{-i\theta\cdot x}d\theta$$
$$\leq \left|\frac{1}{2\pi}\int_{|\theta|<n^\beta} \tilde{q}_t^{n,[L,\infty)}(\theta)\tilde{\zeta}_0^n(\theta)e^{-i\theta\cdot x}d\theta\right| + \left|\frac{1}{2\pi}\int_{|\theta|\geq n^\beta} \tilde{q}_t^{n,[L,\infty)}(\theta)\tilde{\zeta}_0^n(\theta)e^{-i\theta\cdot x}d\theta\right|$$
$$\leq \frac{1}{2\pi}\int_{|\theta|<n^\beta} |e^{t\psi^n(\theta)} - \tilde{q}_t^{n,[0,L)}(\theta)|d\theta + \frac{1}{2\pi}\int_{|\theta|\geq n^\beta} |\tilde{q}_t^{n,[L,\infty)}(\theta)|d\theta.$$

Since $|\tilde{q}_t^{n,[0,L)}(\theta)| = |\mathbb{E}[e^{i\theta\cdot(X_t^n-X_0^n)}\mathbb{1}_{\{K_t<L\}}]| \leq \mathbb{P}[K_t < L]$, we apply Lemmas B.4$(ii)$, B.6 and B.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. $\qquad\square$

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm