Transfer Thesis

Nic Freeman

November 21, 2010

Contents

Chapter 0

Preliminaries

0.1 Abstract

The Spatial Λ-Fleming-Viot process (SΛFV) is a stochastic process developed to model the frequency of genes occurring within a population inhabiting $\mathbb{R}^d$ (although with extra effort it can be defined in a general Lusin space). It is essentially a spatial version of the generalised form of the Fleming-Viot process due to Bertoin and Le Gall (2003).

In Chapter 1 we introduce the Fleming-Viot process and outline the major steps in its development. We start from the original definition in Fleming and Viot (1979) and finish with a description of the duality between the Λ-Fleming-Viot process of Bertoin and Le Gall (2003) and the Λ-coalescents of Pitman (1999), Sagitov (1999) and Donnelly and Kurtz (1999a). From there we move on to our Chapter 2 which begins with a definition of the most basic version of the SΛFV process. We use this basic version to introduce the state space, which has a non-trivial topology coming from Evans (1998), appropriate duality and an informal discussion of existence. We then review the literature to date on the SΛFV process and give our own, more general formulation of the process.

In Chapter 3 we give a proof of the existence of our formulation of the SΛFV process and characterise it as the solution to a martingale problem. We are able to prove uniqueness (via duality) in the case without selection but are unable to give a general proof of uniqueness.

In Chapter 4 we work with a family of processes which we call bursting processes. Bursting processes are a spatially discretized version of a particular type of SΛFV process; they turn out to also be a generalized version of the Voter model. In Cox et al. (2000) it was shown that in dimensions $d \geqslant 2$ the Voter model could be rescaled to super-Brownian motion and we give a proof (in $d \geqslant 3$) extending this result to bursting processes. Our final chapter discusses ideas for further work.

The first two chapters contain no new material and as such we will permit ourselves to discuss known results in an informal style. Chapters 3 and 4 contain new results for which we adopt a properly mathematical approach to our proofs.

0.2 Dependency of the chapters

Figure 1 outlines the dependency between the sections[1]. A solid arrow indicates an important dependency (e.g. carried over notation) whilst a dotted arrow indicates a dependence that is helpful but non-essential.

The shortest self contained routes to understanding the statements of results in Chapters 3 and 4 can be found via the thickened arrows on Figure 1.

The reader who is already familiar with the terminology of population genetics will have little difficulty omitting our first chapter and beginning with Chapter 2. It is preferable but non-essential to be familiar with the Λ-Fleming-Viot process discussed in Sections 1.3.2 and 1.3.3 before reading Chapter 2. A self contained description of the Λ-Fleming-Viot process can be found shortly after Definition 1.3.11.

Chapter 3 cannot be read without understanding Chapter 2, in particular the construction of the state space in Section 2.1.1 and the informal description of our version of the SΛFV process in Section

[1]Not including Chapters 0 and 5, all appendices, and introductory paragraphs at the start of some chapters.

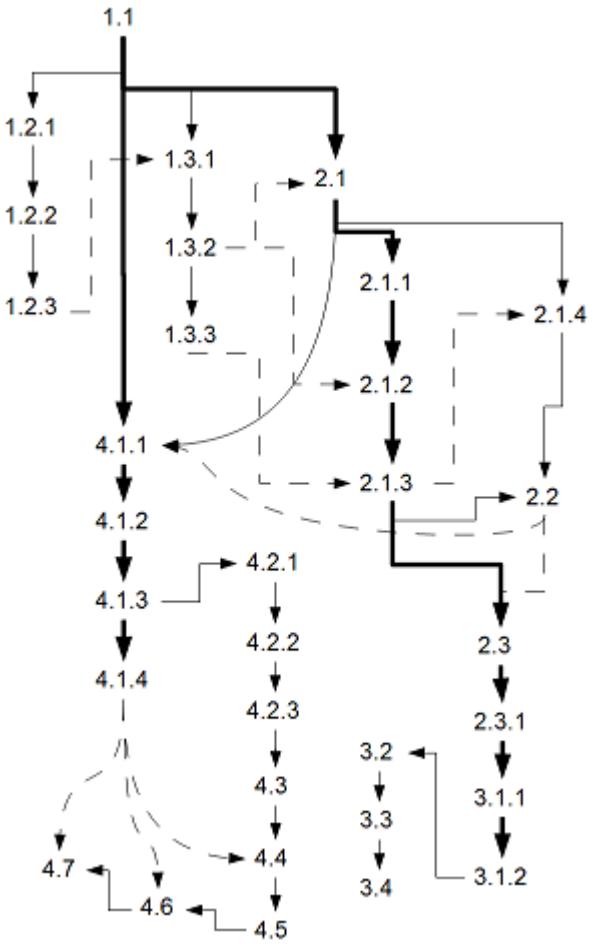

Figure 1: Diagram of dependency between sections.

2.3. In Chapter 3 we continue to use much of the notation set up in Chapter 2.

Chapter 4, on the other hand, stands on its own and (with the exception of the paragraph which portrays bursting processes as a discretization of the SΛFV process) can be understood without having read the preceeding chapters. This does not illustrate a marked change of topic in the thesis but rather the fact that in Chapter 4 we work with a simplified spatially discretized version of the SΛFV process and in order to do so we set up entirely new notation.

0.3 Notation

We will define all our notation when we come to it in the text, but we give here a list of common notation which we use throughout the thesis. In each of the following definitions we use the space A as a dummy space. In each case A is assumed to have the properties appropriate for the definition to make sense. Most of this notation is in common usage, in fact the vast majority is that of Ethier and Kurtz (1986).

Spaces of measures. Let $\mathcal{M}_F(A)$ denote the set of finite signed measures on A, equipped with the weak topology (unless it is explicitly specified otherwise) and let $\mathscr{P}(A)$ denote the closed subspace of probability measures on A. We denote the delta measure of $a \in A$ by $\delta_a(B) = \mathbb{1}\{a \in B\}$.

Spaces of functions. Let $D_A I$ denote the space of càdlàg paths $f : I \to A$, where $I \subseteq \mathbb{R}$ is an interval (usually $I = [0, \infty)$). Equip $D_A I$ with the Skorokhod topology. Whenever we use the terms 'vector space', 'Banach space', etc, we refer to spaces over $\mathbb{R}$ with pointwise operations. Let $C(A)$ denote the vector space of continuous functions on A with the topology of uniform convergence on compact sets. Let $\overline{C}(A)$ denote the continuous functions which are bounded. Then $\overline{C}(A)$ is a Banach space with the

supremum norm $||\cdot||_\infty$. Let $\hat{C}(A)$ denote the subspace of such functions which have compact support and let $C_0^\infty(A)$ denote the further subspace of continuous functions on A with compact support and derivatives of all orders.

Multiplication operators. We denote product measure with the symbol $\bigotimes$ and write $(f \times g)(x, y) = f(x)g(y)$ (but, as normal, $(fg)(x) = f(x)g(x)$).

Integrals. We will use the notations $\int f(x)dx = \int dx\, f(x)$ for integrals interchangeably; Mostly we will write $\int f(x)dx$ but when we come to write down some especially long generators of jump processes we will tend to use $\int dx\, f(x)$.

Euclidean space. We use the letter $d \in \mathbb{N}$ for dimension. We denote balls in $\mathbb{R}^d$ by $B_r(x) = \{z \in \mathbb{R}^d\,;\, |z - x| < r\}$. The Borel subsets of $\mathbb{R}^d$ will be denoted $\mathbb{B}^d$. We use $\mathcal{L}$ for Lebesgue (i.e. d-dimensional Hausdorff) measure in $\mathbb{R}^d$ with the intention that the dimension should be clear from the context.

Other. We will often use superscripts to denote dependencies as well as powers and this also should be clear from the context. Some of our notation has been chosen to agree with particular references. For the SΛFV process these are Evans (1998) and Barton et al. (2010b) whereas for bursting processes we use notation similar to Cox et al. (2000).

0.4 Acknowledgements

This thesis was written at the end of the first year of my D. Phil. at Oxford University under the supervision of Prof. Alison Etheridge and Dr. Ben Hambly. I am very grateful also to Dr. Amandine Véber and Prof. Ed Perkins for their hospitality and helpful discussion.

Chapter 1

An introduction to Fleming-Viot processes

In this Chapter we describe the development of Fleming-Viot processes, beginning with an informal construction in the style of the original paper Fleming and Viot (1979) which we relate to the more commonly cited definitions of Dawson and Hochberg (1982) and Etheridge (2000). We outline the duality between the Fleming-Viot process and Kingman's coalescent before finishing our exploration of the classical Fleming-Viot process with a basic version of the Donelly-Kurtz (1996) lookdown construction.

The SΛFV process finds its clearest connection to the literature via the generalised Fleming-Viot process of Bertoin and Le Gall (2003). We outline their framework for general exchangeable coalescents on $\mathbb{N}$ and the resulting duality between what, following Etheridge et al. (2010), we refer to as the Λ-Fleming-Viot process and the Λ-coalescents of Pitman (1999) and Sagitov (1999). This same duality was implicit in Donnelly and Kurtz (1999a).

1.1 Biological terminology

Let us begin with an informal description of some of the terminology of population genetics. Our intention is to establish common language and we do not imply biological justification for the processes we consider.

In mathematical English, an *individual* is a single organism and a *population* is a set of individuals. We remark that in contrast to the natural language meaning of 'population' we do not require all the individuals in some population to exist at the same instant in time. The *type space* is the set of genetic types an individual could assume. Usually we will choose our notation in such a way as an individual can only assume a single type at any one time.

A *coalesent* is a stochastic process whose initial state is some population X and over time groups the elements of X together. For example the initial state of the coalescent might be $\{1, 2, 3, 4, 5, 6\}$ and at some later time the state of the process might be $\{(1, 4), 2, (3, 5, 6)\}$. Usually this represents looking backwards in time and describing which individuals are descended from common ancestors.

If we view a coalescent in reversed time we see a *branching process*[1], in which we start with some individuals and over time the currently alive individuals give birth to more individuals. Usually when new individuals are born at time t one or more individuals alive at time $t-$ will pass on some of their characteristics (which could constitute spatial position and/or genetic type). These individuals will be referred to as the *parents* and the individuals who are born will be referred to as the *offspring* or *children*. We will almost exclusively look at processes where each reproduction event involves only one parent. Any occurrence within the process which potentially results in individuals being born or killed will be known as a *reproduction event*.

We explain the canonical way in which a coalescent and a branching process are ideologically the same object viewed respectively with time running in opposite directions via the following example. We will

[1]We use term informally here; we do not mean the Galton-Watson process specifically, just any process with a branching structure. Similarly, by coalescent we refer to the concept rather than precise mathematics.

usually parametrize time with the variable t and we will alternate between 'forwards in time' referring to coalescents or branching processes as is convenient. Let $B_r(x) = \{y \in \mathbb{R}\,;\, |y - x| < r\}$.

Example 1.1.1 *Let K be some set, which represents the different genetic types an individual may assume. We define a process F_t which at any time takes its values in the set $\mathscr{F}$ of functions from $\mathbb{R}$ to K. The interpretation is that at all times $t \in [0, \infty)$ we have a single individual at each $x \in \mathbb{R}$ and $F_t(x)$ represents the type of the individual at $x \in \mathbb{R}$ at time $t \in [0, \infty)$.*

The evolution of the process is specified as follows. Pick some initial state F_0. Let Λ be a Poisson point process with points $(t, x) \in [0, \infty) \times \mathbb{R}$ of intensity given by the Lebesgue measure $dt \otimes dx$.

- *If $(t, y) \in \Lambda$ then at time t the individuals at points $x \in B_1(y)$ die and are replaced with new individuals which have the type of the individual at y at time $t-$. In symbols, $F_t(x) = F_{t-}(y)$ for all $x \in B_1(y)$ and $F_t(x) = F_{t-}(x)$ for all $x \notin B_1(y)$.*
- *For each x, in between the jumps caused as above $s \to F_s(x)$ is constant.*

Thus, forwards in time we have a branching process in which at a reproduction event $(t, y) \in \Lambda$ the individual at x gives birth to individuals which instantaneously colonise the surrounding area $B_1(y)$.

Backwards in time we obtain a coalescent: Fix some time $T \in (0, \infty)$ and represent the individuals alive at time T as the set $\mathbb{R}$. To each individual $x \in \mathbb{R}$ associate a random walk $(B_s^x)_{s \in [0,T]}$ with $B_0^x = x$ and dynamics as follows:

- If $(T - s, y) \in \Lambda$ then for all $x \in B_1(y)$ set $B_s^x = y$.
- For each x, in between the jump times caused as above $s \mapsto B_s^x$ is constant.

Define the process

$$C_s = \{[x]_{\sim_s}\,;\, x \in \mathbb{R}^d\}$$

where $\sim_s$ is the equivalence relation $x \sim_s y \Leftrightarrow B_s^x = B_s^y$ and $[x]_{\sim_s}$ is the equivalence class of x. Thus C_s tells us which individuals had a common ancestor in the time interval $[T, T - s]$.

Let us use Example 1.1.1 to introduce the concept of an *ancestral lineage*. In non-spatial settings the ancestral lineage from some individual at time T is the random walk which, looking backwards in time, traces the line of ancestors back over $[T, 0]$. Since we consider processes where reproduction events involve only one parent, this can be thought of as a random walk on the set of individuals which lived in time $[0, T]$. The random walk (i.e. the lineage) moves only when, looking backwards in time, the individual which is the current value was born and at this moment the walk jumps to the parent. In the spatial setting of Example 1.1.1 since we have precisely one individual at any one point of space we carry the same information if we trace the location of the ancestors. Thus in Example 1.1.1 B^x is the ancestral lineage of x from time T.

Remark 1.1.2 *It is customary to take ancestral lineages to be right continuous, even though it would be mathematically more natural to have a left continuous process as the reversed time version of a right continuous process. We will only look at coalescents driven by Poisson point process with a Radon measure on the time component and therefore (by elementary properties of the Poisson point process) we always have ancestral lineages which are stochastically continuous. In this case the left and right versions have the same distribution so we can ignore technicalities and work with the right continuous version.*

A *multiple merger* is said to happen if some reproduction event causes strictly more than two lineages to coalesce. A *simultaneous merger* occurs if two or more reproduction events causing ancestral lineages to coalesce happen in the same instant of time. A coalescent which sees multiple mergers but not simultaneous mergers is known as a Λ-coalescent. A coalescent which sees simultaneous mergers (and usually multiple mergers as well) is known as a Ξ-coalescent.

It is natural to consider branching processes in which the reproductive success of an individual depends on its genetic type. Processes in which this occurs (both branching processes and coalescence) are said to incorporate *selection* and processes in which it does not are said to be *neutral*. We say a process exhibits *mutation* if the genetic type of a child depends on that of its parent and some additional randomness[2].

[2]In this context it could be unclear what is meant by an ancestral lineage. We will only consider lineages in processes without mutation.

Let us conclude this section with some modifications to Example 1.1.1 which would be biologically desirable. It would be natural if a reproduction event didn't completely recolonise the area in which it occurs. We should also to consider reproduction events of varying size and shape. Another noticeable problem is that in Example 1.1.1 we have no sense of how many individuals might inhabit some region, only a sense of genetic type.

It is clear in Example 1.1.1 that the individuals inhabiting any bounded region of space only change at finite rate. From a mathematical point of view it is natural to ask if there are processes similar to Example 1.1.1 which don't posses this property. Clearly some sort of control over the rate is required for the process to exist, but precisely how much? Destroying the homogeneity in time and space would be an unpleasant step towards intractability, but a neat way to aim for 'faster' processes is to allow the radius of the reproduction events to vary. In the style of Lévy processes, big events would have to happen slowly since they affect large numbers of individuals, but small events could happen very fast. We will return to this idea in Sections 1.3.3, 2.1.2 and 2.1.3.

1.2 The Fleming-Viot process.

In the genealogy of mathematics the SΛFV process is a direct descendent of the Fleming-Viot process. Fleming and Viot were interested in modelling the frequency of genetic types found within a population. They did not consider geographical effects; in fact for the duration of this chapter we will work only with non-spatial models.

1.2.1 Rescaling the Moran model

Let K be a compact metric space, which plays the role of the type space. Let $\mathcal{H}$ be the generator of some well behaved K-valued Markov process, H. We describe the Fleming-Viot process as a limit of Moran models.

Definition 1.2.1 (Moran model) *We construct a process with N particles (indexed by $\{1, 2 \ldots, N\}$) moving around in K. The evolution is specified as follows.*

- *Each pair $(i, j) \in \{1, \ldots, N\}^2$ of particles carries an exponential clock of rate $1/2$. When this clock rings the particle j instantaneously moves to the location of the particle i, and both subsequently continue independent motion according to H.*
- *In between jumps caused as above, each particle moves around K according to the process H, independently of the other particles.*

Let $Y_i^N(t)$ denote the type of the i^{th} particle at time t.

The process $Y^N = (Y_i^N)_{i=1}^N$ differs from the classical model of Moran (1958) in the respect that mutation occurs continuously rather than in jumps coinciding with the times of reproduction events.

In Definition 1.2.1 the 'position' of an individual in K corresponds to its genetic type. The operator $\mathcal{H}$ is said to be a mutation operator since it specifies the random mutation of the genetic types individuals. In later sections we will be primarily interested in the case with no mutation ($\mathcal{H} = 0$). The mechanism with the exponential clocks is often known as resampling and corresponds to reproduction with interpretation as follows. First let us note that since we are constrained to keep the population size constant if we are to think of each child as having two parents then one of the parents must die at birth. Suppose the clock corresponding to (i, k) rings. We take the individuals i and k to be parents. The parent i dies and the child is born with the type of the other parent, k.

Fleming and Viot (1979) proposed a rescaling in which one keeps the rate at which individuals mutate and resample constant but allows the number of individuals to tend to ∞. As we have already noted, they were interested in the frequency of genetic types rather than absolute numbers. In mathemetical terms this means they were interested in the empirical distribution of the limiting model.

Definition 1.2.2 *If $W = (W_i)_{i=1}^N$ is a collection of K valued random variables then define the empirical measure of W to be*

$$E^W = \frac{1}{N}\sum_{i=1}^{N} \delta_{W_i}.$$

If $(W_i(t))$ is a particle system write $E^W(t)$ for the $\mathscr{P}(K)$ valued process of the empirical measure of $W(t)$.

Definition 1.2.3 (Fleming-Viot process) *The Fleming-Viot process is the (unique) limit of the processes $E^{Y^N}(t)$ as $N \to \infty$. The limit is taken in the weak topology on the space $\mathscr{P}(D_{\mathscr{P}(K)}[0,\infty))$.*

Definition 1.2.3 does not include the precise technical requirements for the Fleming-Viot process to exist. Fleming and Viot (1979) obtain the same process from a different sequence of prelimiting models in which the mutation process was taken to be Brownian motion. Note also that Definition 1.2.3 is different to that found in Chapter 1 of Etheridge (2000) (which in turn is that of Section 5 of (Dawson and Hochberg, 1982)) since we wish to allow ourselves a general type space. In Etheridge (2000) K is taken to be $\mathbb{Z}^d$ and the mutation process ($\mathcal{L}$) is taken to be the simple random walk. This corresponds to using the stepwise mutation model of Kimura (1953) in place of Definition 1.2.1. Then in addition to the rescaling involved in Definition 1.2.3, space and time is rescaled in such a way as one recovers Brownian mutation in the limit. The limit is then a $\mathscr{P}(\mathbb{R}^d)$ valued process. To obtain the same limiting process via Definition 1.2.3 one simply takes $\mathcal{H} = \frac{1}{2}\Delta$ as the generator for Brownian motion in $K = \mathbb{R}^d$.

Remark 1.2.4 *The Fleming-Viot process has been shown to be a diffusion approximation to many of the classical population models. For example, Section 3 of Ethier and Kurtz (1993) obtains the Fleming Viot process as a limit of rescaled Wright-Fisher models. We will not discuss results of this type.*

In Fleming and Viot (1979) the Fleming-Viot process is characterized by a generator type martingale problem whereas in Etheridge (2000) it is characterised by a superprocess type martingale problem[3]. In both of these formulations existence is proved as a limit of particle systems and uniqueness of the process is proved with duality.

1.2.2 Kingman's coalescent

For the duration of this subsection let us suppose $\mathcal{H} = 0$. In other words no mutation occurs and the Moran models of Definition 1.2.1 change state only at resampling events.

Let $P_{\mathbb{N}}$ be the space of partitions of $\mathbb{N}$ and P_N be the space of partitions of $\{1,\ldots,N\}$. We will usually drop any brackets around a singleton when write down a partition as a set, for example $\{1,(2,3),4\}$ would be the same partition as $\{(1),(2,3),(4)\}$. In words we say this partition has *blocks* $(1),(2,3)$ and (4) and that $(2,3)$ was the *coagulation* of (2) and (3).

For the moment suppose we fix some (deterministic) time $T \in (0,\infty)$ and finite $N \in \mathbb{N}$. We are interested in the genealogy of an N-particle Moran model, that is we are interested in viewing the process backwards in time and recording precisely which individuals were born of which parents and when. Over time it is clear that this defines a binary tree where forwards in time we see branching and backwards in time we see coalescing.

Let us suppose we are looking at a Moran model at time t and we see n individuals (or lineages) which we label $\{1,\ldots,n\}$. Then, looking backwards in time, for each pair (i,j) at rate 1 the lineages corresponding to i and j coalesce (at rate $1/2$ i gave birth to j and at rate $1/2$ j gave birth to i). We then are down to considering only $n-1$ lineages.

Definition 1.2.5 (n-lineage Kingman's coalescent) *Define a P_N valued process κ_t as follows. Initially $\pi_0 = \{1,\ldots,N\}$. Then specify the evolution as follows.*

[3]A generator type martingale problem is the well known $f(X_t) - \int_0^t \mathcal{L}_f(X_s)ds$ type, the theory of which is developed in Ethier and Kurtz (1986). A superprocess type martingale problem is, as one might expect, the form usually used to characterise superprocesses. An example of this type (the martingale problem for super-Brownian motion) is given as Definition 4.1.5 in Chapter 4, and a discussion of how to relate it to the superprocess can be found in Section 4.1.4

- *Enumerate the blocks of κ_t currently in existence as $\{1, \dots, m\}$. To each pair of blocks (i, j) currently alive associate an exponential clock of rate 1. When the clock for the pair (i, j) rings coagulate the blocks labelled i and j, obtaining a partition with $m - 1$ blocks. Relabel the blocks $\{1, \dots, m - 1\}$ and repeat.*
- *In between jumps causing coagulation as above the process is constant (recall $\mathcal{H} = 0$).*

Note that by looking at the lineages of only a subset of m particles in the n particle Moran model we obtain a Kingman coalescent with m lineages embedded inside one for n lineages. This observation suggests that we take a projective limit of the n-lineage Kingman coalescent as $n \to \infty$. We will refer to the projective limit as Kingman's coalescent and we characterize it as follows.

Theorem 1.2.6 (Kingman (1982)) *There is a $\mathbb{P}_{\mathbb{N}}$ valued process K_t with initial value $\mathbb{N}$ (i.e. the partition of $\mathbb{N}$ into singletons) such that for any finite subset A of $\mathbb{N}$ the lineages back from A are well defined and constitute a $|A|$-lineage Kingman coalescent.*

Since the n-lineage coalescent describes the genealogy of the pre-limiting Moran models we naturally expect Kingman's coalescent to describe a particle system carrying the Fleming-Viot process as it's empirical measure. Finding a way to formally express this (without going via the pre-limiting processes) is not trivial and will be the subject of our next section. Note that, in contrast to Kingman's coalescent, we cannot readily embed an N-particle Moran model inside an $(N + 1)$-particle Moran model.

1.2.3 The lookdown process

Donnelly and Kurtz (1996) were concerned with constructing the Fleming-Viot process in such a way as keeps track of the genealogy in the measure valued limit. They did so via a particle system known as the lookdown process. In many ways the lookdown process resembles a Moran model but it contains more structure. Since the original (1996) paper the lookdown construction has been greatly extended to very general settings (see Donnelly and Kurtz (1999a), Donnelly and Kurtz (1999b) and Birkner et al. (2009) for example).

We have already described the Moran model in Definition 1.2.1. Following Donnelly and Kurtz (1996), we will define the particle system which is now known as the (first version of the) lookdown process and show that the limit of its empirical measures is the Fleming-Viot process. We conclude the chapter by using the lookdown process to describe the genealogy of the Fleming-Viot process.

Definition 1.2.7 (The (Original) Lookdown Process.) *We construct a branching system of particles moving around in S. This system will have countably many particles which we will think of as ranked into levels labelled according to $\mathbb{N}$. At all times each level $i \in \mathbb{N}$ contains exactly one particle. The evolution is as follows.*

- *In between jumps, each particle moves around K according to the process H, independently of the other particles.*
- *Each pair (i, j) of levels such that $i < j$ carries an exponential clock of rate 1. When the clock rings the particle at level j 'looks down' on level i; which means the particle in level j at time $t-$ dies and is instantaneously replaced by a new particle with the same type as its parent in i at $t-$.*

Let $X_i(t)$ denote the type of the particle at the i^{th} level at time t.

The idea of the lookdown process is as follows. If one starts both the Moran model and the lookdown process from the same initial configuration $Y(0) = X(0)$ and this configuration is exchangeable then for all time the first N particles of the lookdown process will have the same empirical measure as those of the N particle Moran model. Let us state this as a theorem.

Theorem 1.2.8 *Let $(X_i(0))_{i\in\mathbb{N}}$ be an exchangeable sequence of S valued random variables. Fix $N \in \mathbb{N}$ and let $Y = (Y_i)_{i=1}^N$ be the Moran model with initial configuration $Y_i(0) = X_i(0)$. Let $X = (X_i)_{i\in\mathbb{N}}$ be the lookdown process started from $X(0)$. Then*

$$E^X(t) \text{ is equal in disitribution to } E^Y(t)$$

for all $t \geqslant 0$.

SKETCH OF PROOF: Let us start from X and use it to drive a third particle system W. W will contain N particles and be a stochastic reordering of $(X_i)_{i=1}^N$. That is for all $t \geqslant 0$ there will exist a permutation τ of $\{1, \ldots, N\}$ such that $(X_i(t)) = (W_{\tau_t(i)}(t))$. We will then argue that W is a process with the same distribution as Y, from which the result follows.

To be precise, let $W_i(0) = X_i(0)$ and define the perm(N) valued process σ_t as follows. When (in the process X) $j \leqslant N$ looks down on i we sample an independent Bernoulli random variable with success probability $1/2$ and on a success we exchange i and j. That is, if $j \leqslant N$ looks down on i at time T, with probability $1/2$ we do nothing and with probability $1/2$ set $\sigma_T = \sigma_{T-} \circ \pi_{ij}$ where $\pi_{ij} \in \text{perm}(N)$ is the permutation exchanging i and j. Define $W_i(t) = X_{\sigma_t(i)}(t)$ for all $t > 0$.

We now argue that W is a Moran model in the sense of Definition 1.2.1. Clearly in between jump times W has the correct evolution. Let T be a jump time of X, and let us examine the ways in which pairs of particles W_i and W_j could change type. Without loss of generality consider $i < j$. Clearly W_i and W_j can only be a pair involved in the same reproduction event if it is the clock for (i, j) which rang in the process X, so suppose this happens at time T. In that case, with probability $1/2$ (a failure) we set $W_j(T) = W_i(T-)$ and leave W_i unchanged, and with probability $1/2$ (a success) we do the same thing with i and j swapped; That is we set $W_i(T) = W_j(T-)$ and leave W_t unchanged. Thus we see that at rate $1/2$ (the particle at) j takes on the type of i and with rate $1/2$ i takes on the type of j. This completes the argument. ■

With rather more formal notation, Theorem 2.1 of Donnelly and Kurtz (1996) also shows that for each t, σ_t and $Y(t)$ are independent, whence it follows that X has an exchangeable distribution for all time. We will not prove this but it is certainly something one would expect; Suppose we knew the global distribution of initial types but not their positions, and let us run the lookdown process. Then (ignoring mutation) and tracking the types of individuals, exchangeability implies that any pair of types (descended from the initial state) are equally likely to be ordered above/below one another, so providing we start in an exchangeable state it would be equally likely for any one type in the pair to be looking down on the other. The lookdown mechanism of reproduction thus generates one exchangeable state from another.

Theorem 1.2.9 *Providing the initial distribution X_0 is exchangeable, the empirical measure of X given by*

$$E^X(t) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i(t)}$$

exists and is the Fleming-Viot process.

PROOF: This follows from Theorem 1.2.8 and the fact the Fleming-Viot process is obtained as a limit of Moran models in Definition 1.2.3. ■

Let us briefly set $\mathcal{H} = 0$ again and observe that if we look at the first n particles of the lookdown process backwards in time, their genealogy is precisely that of an n-particle Kingman coalescent. Note also that the evolution backwards in time of these first n particles is not affected by the particles in levels $j > n$. Thus Theorem 1.2.9 tells us that providing X_0 is exchangeabe Kingman's coalescent is the genealogy of the Fleming-Viot process. We can also view the genealogy of the lookdown process as a spine decomposition of the Fleming-Viot process; note that the particle in level 0 does not change type.

We conclude this subsection by obtaining a representation for the generator of the Fleming-Viot process in its particle form. Fix $m \in \mathbb{N}$. For $f \in B(K^m)$ write

$$\tilde{\mathcal{H}} f = \sum_{i=1}^{m} \mathcal{H}_i f$$

where $\mathcal{H}_i f$ denotes $\mathcal{H}$ applied to f as a function only of its i^{th} coordinate. Define $F_f : \mathscr{P}(K) \to \mathbb{R}$ by

$$F_f(\mu) = \langle f, \mu^{\otimes m} \rangle$$

where $\mu^{\otimes m}$ is the m-fold product measure of μ and $\langle g, \kappa \rangle$ denotes integration of g with respect to κ. Then the generator $\mathbb{A} : C(\mathscr{P}(K)) \to C(\mathscr{P}(K))$ of the Fleming Viot process is characterized by

$$\mathbb{A} F_f(\mu) = \langle \tilde{\mathcal{H}} f, \mu^{\otimes m} \rangle + \sum_{1 \leqslant i < j \leqslant m} \left(\langle \Phi_{ij} f, \mu^{\otimes (m-1)} \rangle - \langle f, \mu^{\otimes m} \rangle \right) \tag{1.2.1}$$

where $\Phi_{ij} : B(K^m) \to B(K^{m-1})$ corresponds to setting the i and j^{th} coordinates of f equal and renumbering the coordinates (so for example for $f(x_1, x_2, x_3)$, $\Phi_{12} f(x_1, x_2) = f(x_1, x_1, x_2)$ and $\Phi_{13} f(x_1, x_2) = f(x_1, x_2, x_1)$).

The first term on the right hand side of (1.2.1) corresponds to mutation and the second term corresponds to resampling.

1.3 The generalized Fleming-Viot process

A framework for studying general exchangeable coalescents on $\mathbb{N}$ was provided in Bertoin and Le Gall (2003) where it was shown that all such coalescents can be represented with coagulation driven by a suitable Poisson point process. Building on ideas from Kingman (1982) and Kallenberg (1973) they were able to show that all exchangeable Λ-coalescents on $\mathbb{N}$ could be represented in this way. The duality between the Λ-Fleming-Viot process and Λ-coalescents was first proven in Bertoin and Le Gall (2003) (although it was implicit in the earlier modified lookdown construction from Donnelly and Kurtz (1999b)).

In this section we give an outline of some results from Bertoin and Le Gall (2003). We discuss general exchangeable coalescents before specializing our treatment to Λ-coalescents in Section 1.3.2.

1.3.1 Exchangeable coalescents on $\mathbb{N}$

Definition 1.3.1 *An exchangeable coalescent $(\pi_t)_{t \geqslant 0}$ on $\mathbb{N}$ with initial state η is a $P_{\mathbb{N}}$ valued process such that $\pi_o = \eta$ and*

- *For all $t \geqslant 0$, $b \in \eta$ implies there is some $b' \in \pi_t$ such that $b \subseteq b'$.*
- *For all $t \geqslant 0$, π_t is an exchangeable partition of $\mathbb{N}$.*

Central to the methods of Bertoin and Le Gall (2003) is the idea of an (exchangeable) bridge, which originated in Kallenberg (1973).

Definition 1.3.2 *A bridge B is a $D_{\mathbb{R}}[0,1]$ valued random variable such that*

1. *With probability one, $B(0) = 0$, $B(1) = 1$ and B is both right continuous and non-decreasing.*
2. *$r \mapsto B(r)$ has exchangeable increments.*

Results in Kallenberg (1973) show that every bridge is equal in distribution to a $C([0,1])$ valued random variable of the form

$$\hat{B}(r) = \beta_0 r + \sum_{i=1}^{\infty} \beta_i \mathbb{1}\{U_i \leqslant r\} \tag{1.3.1}$$

where β_i are $[0,1]$ valued random variables such that $\beta_i \geqslant \beta_{i+1}$ for $i \in \mathbb{N}$, $\mathbb{P}\left[\sum_{i=0}^{\infty} \beta_i = 1\right] = 1$ and $(U_i)_{i \in \mathbb{N}}$ are i.i.d. uniform random variables (on $[0,1]$).

We will shortly establish a correspondence between bridges and exchangeable partitions of $\mathbb{N}$. We will then define what it means to coagulate partitions (this corresponds to blocks combining together) and represent it in terms of bridges. We finish this section with a theorem to the effect that any exchangeable coalescent on $\mathbb{N}$ can be represented in terms of composition (as functions) of bridges.

Definition 1.3.3 *For each $\pi \in P_{\mathbb{N}}$ define an equivalence relation $i \overset{\pi}{\sim} j \Leftrightarrow \exists b \in \pi$ such that $i, j \in b$.*

Theorem 1.3.4 (The 'paintbox scheme' of Kingman (1982)) *Let B be a bridge and let B^{-1} denote the (right continuous) inverse of B. Let $(V_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. uniform random variables on $[0,1]$ and define a $P_{\mathbb{N}}$ valued random variable π by*

$$i \overset{\pi}{\sim} j \Leftrightarrow B^{-1}(V_i) = B^{-1}(V_j)$$

Then π is an exchangeable partition of $\mathbb{N}$. Further, (a random variable with the distribution of) any exchangeable partition of $\mathbb{N}$ can be constructed this way.

SKETCH OF PROOF: The first statement, the forwards direction, of Theorem 1.3.4 is easy enough to understand. The sections of B which are linear give rise to the singletons and the jumps give rise to infinite blocks. Note that we do not obtain finite blocks with more than one element. The reverse direction is slightly harder to understand. If $\pi \in P_{\mathbb{N}}$ then the asymptotic frequency of $b \in \pi$ is defined by

$$|b| = \lim_{n\to\infty} \frac{\#(b \cap \{1, 2, \ldots, n\})}{n},$$

provided the limit exists. It can be shown that every exchangeable partition of $\mathbb{N}$ is made up of blocks with asymptotic frequencies. Given some exchangeable partition π of $\mathbb{N}$ the asymptotic frequencies of $b \in \pi$, reordered in order of decreasing size, are taken as the $(\beta_i)_{i\in\mathbb{N}}$ in (1.3.1) and one sets $\beta_0 = 1 - \sum_{i\in\mathbb{N}} \beta_i$. It is a fact that the only finite blocks of an exchangeable partition are singletons. It is now apparent that the construction of in the statement of this theorem produces a partition with the law of π. ■

Thus we have a correspondence between (the laws of) bridges and (the laws of) exchangeable partitions of $\mathbb{N}$. Let us refer to the partition constructed from B by Theorem 1.3.4 as $\pi(B)$.

Definition 1.3.5 *Let $\pi, \pi' \in P_{\mathbb{N}}$. The coagulation of π by π' is $\mathrm{coag}(\pi, \pi') \in P_{\mathbb{N}}$ defined as follows. Write $\pi = \{b_i\,;\, i \in \mathbb{N}\}$ where the blocks b_i are ranked in order of their smallest element. Then set*

$$\mathrm{coag}(\pi, \pi') = \left\{ \bigcup_{i\in c} b_i\,;\, c \in \pi' \right\}.$$

Thus, π' acts as a set of instructions for which blocks of π we coagulate. An obvious question is to ask what coagulation means in terms of bridges. The answer is as follows.

Lemma 1.3.6 *Let B and B' be independent bridges. Then $\mathrm{coag}(\pi(B), \pi(B'))$ is an exchangeable partition of $\mathbb{N}$ with the same distribution as $\pi(B \circ B')$.*

The coagulation of one exchangeable partition by another is exchangeable.

One of the main results of Bertoin and Le Gall (2003) (which they state as Theorem 1) was establishing the following theorem, relating flows of bridges to exchangeable coalescents.

Definition 1.3.7 *A flow of bridges is a collection $\{B_{s,t}\,;\, -\infty < s \leqslant t < \infty\}$ of bridges such that*

1. *For every $s < t < u$, $B_{s,u} = B_{s,t} \circ B_{t,u}$.*
2. *The law of $B_{s,t}$ depends only on $t - s$.*
3. *For $s_1 < s_2 < \cdots < s_n$, the random variables $B_{s_1,s_2}, B_{s_2,s_3}, \ldots, B_{s_{n-1},s_n}$ are independent.*
4. *$B_{0,0}$ is the identity function and and $B_{0,t} \to B_{0,0}$ in probability as $t \to 0$.*

Theorem 1.3.8 (Bertoin and Le Gall (2003)) *Let $B_{s,t}$ be a flow of bridges. Then $(\pi(B_{0,t}))_{t\geqslant 0}$ is an exchangeable coalescent on $\mathbb{N}$ starting from the partition of $\mathbb{N}$ into singletons. Further $(\pi_t)_{t\geqslant 0}$ has a time homogeneous Markov semigroup and is stochastically continuous.*

Conversely, if $(\pi_t)_{t\geqslant 0}$ is a stochastically continuous exchangeable coalescent on $\mathbb{N}$ with a (time homogeneous) Markov semigroup, there is a flow of bridges $\{B_{s,t}\,;\, -\infty < s \leqslant t < \infty\}$ such that $(\pi(B_{0,t}))_{t\geqslant 0}$ and $(\pi_t)_{t\geqslant 0}$ have the same finite dimensional distributions.

1.3.2 The Λ-Fleming-Viot process and Λ-coalescents

As yet, there is nothing to say the coalescent $\pi(B_{0,t})$ constructed from Theorem 1.3.8 is a Λ-coalescent. From Lemma 1.3.6 we see that whether or not $\pi(B_{0,t})$ is a Λ-coalescent is related to whether or not the bridges feature multiple discontinuities.

For $u, x \in [0, 1]$ define $B \in C([0, 1])$ by

$$b_{u,x}(r) = (1 - x)r + x\mathbb{1}\{u \leqslant r\} \tag{1.3.2}$$

and note that if U is a uniform random variable on $[0,1]$, $B_{U,x}(r)$ is a bridge. In fact we can take this idea much further. Let M be a Poisson point process with points $(t,x,u) \in \mathbb{R} \times (0,1) \times (0,1]$ with intensity $dt \otimes dx \otimes \nu(du)$ where ν is a finite measure on $(0,1]$. Then the atoms of M are almost surely countable and without limit points so we may enumerate them ordered by their time coordinate $t_i \in \mathbb{R}$ as $(t_i, x_i, u_i)_{i \in \mathbb{Z}}$. For the remainder of this section we consider the flow

$$B^M_{s,t} = b_{x_j,u_j} \circ b_{x_{j+1},u_{j+1}} \circ \cdots \circ b_{x_k,u_k}$$

where $t_{j-1} < s \leqslant t_j \leqslant \cdots \leqslant t_k \leqslant t < t_{k+1}$. The following result is easy to prove.

Lemma 1.3.9 *$\{B^M_{s,t}\,;\, -\infty < s \leqslant t < \infty\}$ is a flow of bridges.*

Let us describe the evolution of the resulting coalescent $\Pi_t = \pi(B^M_{0,t})$. Clearly t is a jump time for Π_t only if there is a point of M in $\{t\} \times (0,1) \times (0,1]$, and Π_t is constant in between its jump times. The jump times occur at rate $\nu((0,1])$ and at each jump u is independently distribution uniformly on $[0,1]$.

Now let us say t is such a jump time corresponding to $(t,x,u) \in \Lambda$ and thus $\Pi_{t-} = \pi(B^M_{0,t-})$ denotes the state of the coalescent immediately preceding the jump. Enumerate the blocks of Π_{t-} as $(b_i)_{i \in \mathbb{N}}$, ranked in order of their smallest elements and associate an independent uniform $[0,1]$ random variable V_i to each block. Then to obtain Π_t we must coagulate together all blocks for which $b^{-1}_{x,u}(U_i) = b^{-1}_{x,u}(U_j)$. From the form of $b_{x,u}$ this means precisely those blocks for which

$$(1-x)u \leqslant U_i < (1-x)u + x$$

are coagulated into a single block, and all other blocks remain unchanged. Thus Π_t is a Λ-coalescent.

By exchangability (in particular, by Lemma 2 of Bertoin and Le Gall (2003)) we can reformulate this as follows; Label the blocks of Π_{t-} as $(b_i)_{i \in \mathbb{N}}$ (the order does not matter) and to each block associate an independent Bernoulli random variable ξ_i with success probability x. Coagulate all blocks for which $\xi_i = 1$ into a single block and leave all the blocks for which $\xi_i = 0$ unchanged.

Let us now look at time in the other direction.

Definition 1.3.10 *We define the dual flow of $B^M_{s,t}$ to be $\hat{B}^M_{s,t} = B^M_{-t,-s}$.*

Note that $\hat{B}^M$ is not quite a flow of bridges. Since $\hat{B}^M$ travels in reversed time one has $\hat{B}_{t,u} \circ \hat{B}^M_{s,t} = \hat{B}^M_{s,u}$ for $s < t < u$, but all other properties a flow of bridges are carried over from B^M.

Definition 1.3.11 *Define a $\mathscr{P}([0,1])$ process ρ_t by*

$$\rho_t([0,y]) = \hat{B}^M_{0,t}(y)$$

(and extension from the algebra generated by $\{[0,y]\,;\, y \in [0,1]\}$ to the Borel sets of $[0,1]$).

We will shortly see why it is natural to view ρ_t as a generalised version of the Fleming-Viot process and we will adopt the terminology of Etheridge et al. (2010) in calling the family of processes obtained from Definition 1.3.11 Λ-Fleming-Viot processes. Let us first make the important observation that, strictly speaking, ρ_t is a measure valued process and not a branching process. By looking at the mechanism by which the composition of bridges controls the coagulation of $[0,1]$ we see that $\rho_t(A)$ is the mass of the descendants of the set $A \subseteq [0,1]$ at time t. Having realised this is it natural to think of ρ_t as corresponding to a branching process and hope for a duality relationship between π_t and ρ_t. As we have commented, flows of bridges are jump processes. It is straightforward to deduce that ρ_t is also a jump process and hence has càdlàg version, which we will work with from now on.

From the definition of the flow $B^M_{s,t}$ we get the following evolution for ρ_t.

- Reproduction events occur at rate $\nu((0,1])$ and in between reproduction events the process ρ_t is constant.
- At a reproduction event we sample an independent random variable U according to $\nu(\cdot)/\nu((0,1])$ and an independent random variable X with law ρ_t. The change is given by

$$\rho_t = (1-U)\rho_{t-} + U\delta_X. \tag{1.3.3}$$

That is, a parent type X is sampled according of the distribution of the current population and a random proportion U, with law ν, of the population are replaced with that type. The individuals alive at time t are reduced in number through a factor $(1-U)$.

In this formulation ρ_t looks similar to the Fleming Viot process of Section 1.2, but with multiple branching instead of reproduction events producing a single individual. It is also in this form that we will eventually see the clearest connection to the SΛFV process. Note the strong similarity between ρ_t and the process tracking the total mass of a single type in Example 1.1.1.

From our description of the evolution of ρ_t and the standard theory of Markov process it follows that the generator of ρ_t is given by the bounded linear operator $G : C(\mathscr{P}([0,1])) \to C(\mathscr{P}([0,1]))$ defined by

$$G\phi(\rho) = \int_0^1 \rho(dk) \int_0^1 \nu(du) \left[\phi\big((1-u)\rho + u\delta_k\big) - \phi(\rho)\right] \tag{1.3.4}$$

and that the process is characterised as the solution to the following martingale problem.

Definition 1.3.12 *The generalised Fleming Viot process is the unique càdlàg $\mathscr{P}([0,1])$ valued process such that for all $\phi \in C(\mathscr{P}([0,1]))$,*

$$\phi(\rho) - \int_0^t ds\, G\phi(\rho_s)$$

is a martingale.

Remark 1.3.13 *Expressions corresponding to the generators of jump processes should be interpreted as follows. The integrals at the start of the expression are selecting a reproduction event and the integrand is the change exacted by that reproduction event seen through the eyes of the test function. The mass involved in the measures for the integrals corresponds to the rate of the different possible reproduction events.*

The first integral in (1.3.4) *is selecting a parent type and the second integral is selecting the proportion of children to be born of that type. The total rate of events is $\int_0^1 \nu(du) \int_0^1 \rho(dk) = \nu(0,1)$.*

We now establish the duality between Π_t and ρ_t. The generator for Π_t can be characterised as follows. Let P_p denote the space of partitions of $\{1, 2, \cdots, p\}$ and write Π_t^p for the restriction of Π_t to $\{1, \ldots, p\}$ (obtained by simply deleting the natural numbers $> p$). Then Π_t^p is a Markov process. For $\psi : P_p \to \mathbb{R}$ and $\pi \in P_n$ define

$$G^*\psi(\pi) = \sum_{J \subseteq \{1,\ldots,n\}, |J| \geqslant 2} \beta_{n,|J|} \left[\psi(m_J \pi) - \psi(\pi)\right] \tag{1.3.5}$$

where $\pi = \{b_i\,;\, i = 1, \ldots, n\}$, $m_J\pi$ denotes π with the blocks corresponding to $i \in J$ coagulated and

$$\beta_{n,|J|} = \int_0^1 \nu(ds)\, s^{|J|}(1-s)^{n-|J|}.$$

Then G^* is the generator of Π_t^p. Note that, in keeping with our discussion of how to interpret generators of jump processes, the summation in (1.3.5) chooses which blocks coagulate, $\beta_{n,|J|}$ specifies the rates at which the reproduction events corresponding to coagulating the blocks with indices in J occur, and $\psi(m_J\pi) - \psi(\pi)$ is the resulting change to Π_t viewed through the test function ψ.

For each $p \in \mathbb{N}$ and $f \in C([0,1]^p)$ define $\Phi_f : \mathscr{P}([0,1]) \times P_p \to \mathbb{R}$ by

$$\Phi_f(\rho, \pi) = \int_{[0,1]^n} \rho(dx_1) \ldots \rho(dx_n) f\left(Y_\pi(x_1, \ldots, x_n)\right)$$

where $n = |\pi|$ is the number of blocks of π and Y_π is defined by $Y_\pi(x_1, \ldots, x_n) = (y_1, \ldots, y_p)$ where $y_j = x_i$ if (and only if) $i \in b_j$ where $\pi = \{b_j\,;\, j = 1, \ldots, p\}$ are the blocks of π ordered by least element.

The duality relationship between Π_t and ρ_t is formally stated as follows.

Lemma 1.3.14 *It holds that*

$$G\Phi_f(\rho, \pi) = G^*\Phi_f(\rho, \pi) \tag{1.3.6}$$

for all $p \in \mathbb{N}$, $f \in C([0,1]^p)$, $\rho \in \mathscr{P}([0,1])$ and $\pi \in P_p$. On the left hand side of (1.3.6) *G acts on Φ_f as a function of its first coordinate whereas on the right hand side G^* acts on Φ_f as a function of its second coordinate.*

Hence,

$$\mathbb{E}\left[G\Phi_f(\rho_t, \Pi_0)\right] = \mathbb{E}\left[G^*\Phi_f(\rho_0, \Pi_t)\right] \tag{1.3.7}$$

for all $t \geqslant 0$.

Sketch of Proof: The reader can verify (1.3.6) with straightforward (but messy) algebra and the second statement follows from (1.3.6) and the theory of martingale problems (e.g. Section 4.4 of Ethier and Kurtz (1986)). See Lemma 4, Bertoin and Le Gall (2003). ■

Equation (1.3.7) expresses the duality between Π_t and ρ_t. Essentially, the mechanism by which Π_t evolves forwards in time can be recovered from looking at how ρ_t evolves backwards in time (and vice versa). This is sufficient to obtain uniqueness to the relevant martingale problems.

1.3.3 Infinite rate coalescents

By allowing a more general form for (1.3.2) we can adapt the method of Section 1.3.2 to generate a wider class of coalescents. See Section 4 of Bertoin and Le Gall (2003) for details. A natural question to ask is whether or not one can relax the requirement that ν be a finite measure. If ν is infinite then we are unable to label the atoms of ν in order of time coordinate with the integers, and cannot carry through the representation of Section 1.3.2. However, if ν is an infinite measure we can approximate it with a sequence of finite measures and take a limit. We lose the neatness of the representation in the limit but we are able to construct a large family of infinite rate coalescents.

Remark 1.3.15 *Note that the measure ν was constrained to be a measure on $(0,1]$ instead of the more natural $[0,1]$. In Bertoin and Le Gall (2003) the framework is developed in such a way as when ν is point mass at $\{0\}$ one obtains precisely Kingman's coalescent (of Section 1.2.2). A general exchangeable coalescent on $\mathbb{N}$ can then be viewed as a having an evolution which is a mixture of Kingmans coalescent and a Ξ-coalescent which corresponds to the dynamics induced by a flow of bridges and a suitable Poisson point process.*

Clearly we cannot allow ν to be any infinite measure and we should ask precisely what the right condition on ν is for the approximations with a sequence $(\nu_n)_{n\in\mathbb{N}}$ of finite measures to generate a sequence of processes with a unique limit point. If existence of the limit process is given, uniqueness can be covered using the same sort of duality as is exhibited above for the finite rate processes. Thus we are really only concerned with existence of the limit.

The correct condition (with the flows of bridges as in Section 1.3.2) is that

$$\int_0^1 u^2 \nu(du) < \infty. \tag{1.3.8}$$

Note that $\nu((u_0, 1]) < \infty$ for all $u_0 > 0$ but the mass apportioned by ν might blow up around 0. This reflects a principle which is best known in the context of existence of Lèvy processes; if we fix a threshold $u_0 > 0$ and look at the rate of events which have a total effect of magnitude above u_0 we see it must be finite. However, the overall rate of events may still be infinite. We will not discuss why the condition (1.3.8) is precisely the right one but it is essentially a compensation very much in the style of Lèvy processes. For full details see Pitman (1999).

Chapter 2

The Spatial Λ-Fleming-Viot Process

In this chapter we give two informal definitions of the SΛFV process. First, in Definition 2.1.1 we give the most basic formulation and in this context we discuss some important aspects of the process. In particular, we construct the state space Ξ (which comes from Evans (1998)), give a representation of the duality in a similar style to Barton et al. (2010b) and discuss the mechanism by which reproduction events occur at infinite rate. We will then survey the current literature featuring the SΛFV process before giving own own formulation in Definition 2.3.1. Our formulation will feature selection and we give a brief description of why duality then fails in Section 2.3.1.

2.1 The SΛFV process (basic form)

Recall K is a compact metric space which for us plays the role of the type space. Let us associate a $\mathscr{P}(K)$ valued process $\rho_t(x)$ to each $x \in \mathbb{R}^d$.

Definition 2.1.1 (Spatial Λ-Fleming-Viot Process, Basic Form) *Let Λ be a Poisson point process with points $(t, y, r, u) \in [0, \infty) \times \mathbb{R}^d \times (0, \infty) \times [0, 1]$ of rate*

$$dt \otimes dy \otimes \mu(dr)\nu_r(du)$$

where dt and dy correspond to Lebesgue measure and $\mu(dr)\nu_r(du)$ corresponds to a measure on $(0, \infty) \times [0, 1]$ such that such that for each r, $\nu_r \in \mathscr{P}([0, 1])$ and

$$\int_{(0,\infty)\times[0,1]} ur^d \mu(dr)\nu_r(du) < \infty. \tag{2.1.1}$$

From some initial state ρ_0 the evolution is specified by

- *Whenever $(t, y, r, u) \in \Lambda$, sample k according to $\rho_{t-}(y)$. Set*

$$\rho_t(x) = (1 - u)\rho_{t-}(x) + u\delta_k \tag{2.1.2}$$

  for all $x \in B_r(y)$.

- *For each $x \in \mathbb{R}^d$, in between the jump times caused by reproduction events (as above) the process $\rho_t(x)$ is constant.*

We refer to (t, x, r, u) as a reproduction event occurring at time t about y with radius r and killing proportion u. We refer to y as the parent site and k or δ_k as the parent type.

It can be seen from (2.1.1) that we are allowing choices for μ and ν_r such that any open set $A \subseteq \mathbb{R}^d$ is hit (by which we mean intersected) by reproduction events at infinite rate. In fact, we even allow some μ, ν_r such that points are hit at infinite rate. We will discuss this in Sections 2.1.2 and 2.1.3.

Infinite rate occurrences cause a potential problem in defining the process. The result is that for any time $t \in [0, \infty)$ we can only define $\rho_t(x)$ for almost all $x \in \mathbb{R}^d$. It is not known if the exceptional set

of points can be taken independent of time, or even if it needs to be non-empty. We will not formulate proper existence/uniqueness theorems for the SΛFV process until Chapter 3, but we will informally discuss existence and (2.1.1) in Section 2.1.3.

Remark 2.1.2 *We stress that the measure Λ is not intended to correspond to the measure Λ involved in specifying a Λ-coalescent in the formulation of Pitman (1999). However they do share the characteristic that they control the mechanism of reproduction events, which explains our choice of notation.*

Remark 2.1.3 *Most authors do not take the parent type to sampled from the measure at the center of the reproduction event. Our concern for the moment is to discuss the basic features of the process, and to do so Definition 2.1.1 will be quite adequate. See 2.1.4 for possible generalizations of Definition 2.1.1. In Section 2.2 we review the literature to date on the SΛFV process.*

The measure $\rho_t(x)$ should be thought of as representing the local distribution of genetic types. The process is not concerned with the quantity of individuals present in a spatial location, only the relative frequencies of their genetic types. Until further notice (which means, up until Definition 2.3.1) we refer to the process characterised by Definition 2.1.1 as the SΛFV process. We will eventually give a more general definition; this chapter is concerned with understanding the process and for that Definition 2.1.1 will be quite sufficient.

The SΛFV process is a much generalized version of Example 1.1.1. To see this, take $\nu_i = \delta_{\{1\}}$ and start from a state where for each x there is some k_x such that $\rho_0(x) = \delta_{k_x}$. Then it is natural to think of the process as associating a single type to each point and represent the process as a random function mapping each point to a type rather than a probability measure on the type space. We refer to the case $\nu_i = \delta_{\{1\}}$ as *total killing*. We will study this special case in Chapter 4.

Note also the similarity between (2.1.2) and (1.3.3). From this relationship we see that the Definition 2.1.1 is a natural spatial version of the Λ-Fleming-Viot process. In Subsection 2.1.2 we will describe the dual process associated to the SΛFV process and this will be reminiscent of the duality seen in Theorem 1.3.7.

2.1.1 The state space Ξ

In this section we set up the state space for the SΛFV process and characterise its topology. The results of this section are essentially due to Evans (1998). We will relegate all the proofs of results in this Section to Appendix A.

We do not yet have a clear idea of what it means to be an individual in the SΛFV process, but for now let us tacitly assume these individuals will live in $\mathbb{R}^d$, each individual will have a single genetic type and the types of individuals will be drawn from the compact metric space K. Equip $\mathcal{M}_F(K)$, the space of finite signed measures on K, with the total variation norm $||\cdot||_{TV}$. Since K is compact this induces the weak topology on $\mathcal{M}_F(K)$.

Let

$$\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(K)) = \left\{\rho : \mathbb{R}^d \to \mathcal{M}_F(K)\,;\, \rho \text{ is measurable and } \operatorname{esssup}\left\{||\rho(x)||_{TV}\,;\, x \in \mathbb{R}^d\right\} < \infty\right\}.$$

Then $\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(K))$ is a vector space with seminorm

$$||\rho|| = \operatorname{esssup}\left\{||\rho(x)||_{TV}\,;\, x \in \mathbb{R}^d\right\}. \tag{2.1.3}$$

We set

$$\rho_1 \sim_1 \rho_2 \text{ iff } \{x \in \mathbb{R}^d\,;\, \rho_1(x) \neq \rho_2(x)\} \text{ is Lebesgue null,}$$

and define $L^\infty(\mathbb{R}^d, \mathcal{M}_F(K))$ to be the quotient of $\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(K))$ under the equivalence relation $\sim_1$. Then $L^\infty(\mathbb{R}^d, \mathscr{P}(K))$ is a Banach space. Let $[\rho]_{\sim_1}$ denote the equivalence class of ρ under $\sim_1$.

Definition 2.1.4 *Define*

$$\Xi = \left\{[\rho]_{\sim_1} \in L^\infty(\mathbb{R}^d, \mathscr{P}(K))\,;\, \textit{for almost all } x \in \mathbb{R}^d, \rho(x) \in \mathscr{P}(K)\right\}$$

and note this is both well defined and a norm-closed subspace of $L^\infty(\mathbb{R}^d, \mathscr{P}(K))$.

The space Ξ is the state space of the SΛFV process, but with a different topology to the subspace topology inherited from $\mathcal{L}^\infty(\mathbb{R}^d, \mathcal{M}_F(K))$. We characterise the topology on Ξ in Proposition 2.1.7. We will abuse notation in the standard fashion by writing $\rho \in \Xi$ instead of $[\rho]_{\sim_1} \in \Xi$.

Our test functions will be as follows.

Definition 2.1.5 *We say $\Phi : (\mathbb{R}^d)^n \to C(K^n)$ is of the form $(\star, n)$ if it may be written as*

$$(\Phi(z))(k_1, \ldots, k_n) = \psi(z) \prod_{i=1}^{n} \chi_i(k_i)$$

where $\psi : \mathbb{R}^d \to \mathbb{R}$ is continuous with compact support, and $\chi_i \in C(K)$. We say Φ is of the form $(\star)$ if Φ is constant or Φ is of the form $(\star, n)$ for some $n \in \mathbb{N}$.

Remark 2.1.6 *Whenever we say 'let Φ be of the form $(\star, n)$' we implicitly associate the functions ψ and χ to such Φ.*

Recall our convention for multiplication operators, in particular that $(f \times g)(x, y) = f(x)g(y)$ and $\otimes$ denotes product measure. For any $n \in \mathbb{N}$ and $\Phi : (\mathbb{R}^d)^n \to C(K^n)$ such that $\int_{(\mathbb{R}^d)^n} ||\Phi(x)||_\infty dx < \infty$ we define $I_n(\cdot\,; \Phi) : \Xi \to \mathbb{R}$ by

$$I_n(\rho; \Phi) = \int_{(\mathbb{R}^d)^n} \left\langle \Phi(x_1, \ldots, x_n), \bigotimes_{j=1}^{n} \rho(x_j) \right\rangle dx_1 \ldots dx_n.$$

where $\langle \cdot, \cdot \rangle$ denotes integration over K^n. If Φ has the form $(\star, n)$ then this becomes

$$I_n(\rho; \Phi) = \int_{(\mathbb{R}^d)^n} \psi(x_1, \ldots, x_n) \left\langle \bigtimes_{j=1}^{n} \chi_j \,,\, \bigotimes_{j=1}^{n} \rho(x_j) \right\rangle dx_1, \ldots, x_n.$$

Let

$$\mathscr{I} = \operatorname{lin} \{ I_n(\cdot\,; \Phi)\,;\ \Phi \text{ is of the form } (\star) \}\,.$$

Proposition 2.1.7 (mostly Evans (1998)) *There exists a metric r for Ξ such that*

1. *(Ξ, r) is a compact, complete, separable metric space.*
2. *$\mathscr{I}$ is a dense subset of $C(\Xi)$.*
3. *$\mathscr{I}$ is both a separating and convergence determining class of $C(\Xi)$.*
4. *If $\rho^m, \rho \in \Xi$, then $r(\rho^m, \rho) \to 0$ if and only if for all $\psi : \mathbb{R}^d \to \mathbb{R}$ continuous with compact support and $\chi \in C(K)$,*

$$\int_{\mathbb{R}^d} \psi(x) \langle \chi, \rho^m(x) \rangle \, dx \to \int_{\mathbb{R}^d} \psi(x) \langle \chi, \rho(x) \rangle \, dx.$$

PROOF: See Appendix A. ■

From this point on, we use the topology on Ξ induced by r. In general this topology is *not* the topology induced by (2.1.3). The topological properties of Ξ given by Proposition 2.1.7 make it a convenient space to work with.

The topology on Ξ is also natural from the point of view of biological sampling. Consider the test function; first one chooses some $n \in \mathbb{N}$ corresponding to how many sites one wants to consider sampling at once. Then choose $\psi : \mathbb{R}^d \to \mathbb{R}$ integrable and $\chi_j : K \to \mathbb{R}$ measurable and bounded. The function

$$\rho \mapsto \int_{(\mathbb{R}^d)^n} \psi(x_1, \ldots, x_n) \left\langle \bigtimes_{j=1}^{n} \chi_j, \bigotimes_{j=1}^{n} \rho(x_j) \right\rangle dx_1 \ldots dx_n$$

tells us the amount of n-tuplets of $\mathbb{R}^d$ (i.e. $(x_j)_{j=1}^n \in (\mathbb{R}^d)^n$) which in state ρ have the genetic types weighted according to χ_j (for the type of the j^{th} individual) and ψ (for locations). With these test

functions one can extract complete information about the distribution of genetic types in any non-null[1] spatial region.

For biological reasons one wants such functions to be continuous on the state space, whereas from an analytic point of view one wants fewer open sets (since this is what leads to simpler topological structure and easier definition of probability measures etc.). Proposition 2.1.7 says that with Ξ and $\mathscr{I}$ we have achieved the best of both worlds.

Remark 2.1.8 *Using the space Ξ gives us a weak formulation of the process with regards to genetic type but a strong formulation with regards to location. By a 'weak formulation' we mean that we lose track of individuals and can only obtain information about them through looking at large subsets of them via the measures $\rho_t(x)$. However, we maintain precise knowledge of which types exist where in $\mathbb{R}^d$.*

2.1.2 Duality

In Example 1.1.1 it was easy to identify what it meant to be an individual and correspondingly we could define an ancestral lineage. In Definition 2.1.1 it is not immediately clear what one should call an individual. The rate at which a single point $x \in \mathbb{R}^r$ is hit by reproduction events in the process of Definition 2.1.1 is (up to a constant corresponding to the volume of a d dimensional unit ball)

$$\int_0^\infty r^d \mu(dr) \tag{2.1.4}$$

which is potentially infinite, even if (2.1.1) holds. However, a point x is associated to a probability measure $\rho_t(x)$ on K which corresponds to the local distribution of genetic types. The interpretation is as follows; At each site x we have a local neighbourhood containing uncountably many individuals indexed by $[0,1]$. The measure $\rho_t(x)$ specifies the distribution of types of these individuals. Each individual has a single genetic type. At each reproduction event (t, y, r, u) a parent type individual whose type is the variable k is sampled from $\rho_{t-}(y)$. The variable u specifies what proportion of the individuals at x (for $x \in B_r(x)$) are replaced by the parent type k. Note that each genetic type present in the local neighbourhood represented by ρ_t is replaced to an equal extent.

In this *notional formulation* (terminology which we use for the remainder of this section), individuals have a single genetic type and are born/killed at finite rate. In this notional world there are uncountably many individuals associated to each site (and there are uncountable many sites). It is sensible to expect that existence of the SΛFV process is a delicate question.

The notional formulation suggests a way to just define what it means to be an ancestral lineage. The relationship between the dual process and the SΛFV process will then be expressed as a functional identity as in (1.3.7), instead of embedding the lineages directly into the forwards in time model.

Definition 2.1.9 (A dual to the basic SΛFV process) *Let Π be a Poisson point process with points $(t, x, r) \in [0,\infty) \times \mathbb{R}^d \times (0,\infty) \times [0,1]$ of rate*

$$dt \otimes dx \otimes \mu(dr)\nu_r(du).$$

Fix $N \in \mathbb{N}$. We will define a system of N particles (which we refer to as lineages) moving around $\mathbb{R}^d$. Let B_t^n denote the position of the n^{th} lineage at time $t \geqslant$. We will also need a process $\sim_t$, taking values in the equivalence relations on $\{1, \ldots, N\}$. Fix some initial state $(B_0^n)_{n=1}^N \in \mathbb{R}^d$ and define the evolution as follows.

- *If $(t, x, r, u) \in \Pi$, let A denote the set of equivalence classes of $\sim_{t-}$. For each $a \in A$, write $B_t^a = B_t^{a_0}$ where $a_0 \in a$, and note this does not depend on the choice of $a_0 \in a$.*

  For each $a \in A$, if $B_{t-}^a \in B_r(y)$ then, independently of all else, sample a Bernoulli random variable with success probability u and on a success set $B_t^n = y$ for all $n \in a$. Coalesce the blocks $a \in A$ which saw a success to obtain $\sim_t$ from $\sim_{t-}$. On a failure do nothing.

- *In between reproduction events causing jumps as above, the lineages do not move and $t \mapsto \sim_t$ is constant.*

[1] By this we mean, of positive Lebesgue measure.

The first thing to observe is the effect of the equivalence relations, which keep track of blocks of coalesced lineages. If at any point in time two lineages B_t^m and B_t^n are involved in the same reproduction event (i.e. the Bernoulli random variable described above is sampled for the relevant blocks and is a success in each case), they coalesce and remain together for all further time. Having realized that, in order to understand Definition 2.1.9 one needs only to understand the motion of a single lineage. If the lineage B is at the point $B_{t-} = x$ and is hit by a reproduction event (t, y, r, u), which means $x \in B_r(y)$, then with probability $1-u$ the lineage stays put and with probability u it moves to z sampled uniformly from within $B_r(y)$.

Since $\mu(dr)\nu_r(du)$ is potentially an infinite measure we must check that (2.1.1) implies the system in Definition 2.1.9 is well defined. First note that if two lineages are at separation $z \in \mathbb{R}^d \backslash \{0\}$ then they coalesce at rate

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}\{0, z \in B_r(y)\} y^2 \, \nu_r(du)\mu(dr)dy. \tag{2.1.5}$$

Performing the integration with respect to y and using the bound $L_r(0,z) \leqslant C_d r^d$ where $L_r(x,y)$ denotes the volume of $B_r(x) \cap B_r(y)$ and C_d is the volume of a d dimensional unit ball we get

$$(2.1.5) \leqslant C_d \int_0^\infty \int_0^1 u^2 r^d \, \nu_r(du)\mu(dr).$$

Since $u^2 \leqslant u$ the condition (2.1.1) is more than sufficient to guarantee (2.1.5) is finite. A similar calculation can be done for multiple coalescence events.

Secondly, a lineage jumps with increments given by a Poisson point process $\widetilde{\Pi}$ with points $(t, \Delta x) \in [0,\infty) \times \mathbb{R}^d$ and intensity

$$dt \otimes \int_0^\infty \int_0^1 \mathbb{1}\{0 \in B_r(x)\} u \nu_r(du)\mu(dr) \, dx.$$

Thus a lineage would correspond to a well defined Lèvy process $t \mapsto \sum_{(s,\Delta x)\in\widetilde{\Pi}} \mathbb{1}\{s < t\}\Delta x$ if

$$\int_{\mathbb{R}^d} (1 \wedge |x|^2) \left(\int_0^\infty \int_0^1 \mathbb{1}\{0 \in B_r(x)\} u \nu_r(du)\mu(dr) \right) dx$$

was finite. In fact, equation (2.1.1) gives us something much stronger; it says the lineages jump at a finite rate, in particular

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbb{1}\{0 \in B_r(x)\} u \nu_r(du)\mu(dr)dx = C_d \int_0^\infty \int_0^1 r^d u \, \nu_r(du)\mu(dr).$$

These two checks, whilst not a formal proof, establish beyond reasonable doubt that the dual system is well defined. We might also suspect that (2.1.1) is in fact too strong. This leads us to our next section, but first let us record a proper statement of the duality.

Let $B_t = (B_t(m))_{m=1}^n$ be the dual system of Definition 2.1.9 ran for time t from initial state B_0. Define an equivalence relation $\sim_t$ on $\{1,\dots,n\}$ by $n \sim_t m \Leftrightarrow B_{t-}(n) = B_{t-}(m)$. Suppose $\sim_t$ has $l(t)$ equivalence classes and let them be enumerated as $A_t = \{a_1^t, \dots, a_{l(t)}^t\}$. Note $B_{t-}(a_k^t)$ is well defined for $k \leqslant l(t)$. For any bounded measurable $F : [0,1]^n \to \mathbb{R}$ and $\rho \in \Xi$ define Υ_n by

$$\Upsilon_n(\rho, B_t; F) = \int_{K^{l(t)}} F(k_1, \dots, k_{l(t)}) \bigotimes_{i=1}^{l(t)} \rho(B_t(a_i^t))(dk_i).$$

Let $\mathbb{E}_{\{B_0(m)=x_m\}}$ denote expectation on the probability space of the n-particle dual system of Definition 2.1.9 with initial state $B_0(m) = x_m$ for $m = 1,\dots,n$. Let $\mathbb{E}_{\rho_0}$ denote expectation on the probability space of the basic SΛFV process ρ_t with initial condition $\rho_0 \in \Xi$.

Theorem 2.1.10 *The semigroup of the basic SΛFV process is characterized by the relation*

$$\mathbb{E}_{\rho_0}\left[I_n(\rho_t, \Phi)\right] = \int_{(\mathbb{R}^d)^n} \mathbb{E}_{\{B_0(m)=x_m\}}\left[\Upsilon_n\big(\rho_0, B_t; \Phi(x_1,\dots,x_n)\big)\right] dx_1 \dots dx_n$$

for all Φ of the form $(\star, n)$.

The above theorem essentially comes out of the construction in Section 4 of Barton et al. (2010b) (which is really about a very slightly different version of the SΛFV process but the modifications are trivial). We will discuss the construction of Barton et al. (2010b) in more detail in Section 2.2.

2.1.3 The rate of reproduction

Let us compare (2.1.1) to the corresponding equation (1.3.8) for the Λ-Fleming-Viot process. As we have already noted a point $x \in \mathbb{R}^d$ is hit by reproduction events at rate

$$\int_0^\infty r^d \mu(dr)$$

which is potentially infinite. One might examine (1.3.8) and (2.1.4) and guess that condition for existence of the basic SΛFV was

$$\int_{(0,\infty)\times[0,1]} u^2 r^d \mu(dr)\nu_r(du) < \infty. \tag{2.1.6}$$

instead of (2.1.1). This would certainly be enough to handle the dual process of Definition 2.1.9 but equation (2.1.6) turns out to be a bad guess. In Section 2.1.2 we saw a notional formulation of the process in which (2.1.1) corresponds to ancestral lineages jumping at finite rate, obtaining a dual system of coalescing compound Poisson process. As one might imagine from our discussion in Section 1.3.3, equation (2.1.6) corresponds to replacing this with a system of coalescing Lèvy processes. We run into difficulty trying to weave together all the lineages into dynamics driven by a Poisson point process. This is most easily seen if we attempt to write down the generator of the basic SΛFV process.

Let us do so on the test function $I_1(\cdot;\Phi)$ for some Φ of the form $(\star, 1)$. The dynamics of definition 2.1.1 suggest that the result will be

$$GI_1(\rho,\Phi) = \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_0^1 \nu_r(du) \int_K \rho(y)(dk) \int_{\mathbb{R}^d} dx \\ \Big[\phi(x)\mathbb{1}\{x \in B_r(y))\}\Big(\langle\chi,(1-u)\rho(x)+u\delta_k\rangle - \langle\chi,\rho(x)\rangle\Big)\Big].$$

Recall our advice in Section 1.3.2 on interpreting the generators of jump processes. The first three integrals select the parameters (y, r, u) of a reproduction event. The integral over K selects a parent type k according to $\rho(y)$. The remaining pieces are the change the process would experience in the corresponding reproduction event as seen through the test function Φ. A little rearrangement leads us to

$$GI_1(\rho,\Phi) = \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_0^1 \nu_r(du) \int_{\mathbb{R}^d} dx \Big[\phi(x)\mathbb{1}\{x \in B_r(y))u\Big(\langle\chi,\rho(y)\rangle - \langle\chi,\rho(x)\rangle\Big)\Big]$$

The best bound on the terms in large round brackets is in general $2||\chi||_\infty$. Since $||\phi||_1 < \infty$ and $\int_{\mathbb{R}^d} \mathbb{1}\{x \in B_r(y)\}dy = C_d r^d$ we see that (2.1.1) is precisely the right bound for GI_1 to be well defined. A similar calculation can be carried out for any $I \in \mathscr{I}$.

In the Λ-Fleming-Viot process we took essentially the test function I_1 without the spatial component, but in the generator (1.3.4) we did not encounter a finiteness problem since without space the parent was selected from 'the same site' as it reproduces into.

Remark 2.1.11 *We end this section with the comment that it is known (but unpublished) that versions of the SΛFV exist with a corresponding notional formulation in which the ancestral lineages are infinite rate pure jump Levy processes. Since the Poisson point process representation breaks down it is not known how to characterise these processes forwards in time. Very little is known about such processes.*

Expressing the SΛFV process as driven by reproduction events taking place in finite regions where each location is affected equally puts a limitation on the type of interaction between the ancestral lineages. As a consequence the Poisson point process driven SΛFV cannot support a sufficiently complex dependency between the ancestral lineages as would be needed for a system of coalescing infinite rate Lèvy processes.

2.1.4 Generalizations

Some possible generalisations of the basic SΛFV are as follows.

1. The parent location in the reproduction event (t, x, r, u) could be sampled according to some distribution on $B_r(x)$. We will refer to this general form as *non-central parenting*. The sampling of parents as in Definition (2.1.1) we refer to as *central parenting* and if the parent is selected uniformly at random from $B_r(y)$ we refer to it as *uniform parenting*.

2. We could vary the shape of the region affected by a reproduction event with parent y and radius r. We refer to this as *reproduction events of varied shape*. In this case we would no longer be able to parametrise the relevant axis of Λ (which is $(0, \infty) \ni r$) as a radius and would require some indexing space.

3. Multiple parents $z_1, \ldots, z_m$ could be sampled. Then, sampling k_i according to $\rho_{t-}(z_i)$, we would have
$$\rho_t(x) = (1 - u_0)\rho_t(x) + \sum_{i=1}^{m} u_i \delta_{k_i}$$
where Λ is modified so as (u_i) is sampled with $u_0 = \sum_{i=1}^{m} u_i$ as the analogue of (2.1.2).

4. We could allow the killing proportion u to depend on the distance of some affected site from the parent site.

5. We could introduce spatial motion in between reproduction events.

6. We could incorporate selection so as the type δ_k chosen to be the parent type affects to what extent the reproduction event takes place.

7. We could introduce mutation. It would be interesting to let $\rho_t(x)$ evolve randomly in between jumps caused by the reproduction events. It is more realistic biologically to have mutation occurring as part of the reproduction events.

Many other generalizations would be possible, beyond those we list above. With applications in mind, it is sensible to think of SΛFV processes as a framework rather than a particular process. Our construction in Chapter 3 (see also Definition 2.3.1) will include a mechanism selecting event shape and parent location, and will also incorporate selection. We could easily modify the prelimiting particle systems of Definition 3.2.4 in Chapter 3 to generate an SΛFV process with all of the above modifications. Doing so would require even more unwieldy formula than those which already appear in Chapter 3, but would not add significant difficulty to the analysis.

2.2 SΛFV literature

The SΛFV was first introduced in Etheridge (2008) in the form of Definition 2.1.1 but with uniform parenting. A construction of this version of the SΛFV process on a torus, the only published construction to date, was given in Barton et al. (2010b).

1. **Etheridge (2008)** was a survey of models used in modern population genetics. It contains a short section introducing the SΛFV and a list of suggestions of possible generalizations of the process.

2. **Berestycki, Etheridge, and Hutzenthaler (2009)** considers a process which is intended to approximate the SΛFV with only finitely many individuals in any bounded region. In contrast to most finite population models the individuals in their model exhibit correlated reproduction. The article is concerned mostly with discussing the long term survival and ergodicity of such models. Proof that these models really do converge to the SΛFV process is to appear in Etheridge and Kurtz (In Preparation).

3. **Barton, Etheridge, and Véber (2010b)** features the first construction of the SΛFV process, in the form of Definition 2.1.1 but on a Torus with uniform parenting. Neither of these features is crucial to their construction, which is an adaptation of Evans (1998) and relies heavily on ancestral lineage duality. The primary aim of Barton et al. (2010b) is to investigate the asymptotic properties of the dual. Note that for the case with uniform reproduction a dual can be defined analogously to our Definition 2.1.9; on a (t, y, r, u) reproduction event sample a parent site uniformly from $B_r(y)$ rather than always taking it to be y. Under different scaling limits they obtain a Kingman coalescent, a general Λ-coalscent and a system of coalescing Brownian motions with a non-local coalescence mechanism.

4. **Limic, Véber, and Wakolbinger (2010)** is work in progess and is intended to provide a lookdown construction of the SΛFV process. They also give a natural way to express the SΛFV process as a measure valued process (i.e. a $\mathcal{M}(\mathbb{R}^d \times K$ valued process) but using a non-standard topology on $\mathcal{M}(\mathbb{R}^d \times K)$ induced by a bijection with Ξ.

5. **Berestycki, Etheridge, and Véber (2010)** is also work in progress and looks at scaling limits of two versions of the SΛFV process. Firstly, a one dimensional version with a heavy tailed distribution on the radii of reproduction events. Secondly (and in all dimensions d), a case where the ancestral lineages are required to have finite variance and are rescaled to Brownian motion. If (in dimensions $d \geqslant 2$) the initial conditions are suitable sparse, the scaling limit of the SΛFV is a super-Brownian motion. We will explore this ourselves in Chapter 4. Berestycki et al. (2010) covers the non-sparse case.

6. **Barton, Etheridge, and Kelleher (2010a)** is concerned with the biological effects the SΛFV was designed to model. A simulation package (which is used in the paper) can be found at `http://homepages.ed.ac.uk/jkellehe/qps.php`. They consider a version of the process in $\mathbb{R}^2$ where each reproduction event affects all of $\mathbb{R}^d$, and the proportion u of individuals replaced at a site y is the value of a Gaussian centred about the parent location.

2.3 Our formulation

Let us first note that the action of the reproduction events in Definition 2.1.1 is homogeneous in both space and time. Suppose we had non-central parenting in the process of Definition 2.1.1 and suppose the site $x \in \mathbb{R}^d$ is a parent site in the reproduction event (t, x, r, u). We could write down an expression for the conditional distribution of the area in which the reproduction event occurs. Thus we see that specifying the distribution by which the parent is chosen inside a reproduction event is really the same thing as fixing the parent site in the event $(t, x, \cdot, u)$ to always be x and specifying how the shape of the reproduction events vary. In fact, this second formulation is the more general.

Informally for now, let I be some set and for each $i \in I$ let E_i be some open subset of $\mathbb{R}^d$. We then specify the rate of the different shape events with a measure on i. A reproduction event (t, x, i, u) occurs in the region $x + E_i$. We will not require $0 \in E_i$.

We will also incorporate viability selection. For us this means that in a reproduction event once we have selected the parent type δ_k we then carry out a further test in which only we permit the reproduction event to take place with some probability $S(i, k)$. We will call this the *selection test*.

Sampling some parameters of a reproduction event from the Poisson point process and others in a more informal fashion corresponds formally to using a larger Poisson point process. As we will see in the next chapter, selecting all the parameters of a reproduction event properly from a Poisson point process makes for complicated notation. For now let us record a statement of our version of the process in the informal style of Definition 2.1.1. Again, to 'each' point $x \in \mathbb{R}^d$ we associate a $\mathscr{P}(K)$ valued process $\rho_t(x)$.

Definition 2.3.1 (Spatial Λ-Fleming-Viot Process, Our Version) *Let $(I, \mathcal{I})$ be a measure space and let $\mu(di)\nu_i(du)$ be a measure on $I \times [0, 1]$ such that for each $i \in I$, $\nu_i \in \mathscr{P}([0, 1])$. Let $E : I \to \mathbb{B}^d$ be some function such that for each $i \in I$, $E(i) = E_i$ is an open set of $\mathbb{R}^d$. Suppose that*

$$\int_{I \times [0,1]} u\mathscr{D}(i)^d \mu(di)\nu_i(du) < \infty. \tag{2.3.1}$$

where $\mathscr{D}(i) = \sup\{|x-y|\,;\, x,y \in E_i\}$ *is the (maximal) diameter of* E_i*. Let* $S : I \times K \to [0,1]$ *be some function. Let* Λ *be a Poisson point process with points* $(t,y,i,u) \in [0,\infty) \times \mathbb{R}^d \times I \times [0,1]$ *of intensity*

$$dt \otimes dy \otimes \mu(di)\nu_i(du).$$

From some initial state ρ_0 *we specify the evolution*

- *If* $(t,y,i,u) \in \Lambda$ *then select* k *according to* $\rho_{t-}(y)$*. Sample an independent Bernoulli random variable with success probability* $S(i,k)$*. Then, on a success, for each* $x \in y + E_i$ *set*

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

  On a failure do nothing.

- *In between the jump times caused by reproduction events affecting* x *which pass the selection test the process* $\rho_t(x)$ *is constant.*

We call (t,y,u,i) *a reproduction event at time* t *in the region* $y + E_i$ *with parent site* y*, parent type* k*, killing proportion* u *and selection success probability* $S(i,k)$*.*

As with Definition 2.1.1 there are technicalities to do with the topology on the state space which will hold us back from giving a formal definition in this chapter. We have additional technicalities to cope with in Definition 2.3.1 concerning measurability of functions mapping out of I. These issues will be addressed in Chapter 3.

Note that we have replaced (2.1.1) with (2.3.1). Writing

$$C_d \int_0^\infty \int_0^1 u r^d \nu_r(du)\mu(dr) = \int_0^\infty \int_0^1 u\mathcal{L}(B_r(0))\nu_r(du)\mu(dr)$$

(recall $\mathcal{L}$ denotes Lebesgue measure) we expect that the natural analogue of (2.3.1) in the setting with (I,E) specifying non-central parenting would be

$$\int_I \int_0^1 u\mathcal{L}(E_i)\nu_i(du)\mu(di). \tag{2.3.2}$$

Whilst this is probably the correct condition, our construction in the following chapter of the process corresponding to Definition 2.3.1 will require the stronger (2.3.1). Since our construction is the only construction to date that deals with selection, we stick to stating (2.3.1) (but see Remark 3.1.3). Note that if the reproduction events are circular and the reproduction is parent centred then (2.3.1) and (2.3.2) are the same condition. In fact (2.3.1) permits one to construct almost all useful examples of the process. We will state a proper existence theorem (with other, less important conditions too) as Theorem 3.1.8.

2.3.1 The effect of selection on duality

It is very important that at this point we make a note of the influence of selection on the duality expressed in Subsection 2.1.2. As a general principle throughout mathematical genetics, selection and ancestral lineage duality are not happy companions.

The reason for this is that when we are tracing back an ancestral lineage in order to know where to jump next we need to know how strongly the potential parents were competing to give birth. If this competition was influenced by genetic type then in order to know the distribution of where to jump we need to know the distribution of types at a time which (because we are tracing in reverse time) is in the future. This causes duality relationships to break down. Thus the construction of the SΛFV in Barton et al. (2010b) is not able to deal with selection.

Remark 2.3.2 *Fleming and Viot (1979) and Etheridge (2000) both obtain uniqueness of the Fleming-Viot process with selection via a Dawson-Girsanov transform of the neutral case. However, Dawson and Kurtz (1982) gives an example of a Fleming-Viot process with a particular type of selection where a dual can be found. A tool especially worthy of note in this approach is the modified lookdown process of Donnelly and Kurtz (1999a) which deals with both selection and mutation.*

Another idea to get around these issues is the ancestral selection graph, introduced in Krone and Neuhauser (1997). It is known (but unpublished) that a dual exists for the SΛFV process with a particular type of (biologically reasonable) selection, involving a choice between two potential parents. This duality corresponds to an ancestral selection graph with random numbers of branches, and we will not describe it in detail here.

As is well documented in Section 4.4 of Ethier and Kurtz (1986), duality is the major tool for proving uniqueness to martingale problems. This will cause us trouble in the next section where we deal with proving existence/uniqueness for our version of the SΛFV process. In this chapter we have studiously avoided mentioning a martingale problem for the SΛFV process, but we will do so in Chapter 3.

Chapter 3

A construction of the Spatial Λ-Fleming-Viot process

In this chapter we give a construction of the SΛFV process defined informally in Definition 2.3.1. We will characterise the process as the solution to an appropriate generator type martingale problem but we will only be able to prove uniqueness in the case without selection. As we suggested by calling Definition 2.3.1 'our formulation', existence of this process is a new result. We now switch out of the informal style of Chapters 1 and 2 to giving fully rigorous proofs of our results. The final section of this chapter is the subject of ongoing work.

We ask that before beginning this chapter the reader becomes familiar with the construction of the state space Ξ in Section 2.1.1 and Proposition 2.1.7 which characterizes the topology of Ξ. We will use the notation from Section 2.1.1 in this Chapter without comment. We ask also that the reader have read Section 2.3, although we will recall the notation used in our formulation of the SΛFV (Definition 2.3.1) at the start of Section 3.1.1.

3.1 Introduction

The only published construction of the SΛFV process to date has already been mentioned and can be found in Barton et al. (2010b). The construction is indirect and relies heavily on ancestral lineage type duality; as it such cannot be extended to the case with selection. In this chapter we construct the SΛFV process of Definition 2.3.1 forwards in time, as a limit of discrete particle systems.

Our construction is extremely flexible and could be easily extended to include essentially any superposition of effects discussed in 2.1.4. It also would be relatively easy to adapt the convergence arguments in this chapter into showing that the SΛFV process (as a $D_{\Xi}[0,\infty)$ valued process) is continuous in both its initial conditions and its parameters. With regards to the parameters which are measures this would correspond to the appropriate weak topology.

Whilst the strategy of the proof is simple, the expressions involved in dealing with the SΛFV generators are unwieldy. We restrict ourselves to only the existence of the SΛFV process of Definition 2.3.1.

3.1.1 Parameters and conditions

We recap the ingredients of Definition 2.3.1. Let $(I, \mathcal{I}, \mu)$ be a measure space equipped with a σ-finite measure μ. Let $E : I \to \mathbb{B}^d$ be a function, and let us write $E_i = E(i)$ and

$$E_i(y) = y + E_i = \{y + z\,;\, z \in E_i\}.$$

Let $\mathcal{E}(di, du) = \mu(di)\nu_i(du)$ be a σ-finite measure on $I \times [0,1]$ such that the conditional measures $\nu_i(du)$ are probability measures on $[0,1]$. Let $S : I \times K \to [0,1]$ be a measurable function.

Remark 3.1.1 *I is just an index for different shapes of reproduction event and E is the function mapping the index to the event shapes. The measures μ and ν_i control the event rate and the killing proportion.*

S specifies which types in K are at a selective advantage. A high value of $S(i,k)$ increases the chance that a parent of type k has a reproduction event affecting an area of shape E_i. We are consistent with terminology of Definition 2.3.1; the probability that a a reproduction even corresponding to (t,x,i,u) with parent type chosen to be k passes the selection test is $S(i,k)$. If S is constant the process is neutral (i.e. without selection).

Remark 3.1.2 *The same construction as given below is valid if S is also allowed to depend on u.*

We intend all the action to take place over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which will, for the most part, remain in the background. As normal, elements of the sample space Ω are denoted by ω.

Let κ be a Ξ valued random variable, which will be our initial state. Recall

$$\mathscr{D}(i) = \sup\{|x-y|\,;\, x,y \in E_i\}$$

is the (maximal) diameter of E_i.

The parameters $(I, E, \mu, \nu, S, \kappa)$ must satisfy

($\mathscr{H}1$) Each E_i is bounded and open.

($\mathscr{H}2$) For each $x \in \mathbb{R}^d$, $i \to \mathbb{1}\{x \in E_i\}$ is measurable.

($\mathscr{H}3$) $\int_I \mu(di) \int_0^1 \nu_i(du)\, u\mathscr{D}(i)^d < \infty$.

($\mathscr{H}4$) The set

$$\left\{x \in \mathbb{R}^d\,;\, \exists (x_m) \subseteq \mathbb{R}^d, x_m \to x \text{ and } \kappa(x_m) \to \kappa(x) \text{ in } \mathcal{M}_F(K)\right\}$$

is almost surely a Lebesgue-null subset of $\mathbb{R}^d$.

Obviously there is something to prove here in that ($\mathscr{H}2$) guarantees enough measurability for ($\mathscr{H}3$) to make sense. We will prove this (and more) in Lemma B.1. We will not make further comment on matters of $\mathcal{I}$-measurability; All the $\mathcal{I}$-measurability required in what follows can be easily deduced from Lemma B.1 and the standard algebra of measurable functions. ($\mathscr{H}4$) is what we require for the natural sequence of lattice approximations to κ to converge to κ.

Remark 3.1.3 *The fact that we require ($\mathscr{H}4$) is a limitation of the discrete space particle system approach. It is the regularity one needs to make the appropriate lattice approximation for the initial states. To get existence (potentially without particle system approximations) the natural condition to expect is simply $\kappa \in \Xi$. ($\mathscr{H}4$) poses little restriction from a biological point of view, since if $\mathbb{R}^d$ can be divided into countably many regions with say, Lipschitz boundary, and on each region $x \to \kappa(x)$ is continuous, ($\mathscr{H}4$) clearly holds.*

For our purposes, condition ($\mathscr{H}3$) could be replaced with the weaker (but unwieldy) condition that for some $\beta \in (0,\infty)$,

$$\int_{I_1} \mu(di) \int_0^1 \nu_i(du) u\mathscr{D}(i)^d < \infty \text{ and } \int_{I_2} \mu(di) \int_0^1 \nu_i(du) u\mathcal{L}(E_i) < \infty$$

where $I_1 = \{i \in I\,;\, \mathscr{D}(i) < \beta\}$ and $I_2 = \{i \in I\,;\, \mathscr{D}(i) \geqslant \beta\}$. In other words, it is enough to have the bound with $u\mathcal{L}(E_i)$ for the large events but for the small events we must use the stronger $u\mathscr{D}(i)^d$.

3.1.2 Results

Definition 3.1.4 *Define a linear operator $G : \mathscr{I} \to C(\Xi)$ by*

$$\begin{aligned} G^\alpha(I_n(\rho,\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu(di) \int_0^1 \nu_i(du) \int_K \rho(y)(dk) \int_{(\mathbb{R}^d)^n} dx_1,\ldots,dx_n \\ S(i,k)\psi(x_1,\ldots,x_n) \left(\prod_{\{j\,;\,x_j \notin E_i(y)\}} \langle \chi_j, \rho(x_j)\rangle\right) \\ \times \left(\prod_{\{j\,;\,j\in E_i(y)\}} \langle \chi_j, (1-u)\rho(x_j) + u\delta_k\rangle - \prod_{\{j\,;\,j\in E_i(y)\}} \langle \chi_j, \rho(x_j)\rangle\right), \end{aligned}$$

and $G(f) = 0$ if $f \in C(\Xi)$ is constant.

Note that this is precisely the generator we would expect to correspond to Definition 2.3.1. The first four integrals select a location y, a shape reproduction event i (in the form of E_i), a killing proportion u and a parent type k. The term $S(i,k)$ modifies the rate at which reproduction events corresponding to y, i, u, and k can take place. The remainder of the expression is the change we would expect to see at such a reproduction event as viewed through the test function $I_n(\cdot;\Phi)$.

We define the Spatial Λ-Fleming-Viot process to be the solution to the following martingale problem. We say 'the' somewhat tongue in cheek since there is currently no proof of uniqueness in the case with selection.

Definition 3.1.5 (Martingale Problem for the SΛFV process.) *The Spatial Λ-Fleming-Viot process with initial distribution κ is the càdlàg Ξ valued process $t \to \rho_t$ such that $\mathbb{P}[\rho_0 = \kappa] = 1$ and*

$$\phi(\rho_t) - \int_0^t G_s(\phi)(\rho_s)ds$$

is a martingale for all $\phi \in \mathscr{I}$.

Our pre-limiting particle systems are indexed by $\alpha \in (0,1]$ and can be informally described as follows. We will give a formal description as Definition 3.2.2. Note that they are not formulated as Ξ-valued processes, which is an issue we will address in Section 3.2.

Define

$$I_0^\alpha = \{i \in I\,;\, \mathscr{D}(i) < \alpha \text{ or } 1/\alpha < \mathscr{D}(i)\}$$

and let μ^α be the measure on $(I, \mathcal{I})$ given by

$$\mu^\alpha(A) = \mu(A \backslash I_0^\alpha).$$

Remark 3.1.6 *We could use any sequence $f(\alpha)$ such that $f(\alpha) \to \infty$ as $\alpha \downarrow 0$ in place of $1/\alpha$ in the definition of I_0^α. The point of μ^α is that we remove the small events and the large events.*

Definition 3.1.7 *Fix $\alpha \in (0,1]$. To each $x \in \alpha\mathbb{Z}^d$ associate a K valued process $t \to \xi_t^\alpha(x)$ with initial state defined by the relation $\delta_{\xi_0(x)} = \kappa(x)$. Let L be a Poisson point process in $[0,\infty) \times \alpha\mathbb{Z}^d \times I \times [0,1]$ of rate*

$$dt \otimes \alpha^d dx \otimes \mu^\alpha(di)\nu_i(du)$$

where dt corresponds to Lebesgue measure on $[0,\infty)$ and dx corresponds to the measure giving each point of $\alpha\mathbb{Z}^d$ mass 1. Define the dynamics as follows.

- *If $(t, y, i, u) \in L$ then, independently of all else, sample k according to $\rho_t(y)$ and a Bernoulli random variable $\mathscr{S}$ with success probability $S(i,k)$. Do the following.*
  - *If $\mathscr{S}$ is a success, for each $x \in E_i(y) \cap \alpha\mathbb{Z}^d$, independently of all else, take a uniform random variable U and if $U < u$ set $\xi_t^\alpha(x) = \xi_{t-}^\alpha(y)$.*
  - *If $\mathscr{S}$ is a failure, do nothing.*
- *The process $t \to \xi_t^\alpha(y)$ is constant in between the jumps caused by the reproduction events above.*

Our main results are:

Theorem 3.1.8 *The martingale problem for the Spatial Λ-Fleming-Viot process has a solution. In fact, any limit point (in $\mathscr{P}(D_\Xi[0,\infty))$) of the set of particle systems defined in Definition 3.1.7 is a solution.*

Theorem 3.1.9 *If S is constant the martingale problem for the Spatial Λ-Fleming-Viot process has at most one solution. In this case, the particle systems of Definition 3.1.7 have a unique limit point.*

Remark 3.1.10 *As we have already commented, we believe Theorem 3.1.9 to be true without the condition that S be constant.*

We discuss the pre-limiting processes in Section 3.2. Proof of Theorem 3.1.8 is given in Section 3.3 and proof of Theorem 3.1.9 is given in Section 3.4.

Recall that at the end of Section 2.2 we remarked that our pre-limiting particle systems would allow for easy modification to produce even more general version of the process. We will not describe the precise modification for each case but hope that the principles of how to discretize the desired limit process are now clear. The convergence argument in each case will be essentially the same as that given below.

The estimates involved in the proof of Theorem 3.1.8 require a huge amount of notation, and the reader who does not wish to verify the proof is advised skip the remainder of this chapter.

3.2 The pre-limiting processes

In this section we set up the pre-limiting particle systems and prove some results about their generators. For the duration of this section fix $\alpha \in (0,1]$. We prove existence and characterize the generator, formulated as a $K^{\alpha\mathbb{Z}^d}$ valued process. Then we translate these results into statements about the Ξ valued version of the same system and take a limit as $\alpha \to 0$.

For $x \in \mathbb{R}^d$ define

$$[x]_\alpha = y \text{ where } y \in \alpha\mathbb{Z}^d \text{ and } x \in [y+\alpha/2, y-\alpha/2)^d.$$

to be (off a null set) the nearest neighbour in $\alpha\mathbb{Z}^d$ of $x \in \mathbb{R}^d$.

In order to properly describe our pre-limiting particle systems of Definition 3.1.7 we need a method for constructing countably many uniform random variables within a larger Poisson point process. Let

$$\mathcal{U} = \{f : \alpha\mathbb{Z}^d \to [0,1]\,;\, f \text{ is measurable.}\}$$

with the $||\cdot||_\infty$ norm and corresponding σ-field. The following result is well known.

Lemma 3.2.1 *There exists a measure U on $\mathcal{U}$ such that if F is a $\mathcal{U}$ valued random variable with law U, the set of functions $\{\omega \mapsto F(x)(\omega)\,;\, x \in \alpha\mathbb{Z}^d\}$ is a set of independent random variables, each of which is uniform on $[0,1]$.*

Our initial state for the limiting process is $\kappa \in \Xi$. For each $\alpha \in (0,1]$ there is a (unique, in distribution) $K^{\alpha\mathbb{Z}^d}$ valued random variable satisfying the relation

$$\delta_{\xi_0^\alpha(x)} = \kappa(x) \tag{3.2.1}$$

Our pre-limiting processes are described as follows.

Definition 3.2.2 *Fix $\alpha \in (0,1]$. The α^{th} process $(\xi_t^\alpha)_{t=0}^\infty$ takes values in $K^{\alpha\mathbb{Z}^d}$ and its value at time $t \in [0,\infty)$ will be denoted ξ_t^α. The initial state ξ_0^α is given by* (3.2.1)*. Let Λ^α be a Poisson point process (independent of κ) with points (t,x,i,u,f,g) in*

$$[0,\infty) \times \alpha\mathbb{Z}^d \times I \times [0,1] \times [0,1]^{\alpha\mathbb{Z}^d} \times [0,1]^{\alpha\mathbb{Z}^d}$$

of rate

$$dt \otimes \alpha^d dx \otimes \mu^\alpha(di)\nu_i(du) \otimes U(df) \otimes U(dg)$$

where dt is Lebesgue measure on $[0,\infty)$ and dx is the measure assigning mass 1 to each lattice point of $\alpha\mathbb{Z}^d$. The dynamics are

- *If $(t,y,i,u,f,g) \in \Lambda^\alpha$, then for every $x \in E_i(y) \cap \alpha\mathbb{Z}^d$, if $S\left(i, \xi_{t-}^\alpha(y)\right) > f(y)$ and $u > g(x)$ set $\xi_t^\alpha(y) = \xi_{t-}^\alpha(x)$.*
- *The process $t \to \xi_t^\alpha(y)$ is constant in between the jumps caused the the reproduction events as above.*

Remark 3.2.3 *By Lemma 3.2.1 this defines the same particle system as Definition 3.1.7.*

Let

$$\mathscr{Y} = \mathbb{R}^d \times I \times [0,1] \times \mathcal{U}^2$$

be the underlying space for Λ^α without the time component. For $x \in \mathbb{R}^d$, $(y,i,u,f,g) \in \mathscr{Y}$ and $\zeta \in K^{\alpha\mathbb{Z}^d}$ let

$$T(x,y,i,u,f,g,\zeta) = \mathbb{1}\Big\{[x]_\alpha \in E_i\left([y]_\alpha\right), u > g\left([x_j]_\alpha\right), S\left(i,\zeta([y]_\alpha)\right) > f\left([y]_\alpha\right)\Big\}.$$

We will usually permit ourselves to suppress dependence on all but x and write $T(x;)$. In words, $T(x;)$ is a test to see if the reproduction event $(\cdot,y,i,u,f,g)$ is successful in overwriting the type of x from state ζ. Let

$$\hat{T}\left(x;\right) = 1 - T\left(x;\right).$$

be the corresponding test for failure.

We equip the space $C(K^{\alpha\mathbb{Z}^d})$ of continuous real valued functions on $K^{\alpha\mathbb{Z}^d}$ with the $||\cdot||_\infty$ norm (which is defined since K is a metric space) and corresponding topology. For $\eta \in K^{\alpha\mathbb{Z}^d}$ set

$$\eta^{y,i,u,f,g}(x) = T(x;)\eta([y]_\alpha) + \hat{T}(x;)\eta(x).$$

In words, $\eta^{y,i,u,f,g}$ is η after being affected by a reproduction event $(\cdot,y,i,u,f,g)$.

Lemma 3.2.4 *The $K^{\alpha\mathbb{Z}^d}$ valued process $\xi_t^\alpha(\cdot)$ is well defined by Definition 3.2.2. Further, ξ^α has Markov pre-generator Ω^α defined by*

$$\Omega^\alpha J(\eta) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg)\left[J(\eta^{y,i,u,f,g}) - J(\eta)\right]$$

for $J \in \Delta^\alpha$ where

$$\Delta^\alpha = \left\{ J \in C\left(K^{\alpha\mathbb{Z}^d}\right) ; \sum_{x \in \alpha\mathbb{Z}^d} \sup\left\{ |J(\eta) - J(\zeta)| ; \eta,\zeta \in K^{\alpha\mathbb{Z}^d} \text{ and for } y \neq x, \eta(y) = \zeta(y)\right\} < \infty \right\}.$$

The closure $\overline{\Omega} : C(K^{\alpha\mathbb{Z}^d}) \to C(K^{\alpha\mathbb{Z}^d})$ of Ω is a Markov generator.

PROOF: Our proof of this Lemma will rely on results from Ligget (1985). See Appendix C. ■

We now move straight on to considering the particle system of Definition 3.2.2 as a Ξ valued process. Define a process $t \to \rho_t^\alpha$ by

$$\rho_t^\alpha(x) = \delta_{\xi_t^\alpha([x]_\alpha)}.$$

Note that taking $t = 0$ gives $\rho_0^\alpha(x) = \kappa([x]_\alpha)$. Let

$$\mathscr{X}^\alpha = \{\rho \in \Xi ; \text{ for almost all } x \in \mathbb{R}^d, \rho(x) = \rho([x]_\alpha) \text{ and for some } k_x \in K, \rho(x) = \delta_{k_x}\}.$$

Lemma 3.2.5 *$t \to \rho_t^\alpha$ is Ξ valued and almost surely càdlàg . For all t, $\rho_t^\alpha \in \mathscr{X}^\alpha$.*

PROOF: The fact that $\rho_t^\alpha \in \mathscr{X}^\alpha \subseteq \Xi$ is immediate from the definition of ρ^α. It remains only to prove that ρ^α almost surely has càdlàg paths. Let $\psi : \mathbb{R}^d \to \mathbb{R}$ be continuous with compact support and let $\chi \in C(K)$. By Definition 3.2.2 for each $\alpha \in (0,1]$ the rate of reproduction events affecting the bounded region $\text{supp}(\psi)$ is finite and so

$$\mathbb{P}\left[\forall t \in [0,\infty)\, \exists \epsilon > 0\, \forall s \in [t, t+\epsilon), \rho_s(x) = \rho(x) \text{ for all } x \in \text{supp}(\psi)\right] = 1.$$

Hence

$$\mathbb{P}\left[\forall t \in [0,\infty), \text{as } s \downarrow t, \int_{\mathbb{R}^d} \psi(x)\left\langle \chi, \rho_s^\alpha(x)\right\rangle dx \to \int_{\mathbb{R}^d} \psi(x)\left\langle \chi, \rho_t^\alpha(x)\right\rangle dx\right] = 1$$

so by Lemma 2.1.7,

$$\mathbb{P}\left[\forall t \in [0,\infty), \text{ if } s \downarrow t \text{ then } \rho_s^\alpha \to \rho_t^\alpha\right] = 1.$$

A similar argument covers the left limits. ■

For $\rho \in \mathscr{X}^\alpha$, let $p_\rho(y) = k$ where $k \in K$ is unique such that $\rho(y) = \rho([y]_\alpha) = \delta_k$. Note in particular that by Lemma 3.2.5,

$$\delta_{p_{\rho_t(y)}} = \rho_t(y).$$

With mild abuse of notation let us write

$$T(x,y,i,u,f,g,\rho) = \mathbb{1}\left\{[x]_\alpha \in E_i\left([y]_\alpha\right), u > g\left([x_j]_\alpha\right), S\left(i, p_\rho([y]_\alpha)\right) > f\left([y]_\alpha\right)\right\}.$$

for $\rho \in \mathscr{X}^\alpha$ and similarly for $\hat{T}(x;) = 1 - T(x;)$.

Let G^α be the linear operator defined on $\mathscr{I} \to C(\Xi)$ by

$$\begin{aligned} G^\alpha(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \\ \psi(x_1,\dots,x_n)\Bigg\{ \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \left[T(x_j;)\rho\left([y]_\alpha\right) + \hat{T}(x_j;)\rho\left([x_j]_\alpha\right)\right] \right\rangle \\ - \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \rho([x_j]_\alpha) \right\rangle \Bigg\} \end{aligned} \tag{3.2.2}$$

and $G(I_0) = 0$ if $I_0 \in C(\Xi)$ is constant.

Remark 3.2.6 *We will not prove whether or not G^α extends to a Markov generator. It is relatively easy to show (using, for example, Theorem 2.2 of Ligget (1985)) that G^α is an unbounded Markov pre-generator.*

Proposition 3.2.7 *For all $h \in \mathscr{I}$,*

$$h(\rho_t^\alpha) - \int_0^t G^\alpha h(\rho_s^\alpha) ds \tag{3.2.3}$$

is a martingale.

Proof: Since a linear combinations of martingales is a martingale it suffices to prove (3.2.3) for $h = I_n(\cdot;\Phi)$ where for some $n \in \mathbb{N}$ Φ has the form $(\star, n)$.

So let Φ have the form $(\star, n)$. Define $J_n^\alpha(\cdot;\Phi) : K^{\alpha\mathbb{Z}^d} \to \mathbb{R}$ by

$$J_n^\alpha(\eta;\Phi) = \int_{(\mathbb{R}^d)^n} \psi(x_1,\dots,x_n) \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \delta_{\eta([x_j]_\alpha)} \right\rangle dx_1 \dots dx_n.$$

Note that $\delta_{\xi_t^\alpha([x_j]_\alpha)} = \rho_t^\alpha([x_j]_\alpha)$ and so

$$J_n^\alpha(\xi_t^\alpha;\Phi) = I_n(\rho_t^\alpha;\Phi). \tag{3.2.4}$$

By Lemma 3.2.4,

$$\begin{aligned} \Omega^\alpha J_n(\xi_t^\alpha;\Psi) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \Bigg\{ \psi(x_1,\dots,x_n) \\ \times \left(\left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \delta_{T([x_j]_\alpha;)\xi_t^\alpha([y]_\alpha) + \hat{T}([x_j]_\alpha;)\xi_t^\alpha([x_j]_\alpha)} \right\rangle - \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \delta_{\xi_t^\alpha([x_j]_\alpha)} \right\rangle \right) \Bigg\}. \end{aligned}$$

Since $\rho_t^\alpha \in \mathscr{X}^\alpha$,

$$\delta_{T([x_j]_\alpha;)\xi_t^\alpha([y]_\alpha) + \hat{T}([x_j]_\alpha;)\xi_t^\alpha([x_j]_\alpha)} = T([x_j]_\alpha;)\rho_t^\alpha([y]_\alpha) + \hat{T}([x_j]_\alpha;)\rho_t^\alpha([x_j]_\alpha),$$

so as

$$\Omega^\alpha J_n(\xi_t^\alpha;\Psi) = G^\alpha I_n(\rho_t^\alpha;\Psi). \tag{3.2.5}$$

Using (3.2.4) and (3.2.5), (3.2.3) is a martingale if and only if

$$J_n^\alpha(\xi_t^\alpha) - \int_0^t \Omega^\alpha J_n^\alpha(\xi_s^\alpha;\Phi)ds \tag{3.2.6}$$

is a martingale.

Recall that our definition of the form $(\star, n)$ required ψ to have compact support. Hence if $x \in \alpha\mathbb{Z}^d \backslash \operatorname{supp}(\psi)$ and $\eta(y) = \zeta(y)$ for all $y \neq x$ we have $J_n^\alpha(\eta;\Psi) = J_n^\alpha(\zeta;\Psi)$. Note also that for any $\eta \in K^{\alpha\mathbb{Z}^d}$,

$$|J_n^\alpha(\eta,;\Psi)| \leqslant ||\psi||_1 \prod_{j=1}^n ||\chi_j||_\infty .$$

From these two observations it follows that

$$\begin{aligned}\sup\Big\{|J(\eta) - J(\zeta)|\,;\, \eta,\zeta \in K^{\alpha\mathbb{Z}^d} \text{ and for } y \neq x, \eta(y) = \zeta(y)\Big\}\\ \leqslant 2\#(\alpha\mathbb{Z}^d \cap \operatorname{supp}(\psi))||\psi||_1 \prod_{j=1}^n ||\chi_j||_\infty < \infty\end{aligned}$$

and hence $J_n^\alpha(\cdot;\Psi) \in \Delta^\alpha$. In Lemma 3.2.4 we showed that the closure of the pre-generator Ω^α is a Markov Generator for ξ_t^α. By the Hille-Yosida theorem there is a semigroup corresponding to this Markov generator. The fact that (3.2.6) is a martingale now follows from Lemma 2.1.7 of Ethier and Kurtz (1986). ■

Let

$$A_n^\alpha = \left\{x \in (\mathbb{R}^d)^n\,;\, \forall j \neq k, [x_j]_\alpha \neq [x_k]_\alpha\right\}.$$

As $\alpha \to 0$ note that $\mathbb{1}_{A_n^\alpha} \to 1$ pointwise.

Lemma 3.2.8 *Let $n \in \mathbb{N}$. There exists a bounded real valued function b_n^α such that for all $\rho \in \mathscr{X}^\alpha$ and all Φ of the form $(\star, n)$*

$$\begin{aligned}G^\alpha(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_K \rho(y)(dk) \int_{(\mathbb{R}^d)^n} dx_1,\ldots,dx_n\\ \Bigg\{\mathbb{1}\{(x_j) \in A_n^\alpha\} S(i,p_\rho(y))\psi(x_1,\ldots,x_n)\left(\prod_{\{j\,;\,[x_j]_\alpha \notin E_i([y]_\alpha)\}} \langle\chi_j, \rho(x_j)\rangle\right)\\ \times\left(\prod_{\{j\,;\,[x_j]_\alpha \in E_i([y]_\alpha)\}} \langle\chi_j, (1-u)\rho(x_j) + u\delta_{p_\rho(y)}\rangle - \prod_{\{j\,;\,j \in E_i([y]_\alpha)\}} \langle\chi_j,\rho(x_j)\rangle\right)\\ + \mathbb{1}\{(x_j) \notin A_n^\alpha\} b_n^\alpha(x_1,\ldots,x_n,y,i,u,\Phi)\Bigg\}.\end{aligned} \tag{3.2.7}$$

where

$$|b_n^\alpha(x_1,\ldots,x_n,y,i,u,\Phi)| \leqslant u\mathbb{1}\Big\{\exists j, [x_j]_\alpha \in E_i([y]_\alpha)\Big\}|\psi(x_1,\ldots,x_n)|2n\prod_{j=1}^n ||\chi_j||_\infty < \infty. \tag{3.2.8}$$

Further, there exists $C_n \in (0,\infty)$ such that for all $\rho \in \Xi$ and Φ of the form $(\star, n)$,

$$|G^\alpha(I_n(\rho;\Phi))| \leqslant C_n||\varphi||_1 \prod_{j=1}^n ||\chi_j||_\infty . \tag{3.2.9}$$

Proof: We first note that $\mathbb{1}\{S(i,p_\rho([y]_\alpha) > g([y]_\alpha)\}$ does not depend on j, so if $S(i,p_\rho([y]_\alpha) \leqslant g([y]_\alpha)$ we have

$$\left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \Big[T(x_j;)\rho\left([y]_\alpha\right) + \hat{T}(x_j;)\rho\left([x_j]_\alpha\right)\Big]\right\rangle - \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \rho([x_j]_\alpha)\right\rangle = 0$$

and thus from (3.2.2),

$$G^\alpha(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n$$
$$\psi(x_1,\dots,x_n)\mathbb{1}\{S(i,p_\rho([y]_\alpha)) > g([y]_\alpha)\}$$
$$\times \left\{ \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \left[T'(x_j;)\rho\left([y]_\alpha\right) + \hat{T}'(x_j;)\rho\left([x_j]_\alpha\right)\right]\right\rangle - \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \rho([x_j]_\alpha)\right\rangle \right\} \tag{3.2.10}$$

where

$$T'(x;) = T'(x,y,i,u,f,g,\rho) = \mathbb{1}\left\{[x]_\alpha \in E_i\left([y]_\alpha\right), u > g\left([x_j]_\alpha\right)\right\}$$

and $\hat{T}'(x;) = 1 - T'(x;)$. For $m = 0, 1$ define

$$\sigma(m) = \{j \in \{1,\dots,n\}\,;\, [x_j]_\alpha \in E_i([y]_\alpha)\}$$

(we supress dependence on α, n, (x_j) and y) and then

$$G^\alpha(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n$$
$$\psi(x_1,\dots,x_n)\mathbb{1}\{S(i,p_\rho([y]_\alpha)) > g([y]_\alpha)\}\left(\prod_{j\in\sigma(0)} \langle \chi_j, \rho([x_j]_\alpha)\rangle \right)$$
$$\times \left\{ \left\langle \bigtimes_{j\in\sigma(1)} \chi_j, \bigotimes_{j\in\sigma(1)} \left[\mathbb{1}\{u > f([x_j]_\alpha)\}\rho([y]_\alpha) + \mathbb{1}\{u \leqslant f([x_j]_\alpha)\}\rho([x_j]_\alpha)\right]\right\rangle \right.$$
$$\left. - \left\langle \bigtimes_{j\in\sigma(1)} \chi_j, \bigotimes_{j\in\sigma(1)} \rho([x_j]_\alpha)\right\rangle \right\}. \tag{3.2.11}$$

We seek to perform the integration over $\mathcal{U}^2$. In the above expression we have isolated the piece depending on g, and by Lemma 3.2.1,

$$\int_{\mathcal{U}} U(dg)\, \mathbb{1}\{S(i,p_\rho([y]_\alpha)) > g([y]_\alpha)\} = S(i,p_\rho([y]_\alpha)). \tag{3.2.12}$$

We have also isolated the piece depending on f, namely the expression in curly brackets making up third and fourth lines of (3.2.11). Note that the term from the fourth line does not depend on f and $\mathcal{U}(U) = 1$ by Lemma 3.2.1. We now approach the term from the third line. Let

$$\mathscr{K}_{\sigma(1)} = \{\tau : \sigma(1) \to \{0,1\}\}$$

and define

$$\mathcal{T}_1(x,y,u) = \mathbb{1}\{u > f([x_j]_\alpha)\},$$
$$\mathcal{T}_0(x,y,u) = 1 - \mathcal{T}(x,y,u),$$

$\mathcal{S}_1(x,y) = \rho([y]_\alpha)$ and finally $\mathcal{S}_0(x,y) = \rho([x_j]_\alpha)$. Then

$$\int_{\mathcal{U}} U(df) \left\{ \left\langle \bigtimes_{j\in\sigma(1)} \chi_j, \bigotimes_{j\in\sigma(1)} \left[\mathbb{1}\{u > f([x_j]_\alpha)\}\rho([y]_\alpha) + \mathbb{1}\{u \leqslant f([x_j]_\alpha)\}\rho([x_j]_\alpha)\right]\right\rangle \right\}$$
$$= \int_{\mathcal{U}} U(df) \left\{ \sum_{\tau\in\mathscr{K}_{\sigma(1)}} \int_{K^{|\sigma(1)|}} \bigotimes_{j\in\sigma(1)} \mathcal{S}_{\tau(j)}(x,y)(dk_j) \left[\prod_{j\in\sigma(1)} \chi_j(k_j)\mathcal{T}_{\tau(j)}(x,y,u)\right]\right\}$$
$$= \sum_{\tau\in\mathscr{K}_{\sigma(1)}} \int_{K^{|\sigma(1)|}} \bigotimes_{j\in\sigma(1)} \mathcal{S}_{\tau(j)}(x,y)(dk_j) \left[\left(\prod_{j\in\sigma(1)} \chi_j(k_j)\right) \int_{\mathcal{U}} U(df) \left\{ \prod_{j\in\sigma(1)} \mathcal{T}_{\tau(j)}(x,y,u)\right\}\right]. \tag{3.2.13}$$

To obtain (3.2.13) from the line immediately preceding (3.2.13) we note that all the products over j are separate entities and we are not really overusing the variable j. We are now able to evaluate the integral in (3.2.13) over $\mathcal{U}$, but recalling Lemma 3.2.1 and the definition of $\mathcal{T}_m(x,y,u)$ we require to know which of the x_j are such that $[x_{j_1}]_\alpha = [x_{j_2}]$. Looking at the case where the $[x_j]_\alpha$ are distinct we obtain from Lemma 3.2.1 that

$$\mathbb{1}\{(x_j) \in A_n^\alpha\} \int_{\mathcal{U}} U(df) \left\{ \prod_{j \in \sigma(1)} \mathcal{T}_{\tau(j)}(x,y,u) \right\} = \mathbb{1}\{(x_j) \in A_n^\alpha\} u^{|\tau|} (1-u)^{|\sigma(1)|-|\tau|} \tag{3.2.14}$$

where we define $|\tau| = \sum_{j \in \sigma(1)} \tau(j)$. For the case involving $\mathbb{1}\{(x_j) \notin A_n^\alpha\}$ we create the error term b_n^α described in the statement of this Lemma. So, by (3.2.14)

$$\begin{aligned}
&\mathbb{1}\{(x_j) \in A_n^\alpha\} \times (3.2.13) \\
&= \mathbb{1}\{(x_j) \in A_n^\alpha\} \sum_{\tau \in \mathscr{K}_{\sigma(1)}} \int_{K^{|\sigma(1)|}} \bigotimes_{j \in \sigma(1)} \mathcal{S}_{\tau(j)}(x,y)(k_j) \left[\left(\prod_{j \in \sigma(1)} \chi_j(k_j) \right) u^{|\tau|} (1-u)^{|\sigma(1)|-|\tau|} \right] \\
&= \mathbb{1}\{(x_j) \in A_n^\alpha\} \sum_{\tau \in \mathscr{K}_{\sigma(1)}} \int_{K^{|\sigma(1)|}} \bigotimes_{j \in \sigma(1)} \mathcal{S}'_{\tau(j)}(x,y)(k_j) \left[\prod_{j \in \sigma(1)} \chi_j(k_j) \right] \\
&= \mathbb{1}\{(x_j) \in A_n^\alpha\} \left\langle \bigtimes_{j \in \sigma(1)} \chi_j, \bigotimes_{j \in \sigma(1)} \Big[u\rho([y]_\alpha) + (1-u)\rho([x_j]_\alpha) \Big] \right\rangle
\end{aligned} \tag{3.2.15}$$

where $\mathcal{S}'_1(x,y) = u\rho([y]_\alpha)$ and $\mathcal{S}'_1(x,y) = (1-u)\rho([x]_\alpha)$. We now look to be in good shape! Putting (3.2.12) and (3.2.15) into (3.2.11) and creating an error term to account for using $\mathbb{1}\{(x_j) \in A_n^\alpha\}$ in (3.2.14),

$$\begin{aligned}
G^\alpha(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{(\mathbb{R}^d)^n} dx_1, \ldots, dx_n \\
\Bigg\{ \mathbb{1}\{(x_j) \in A_n^\alpha\} S\left(i, p_\rho([y]_\alpha)\right) \psi(x_1, \ldots, x_n) \left(\prod_{\{j\,;\,[x_j]_\alpha \notin E_i([y]_\alpha)\}} \langle \chi_j, \rho([x_j]_\alpha) \rangle \right) \\
\times \left(\prod_{\{j\,;\,[x_j]_\alpha \in E_i([y]_\alpha)\}} \langle \chi_j, (1-u)\rho([x_j]_\alpha) + u\rho([y]_\alpha) \rangle - \prod_{\{j\,;\,j \in E_i([y]_\alpha)\}} \langle \chi_j, \rho([x_j]_\alpha) \rangle \right) \\
+ \mathbb{1}\{(x_j) \notin A_n^\alpha\} b_n^\alpha(x_1, \ldots, x_n, y, i, u) \Bigg\}.
\end{aligned} \tag{3.2.16}$$

where b_n^α is given by

$$\begin{aligned}
b_n^\alpha(x_1, \ldots, x_n, y, i, u, \rho, \Psi) = \int_{\mathcal{U}^2} U(df) U(dg) \Bigg\{ \psi(x_1, \ldots, x_n) \mathbb{1}\{S(i, p_\rho([y]_\alpha) > g([y]_\alpha)\} \\
\times \left(\left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \Big[T(x_j;)\rho\left([y]_\alpha\right) + \hat{T}(x_j;)\rho\left([x_j]_\alpha\right) \Big] \right\rangle - \left\langle \bigtimes_{j=1}^n \chi_j, \bigotimes_{j=1}^n \rho([x_j]_\alpha) \right\rangle \right) \Bigg\}.
\end{aligned} \tag{3.2.17}$$

The fact that $\rho \in \mathscr{X}^\alpha$ means that for almost all $y \in \mathbb{R}^d$, $\int_K \rho(y)(dk) = \delta_{p_\rho([y]_\alpha)} = \delta_{p_\rho(y)}$. Thus (3.2.16) is

equal to

$$G^{\alpha}(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d} dy \int_I \mu^{\alpha}(di) \int_0^1 \nu_i(du) \int_K \rho(y)(dk) \int_{(\mathbb{R}^d)^n} dx_1, \ldots, dx_n$$
$$\left\{ \mathbb{1}\{(x_j) \in A_n^{\alpha}\} S(i,k)\, \psi(x_1,\ldots,x_n) \left(\prod_{\{j\,;\,[x_j]_{\alpha} \notin E_i([y]_{\alpha})\}} \langle \chi_j, \rho([x_j]_{\alpha})\rangle \right) \right.$$
$$\times \left(\prod_{\{j\,;\,[x_j]_{\alpha} \in E_i([y]_{\alpha})\}} \langle \chi_j, (1-u)\rho([x_j]_{\alpha}) + u\delta_k)\rangle - \prod_{\{j\,;\,j \in E_i([y]_{\alpha})\}} \langle \chi_j, \rho([x_j]_{\alpha})\rangle \right)$$
$$\left. + \mathbb{1}\{(x_j) \notin A_n^{\alpha}\} b_n^{\alpha}(x_1,\ldots,x_n,y,i,u) \right\}.$$

which is precisely (3.2.7).

All that remains to do is prove (3.2.8) and (3.2.9). Let us note first that if $[x_j] \notin E_i([y]_{\alpha})$ for each $j = 1,\ldots,n$ then the term in the second line of (3.2.17) is zero. Similarly if $u \leqslant f([x_j]_{\alpha})$ for all j Then the second line of (3.2.17) is zero. So from (3.2.17) we obtain

$$|b_n^{\alpha}(x_1,\ldots,x_n,y,i,u,\rho,\Psi)|$$
$$\leqslant \int_{\mathcal{U}^2} U(df)U(dg) \left\{ \mathbb{1}\Big\{\exists j, [x_j]_{\alpha} \in E_i([y]_{\alpha})\Big\} \mathbb{1}\Big\{\exists j, u > f([x_j]_{\alpha})\Big\} |\psi(x_1,\ldots,x_n)| 2 \prod_{j=1}^n ||\chi_j||_{\infty} \right\}.$$

Carrying out the integral

$$\int_{\mathcal{U}^2} U(df)U(dg)\mathbb{1}\Big\{\exists j, u > f([x_j]_{\alpha})\Big\} \leqslant \int_{\mathcal{U}^2} U(df)U(dg) \sum_{j=1}^n \mathbb{1}\{u > f([x_j]_{\alpha})\}$$
$$= \sum_{j=1}^n \int_{\mathcal{U}} U(df)\, \mathbb{1}\{u > f([x_j]_{\alpha})\}$$
$$= nu$$

by Lemma 3.2.1 proves (3.2.8). Similar considerations obtain from (3.2.10) that

$$|G^{\alpha}(I_n(\rho,\Phi))| \leqslant \int_{\mathbb{R}^d} dy \int_I \mu^{\alpha}(di) \int_0^1 \nu_i(du) \int_{(\mathbb{R}^d)^n} \left[u\mathbb{1}\Big\{\exists j, [x_j]_{\alpha} \in E_i([y]_{\alpha})\Big\} |\psi(x_1,\ldots,x_n)| 2n \prod_{j=1}^n ||\chi_j||_{\infty} \right].$$

We note that

$$\int_{\mathbb{R}^d} dy\, \mathbb{1}\Big\{\exists j, [x_j]_{\alpha} \in E_i([y]_{\alpha})\Big\} \leqslant \int_{\mathbb{R}^d} dy \sum_{j=1}^n \mathbb{1}\{[x_j]_{\alpha} \in E_i([y]_{\alpha})\}$$
$$= \sum_{j=1}^n \int_{\mathbb{R}^d} dy\, \mathbb{1}\{-[y]_{\alpha} \in E_i(-[x_j]_{\alpha})\}$$
$$= \sum_{j=1}^n \int_{\mathbb{R}^d} dy\, \mathbb{1}\{-[y]_{\alpha} \in E_i\}$$
$$= n \int_{\mathbb{R}^d} dy\, \mathbb{1}\{|y| \leqslant \mathscr{D}(i) + \alpha\}$$

which gives us

$$
\begin{aligned}
|G^\alpha(I_n(\rho,\Phi))| &\leqslant 2n^2\left(||\varphi||_1\prod_{j=1}^n||\chi_j||_\infty\right)\int_I\mu^\alpha(di)\int_0^1\nu_i(du)\,u\int_{\mathbb{R}^d}dy\,\mathbb{1}\{|y|\leqslant\mathscr{D}(i)+\alpha, i\notin I_n^\alpha\}\\
&\leqslant 2n^2\left(||\varphi||_1\prod_{j=1}^n||\chi_j||_\infty\right)\int_I\mu^\alpha(di)u\int_{\mathbb{R}^d}dy\,\mathbb{1}\{|y|\leqslant 2\mathscr{D}(i)\}\\
&\leqslant 2n^2C\left(||\varphi||_1\prod_{j=1}^n||\chi_j||_\infty\right)\int_I\mu^\alpha(di)uD(\mathscr{I})^d<\infty
\end{aligned}
$$

where to get from the second to third lines we used that $i\in I\backslash I_0^\alpha$ implies $\alpha\leqslant\mathscr{D}(i)$. Finiteness follows from ($\mathscr{H}3$). ■

The corresponding result for G is much easier to prove. But note that we have an extra term in (3.2.18) since we are also comparing μ to μ^α.

Lemma 3.2.9 *Let $n\in\mathbb{N}$. There exists a bounded real valued function b_n such that for all $\rho\in\Xi$ and all Φ of the form $(\star,n)$*

$$
\begin{aligned}
G(I_n(\rho;\Phi)) = \int_{\mathbb{R}^d}dy\int_I\mu^\alpha(di)\int_0^1\nu_i(du)\int_K\rho(y)(dk)\int_{(\mathbb{R}^d)^n}dx_1\dots dx_n&\\
\Bigg\{\mathbb{1}\{(x_j)\in A_n^\alpha\}S(i,p_\rho(y))\psi(x_1,\dots,x_n)\left(\prod_{\{j\,;\,x_j\notin E_i(y)\}}\langle\chi_j,\rho(x_j)\rangle\right)&\\
\times\left(\prod_{\{j\,;\,x_j\in E_i(y)\}}\langle\chi_j,(1-u)\rho(x_j)+u\delta_k\rangle-\prod_{\{j\,;\,j\in E_i(y)\}}\langle\chi_j,\rho(x_j)\rangle\right)&\\
+\mathbb{1}\{(x_j)\notin A_n^\alpha\}b_n(x_1,\dots,x_n,y,i,u,\Phi)&\\
+\left(2n||\varphi||_1\prod_{j=1}^n||\chi_j||_\infty\right)\int_{I_0^\alpha}\mu(di)\int_0^1\nu_i(du)\,u\mathscr{D}(i)^d\Bigg\}.&
\end{aligned}
\tag{3.2.18}
$$

where

$$
|b_n^\alpha(x_1,\dots,x_n,y,i,u,\Phi)|\leqslant u\mathbb{1}\Big\{\exists j,x_j\in E_i(y)\Big\}|\psi(x_1,\dots,x_n)|2\prod_{j=1}^n||\chi_j||_\infty.
\tag{3.2.19}
$$

Further, there exists $C_n\in(0,\infty)$ such that for all $\rho\in\Xi$ and Φ of the form $(\star,n)$,

$$
|G^\alpha(I_n(\rho;\Phi))|\leqslant C_n||\varphi||_1\prod_{j=1}^n||\chi_j||_\infty.
\tag{3.2.20}
$$

PROOF: First let us split off the comparison between μ and μ^α. To ease the length of our formulae let

$$
\begin{aligned}
\mathscr{F}_n(y,i,u,k,\Phi) = \int_{(\mathbb{R}^d)^n}dx_1,\dots,dx_n\Bigg[S(i,k)\psi(x_1,\dots,x_n)\left(\prod_{\{j\,;\,x_j\notin E_i(y)\}}\langle\chi_j,\rho(x_j)\rangle\right)&\\
\times\left(\prod_{\{j\,;\,j\in E_i(y)\}}\langle\chi_j,(1-u)\rho(x_j)+u\delta_k\rangle-\prod_{\{j\,;\,j\in E_i(y)\}}\langle\chi_j,\rho(x_j)\rangle\right)\Bigg].&
\end{aligned}
$$

If for all $j=1,\dots,n$ one has $x_j\notin E_i(y)$ then the second line of the above is expression zero and if j_0 is

such that $x_{j_0} \in E_i(y)$,

$$\begin{aligned}
&\left| \prod_{\{j\,;\,j\in E_i(y)\}} \langle \chi_j, (1-u)\rho(x_j) + u\delta_k\rangle - \prod_{\{j\,;\,j\in E_i(y)\}} \langle \chi_j, \rho(x_j)\rangle \right| \\
&\qquad \leqslant \left(2 \prod_{\substack{j=1 \\ j\neq j_0}}^{n} ||\chi_j||_\infty \right) \langle \chi_{j_0}, -u\rho(x) + u\rho(y)\rangle \\
&\qquad \leqslant u \left(2 \prod_{j=1}^{n} ||\chi_j||_\infty \right)
\end{aligned}$$

Thus we have the bound

$$|\mathscr{F}_n(y,i,u,k,\Phi)| \leqslant u \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \left[\mathbb{1}\{\exists j, x_j \in E_i(y)\} \psi(x_1,\dots,x_n) 2 \prod_{j=1}^{n} ||\chi||_\infty \right]. \tag{3.2.21}$$

From Definition 3.1.4, recalling that $\mu^\alpha(A) = \mu(A \backslash I_0^\alpha)$,

$$\begin{aligned}
&\left| G(I_n(\rho;\Phi)) - \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_K \rho(y)(dk) \mathscr{F}_n(y,i,u,k,\Phi) \right| \\
&\quad = \int_{\mathbb{R}^d} dy \int_{I_0^\alpha} \mu(di) \int_0^1 \nu_i(du) \int_K \rho(y)(dk) \mathscr{F}_n(y,i,u,k;\Phi) \\
&\quad \leqslant \int_{I_0^\alpha} \mu(di) \int_0^1 \nu_i(du) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \\
&\qquad \left[u \sum_{j=1}^{n} \left(\int_{\mathbb{R}^d} dy\, \mathbb{1}\{x_j \in E_i(y)\} \right) \psi(x_1,\dots,x_n) 2 \prod_{j=1}^{n} ||\chi||_\infty \right] \\
&\quad \leqslant \left(2n ||\psi||_1 \prod_{j=1}^{n} ||\chi||_\infty \right) \int_{I_0^\alpha} \mu(di) \int_0^1 \nu_i(du) \left[u\mathcal{L}(E_i) \right]
\end{aligned} \tag{3.2.22}$$

where we used (3.2.21) to get from the second to third lines.

We now split off the term with $\mathbb{1}\{(x_j) \in A_n^\alpha\}$. From (3.2.22) we obtain that (3.2.18) is satisfied with

$$\begin{aligned}
b_n^\alpha(x_1,\dots,x_n,y,i,u,\Phi) = S(i,p_\rho(y))\psi(x_1,\dots,x_n) &\left(\prod_{\{j\,;\,x_j\notin E_i(y)\}} \langle \chi_j, \rho(x_j)\rangle \right) \\
&\times \left(\prod_{\{j\,;\,x_j\in E_i(y)\}} \langle \chi_j, (1-u)\rho(x_j) + u\delta_{p_\rho(y)}\rangle - \prod_{\{j\,;\,j\in E_i(y)\}} \langle \chi_j, \rho(x_j)\rangle \right)
\end{aligned}$$

We bound this using the same technique as we did on $\mathscr{F}(y,i,u,k)$. The bottom line of the above expression is only non zero if for some $j = 1,\dots,n$ we have $x_j \in E_i(y)$. Fix j_0 to be such a j. So

$$\begin{aligned}
|b_n^\alpha(x_1,\dots,x_n,y,i,u,\Phi)| &\leqslant \mathbb{1}\{\exists j, x_j \in E_i(y)\} |\psi(x_1,\dots,x_n)| \left(2 \prod_{\substack{j=1 \\ j\neq j_0}}^{n} ||\chi_j||_\infty \right) \left| \langle \chi_{j_0}, -u\rho(x) + u\rho(y)\rangle \right| \\
&\leqslant \mathbb{1}\{\exists j, x_j \in E_i(y)\} |\psi(x_1,\dots,x_n)| \left(2 \prod_{j=1}^{n} ||\chi_j||_\infty \right) u
\end{aligned}$$

which proves (3.2.19). Similar considerations applied to the expression in Definition (3.1.4) give us that

$$\begin{aligned}
|G\, I_n(\rho,\Psi)| &\leqslant \int_{\mathbb{R}^d} dy \int_I \mu(di) \int_0^1 \nu_i(du) \int_K \rho(y)(dk) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \\
&\qquad u\mathbb{1}\{\exists j, x_j \in E_i(y)\} |\psi(x_1,\dots,x_n)| \left(2\prod_{j=1}^n ||\chi_j||_\infty\right) \\
&\leqslant \int_I \mu(di) \int_0^1 \nu_i(du) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \\
&\qquad u|\psi(x_1,\dots,x_n)| \left(2\prod_{j=1}^n ||\chi_j||_\infty\right) \sum_{j=1}^n \int_{\mathbb{R}^d} dy \mathbb{1}\{x_j \in E_i(y)\} \\
&\leqslant 2n||\psi||_1 \left(\prod_{j=1}^n ||\chi_j||_\infty\right) \int_I \mu(di) \int_0^1 \nu_i(du) u\mathcal{L}(E_i)^d < \infty
\end{aligned}$$

as required. Finiteness follows from ($\mathscr{H}3$). ■

3.3 Existence

Proposition 3.3.1 *The set of processes* $\{\rho^\alpha_\cdot\,;\,\alpha\in(0,1]\}$ *is tight in* $D_\Xi[0,\infty)$.

Proof: We look to use the Aldous-Rebolledo Criterion which is recalled as Theorem D.2. Our process ρ^α is Ξ valued rather than real valued and to remedy this we use Theorem D.1. Fix some sequence $(\alpha_m)_{m\in\mathbb{N}} \subseteq (0,1]$ such that $\alpha_m \downarrow 0$ as $m\to\infty$. and write $\rho^{\alpha_m}_t = \rho^m_t$ for the duration of this proof. Proposition 2.1.7 gives us that Ξ is separable and hence $D_\Xi[0,\infty)$ is also separable. Thus it suffices to check that $\{\rho^m\,;\,m\in\mathbb{N}\}$ is tight for the arbitrary sequence $\alpha_m\downarrow 0$.

Ξ is compact by Proposition 2.1.7 and thus ρ^m automatically satisfies the compact containment condition of Theorem D.1. Also by Proposition 2.1.7, $\mathscr{I}$ is dense in $C(\Xi)$ under the $||\cdot||_\infty$ topology, so it is certainly dense in the topology of uniform convergence on compact sets. Thus by Theorem D.1 and the fact that a linear combination of tight processes is necessarily tight, $\{\rho^m\,;\,m\in\mathbb{N}\}$ is tight as a $D_\Xi[0,\infty)$ valued process if and only if for every $n\in\mathbb{N}$ and Ψ of the form $(\star,n)$, $\{I_n(\rho^m_\cdot;\Psi)\,;\,m\in\mathbb{N}\}$ is tight as a sequence of $D_\mathbb{R}[0,\infty)$ valued processes.

Since Ξ is compact condition *(1)* of Theorem D.2 is automatic. We now check *(2)*. Note that

$$\begin{aligned}
&\mathbb{E}\Big[\,\big|I_n(\rho^m_{t+\theta};\Psi) - I_n(\rho^m_t;\Psi)\big|\,\Big] \\
&\quad = \int_{(\mathbb{R}^d)^n} dx_1\dots dx_n \psi(x_1,\dots,x_n)\mathbb{E}\left[\left\langle \bigtimes_{j=1}^n \xi_j, \bigotimes_{j=1}^n \rho^m_{t+\theta}(x_j)\right\rangle - \left\langle \bigtimes_{j=1}^n \xi_j, \bigotimes_{j=1}^n \rho^m_t(x_j)\right\rangle\right].
\end{aligned}$$

Let

$$A^m_n(\theta,x_1,\dots,x_n) = \{\rho^m_t(x_j) = \rho^m_{t+\theta}(x_j) \text{ for all } j=1,\dots,n\}.$$

be the event that each of the sites $\rho^\alpha_t(x_j)$ does not change type over the time interval $(t,t+\theta]$. By the dynamics of Definition 3.2.2, the rate at which any fixed site is hit by reproduction events is bounded above by $\int_I \mu^\alpha(di)\int_0^1 \nu_i(du)\, u\alpha^d \#(E_i\cap\alpha\mathbb{Z}^d)$. By Lemma 3.3.2 there is a constant C such that for all m

$$\mathbb{P}[A^m_n(\theta,x_1,\dots,x_n)] \leqslant \exp\left(-n\theta C\int_I \mu(di)\int_0^1 u\mathscr{D}(i)^d\right).$$

Note also that off the event $A^m_n(\theta,x_1,\dots,x_n) = \{\rho^m_t(x_j)$, for all j $\rho^m_{t+\theta}(x_j) = \rho^m_t(x_j)$ so as

$$\mathbb{E}\left[\left\langle \bigtimes_{j=1}^n \xi_j, \bigotimes_{j=1}^n \rho^m_{t+\theta}(x_j)\right\rangle - \left\langle \bigtimes_{j=1}^n \xi_j, \bigotimes_{j=1}^n \rho^m_t(x_j)\right\rangle \Big| A^m_n(\theta,x_1,\dots,x_n)^c\right] = 0.$$

Hence

$$\mathbb{E}\Big[\ \big|I_n(\rho^m_{t+\theta};\Psi) - I_n(\rho^m_t;\Psi)\big|\ \Big] \leqslant ||\psi||_1 \left(\prod_{j=1}^n ||\chi_j||_\infty\right) \exp\left(-n\theta C \int_I \mu(di) \int_0^1 u\mathscr{D}(i)^d\right)$$

which tends to 0 uniformly in m as $\theta \to 0$. *(2)* follows by Markov's inequality, completing the proof. ■

The following Lemma essentially states that, ignoring whether or not the selection test passes, the rate of reproduction events occurring in the pre-limiting processes converges to that of the SΛFV process.

Lemma 3.3.2 *As $\alpha \downarrow 0$,*

$$\int_{I_0^\alpha} \mu(di) \int_0^1 \nu_i(du)\, u\mathcal{L}(E_i) \to 0 \tag{3.3.1}$$

and for some constant $C \in (0,\infty)$, for all $\alpha \in (0,1]$, $\int_{I_0^\alpha} \mu(di) \int_0^1 \nu_i(du)\, u\mathcal{L}(E_i) \leqslant C$.

Proof: Note that by $(\mathscr{H}1)$

$$\mathbb{1}\{i \in I_0^\alpha\} u\mathcal{L}(E_i) \to 0$$

pointwise, as $\alpha \to 0$. Also note that $\mathcal{L}(E_i) \leqslant C_d \mathscr{D}(i)^d$ where C_d is the volume of a d dimensional unit ball and thus by $(\mathscr{H}2)$ we have

$$\int_I \mu(di) \int_0^1 \nu_i(du)\, \mathbb{1}\{i \in I_0^\alpha\} u\mathcal{L}(E_i) \leqslant \int_I \mu(di) \int_0^1 \nu_i(du) u C_d \mathscr{D}(i)^d < \infty$$

By the Dominated Convergence Theorem we have the result. ■

Let $\mathbb{Q}^\alpha \in \mathscr{P}(D_\Xi[0,\infty))$ be the law of $\rho^\alpha_\cdot$.

Proposition 3.3.3 *Any limit point of the set $\{\mathbb{Q}^\alpha\,;\, \alpha \in (0,1]\} \subseteq \mathscr{P}(D_\Xi[0,\infty))$ is the law of a Ξ valued process which solves the martingale problem for the Spatial Λ-Fleming-Viot process.*

Proof: Suppose $\mathbb{Q}$ is a limit point of $\{\mathbb{Q}^\alpha\,;\, \alpha \in (0,1]\}$. Then there exists a sequence $\alpha_m \to 0$ such that $Q^{\alpha_m} \to Q$. By Lemma 3.3.1 and Skorokhods theorem (3.1.8 in Ethier and Kurtz (1986)) there is a probability space $(\hat\Omega, \hat{\mathcal{F}}, \hat P)$ equipped with $D_\Xi[0,\infty)$ valued random variables $\hat\rho^m$ with law $\mathbb{Q}^{\alpha_m}$ and $\hat\rho$ with law $\mathbb{Q}$ such that $Q^{\alpha_m} \to Q$ in $\mathscr{P}(D_\Xi[0,\infty))$. By passing to a subsequence we may assume

$$\hat\rho^m \to \hat\rho \tag{3.3.2}$$

almost surely as $m \to \infty$.

By Proposition 3.2.7, for all $m \in \mathbb{N}$ and Ψ of the form $(\star, n)$

$$I_n(\hat\rho^m_t;\Psi) - \int_0^t G^{\alpha_m} I_n(\hat\rho^m_s;\Psi) ds \tag{3.3.3}$$

is a martingale. We aim to take a limit of this expression as $m \to \infty$. Proposition 2.1.7 gave us $I_n(\cdot;\Psi) \in C(\Xi)$ and thus by (3.3.2) we have

$$I_n(\hat\rho^m_t;\Psi) \to I_n(\hat\rho_t;\Psi) \tag{3.3.4}$$

almost surely. Note that this convergence is dominated by the constant $||\psi||_1 \prod_{j=1}^n ||\chi_j||_\infty$. By Lemma 3.2.5 we have $\rho^\alpha \in \mathscr{X}^\alpha$ and so we may apply Lemmas 3.2.8 and 3.2.9 to get

$$\begin{aligned}
&|G^{\alpha_m} I_n(\hat\rho_s;\Psi) - G I_n(\hat\rho_s;\Psi)| \\
&\quad\leqslant \int_{\mathbb{R}^d} dy \int_I \mu(di) \int_0^1 \nu_i(du) \int_{(\mathbb{R}^d)^n} dx_1 \dots dx_n \Bigg\{ \mathbb{1}\Big\{(x_j) \notin A_n^{\alpha_m}\Big\} \\
&\qquad u\left(2n|\psi(x_1,\dots,x_n)| \prod_{j=1}^n ||\chi_j||_\infty\right) \left(\mathbb{1}\Big\{\exists j, [x_j]_\alpha \in E_i([y]_\alpha)\Big\} \mathbb{1}\Big\{i \notin I_0^\alpha\Big\} + \mathbb{1}\Big\{\exists j, x_j \in E_i(y)\Big\}\right) \Bigg\} \\
&\quad + \left(2n||\varphi||_1 \prod_{j=1}^n ||\chi_j||_\infty\right) \int_{I_0^\alpha} \mu(di) \int_0^1 \nu_i(du)\, u\mathscr{D}(i)^d.
\end{aligned}$$

The term in the final line of the above expression tends to zero as $\alpha \downarrow 0$ by Lemma 3.3.2. We bound the other term of this expression (which makes up the second and third lines) as follows.

$$\begin{aligned}
\ldots &\leqslant C(n,\Psi)\int_{\mathbb{R}^d} dy \int_I \mu(di)\int_0^1 \nu_i(du)\int_{(\mathbb{R}^d)^n} dx_1\ldots dx_n \\
&\qquad\qquad u\mathbb{1}\Big\{(x_j)\notin A_n^{\alpha_m}\Big\}|\psi(x_1,\ldots,x_n)|\sum_{j=1}^n \mathbb{1}\Big\{|y-x_j|\leqslant 2\mathscr{D}(i)\Big\} \\
&= C(n,\Psi)\int_I \mu(di)\int_0^1 \nu_i(du)\, u\int_{(\mathbb{R}^d)^n} dx_1\ldots dx_n \\
&\qquad\qquad \mathbb{1}\Big\{(x_j)\notin A_n^{\alpha}\Big\}\mathbb{1}\Big\{(x_j)\notin A_n^{\alpha_m}\Big\}|\psi(x_1,\ldots,x_n)|\sum_{j=1}^n\int_{\mathbb{R}^d} dy\mathbb{1}\Big\{|x_j-y|\leqslant 2\mathscr{D}(i)\Big\} \\
&\leqslant C'(n,\Psi)\int_I \mu(di)\int_0^1 \nu_i(du)\, u\int_{(\mathbb{R}^d)^n} dx_1\ldots dx_n\mathbb{1}\Big\{(x_j)\notin A_n^{\alpha_m}\Big\}|\psi(x_1,\ldots,x_n)|\mathscr{D}(i)^d \\
&\leqslant C''(n,\Psi)\int_{(\mathbb{R}^d)^n} dx_1\ldots dx_n\mathbb{1}\Big\{(x_j)\notin A_n^{\alpha_m}\Big\}|\psi(x_1,\ldots,x_n)| \qquad (3.3.5)
\end{aligned}$$

where to get from the second to third lines we used that $i \in I\backslash I_0^\alpha$ implies $\alpha \leqslant \mathscr{D}(i)$ and to get from the fifth to the sixth lines we used ($\mathscr{H}3$). Since $||\psi||_1 < \infty$ and $\mathbb{1}\Big\{(x_j)\notin A_n^{\alpha_m}\Big\} \to 0$ as $m \to \infty$, by the Dominated Convergence Theorem we obtain

$$|G^{\alpha_m} I_n(\hat{\rho}_s;\Psi) - GI_n(\hat{\rho}_s;\Psi)| \to 0 \qquad (3.3.6)$$

as $m \to \infty$. By Lemma 3.3.2 and (3.3.5) this convergence is bounded by the some constant $C'''(\Psi,n)$. It follows from (3.3.2) and the dominated convergence taking place in (3.3.4) and (3.3.6) that if $L \in \mathbb{N}$ and for $l = 1,\ldots,L$, $b_l : \Xi \to \mathbb{R}$ is bounded and measurable and $0 < r_1,\ldots,r_L \leqslant s < t$,

$$\begin{aligned}
&\mathbb{E}\left[\left(I_n(\hat{\rho}_t;\Psi) - I_n(\hat{\rho}_s;\Psi) - \int_s^t G\, I_n(\hat{\rho}_a;\Psi)da\right)\prod_{l=1}^L b_l(\hat{\rho}^m_{r_l})\right] \\
&\qquad = \lim_{m\to\infty}\mathbb{E}\left[\left(I_n(\hat{\rho}_t;\Psi) - I_n(\hat{\rho}^m_s;\Psi) - \int_s^t G\, I_n(\hat{\rho}^m_a;\Psi)da\right)\prod_{l=1}^L b_l(\hat{\rho}^m_{r_l})\right].
\end{aligned}$$

Since (3.3.3) is a martingale the above line is equal to 0. Hence

$$I_n(\hat{\rho}_t;\Psi) - \int_0^t G^{\alpha_m} I_n(\hat{\rho}_s;\Psi)ds$$

is a martingale. Since a linear combination of martingales is a martingale it follows that

$$\phi(\rho_t) - \int_0^t G_s(\phi)(\rho_s)ds$$

is a martingale for all $\phi \in \mathscr{I}$.

It remains only to show that $\rho_0^\alpha \to \kappa$. Recall that the initial states ρ_0^α are given by $\rho_0^\alpha(x) = \kappa([x]_\alpha)$. According to Lemma 2.1.7 it suffices to show that for each $I_n(\cdot,\Phi)$ of the form $(\star,n)$, $I_n(\rho_0^\alpha,\Psi) \to I_n(\kappa,\Psi)$. By ($\mathscr{H}4$) we have

$$\mathbb{P}\left[||\rho_0^\alpha(x) - \kappa(x)||_{TV} \to 0 \text{ for almost all } x = (x_j)_{j=1}^d \in \mathbb{R}^d\right] = 1$$

(recall $||\cdot||_{TV}$ denotes the total variational norm on $\mathscr{P}(K)$). Hence,

$$\mathbb{P}\left[\left\langle \bigtimes_{j=n}\chi_j, \bigotimes_{j=1}^n \rho_0^\alpha(x_j)\right\rangle \to \left\langle \bigtimes_{j=n}\chi_j, \bigotimes_{j=1}^n \kappa(x_j)\right\rangle \text{ for almost all } x = (x_j)_{j=1}^d \in \mathbb{R}^d\right] = 1.$$

It now follows by dominated convergence that $I_n(\rho_0^\alpha,\Psi) \to I_n(\kappa,\Psi)$ almost surely. ■

PROOF: [Of Theorem 3.1.8] This now follows from Propositions 3.3.1 and 3.3.3. ■

3.4 Uniqueness

As we have already said, proof of uniqueness to the martingale problem of Definition 3.1.5 is work in progress.

PROOF: [Of Theorem 3.1.9] In Section 2.1.2 we constructed a dual for the basic version of the SΛFV process. Essentially the same method constructs a dual in the case with non-central parenting and since one of our assumptions was that there is no selection, this dual is then enough to prove uniqueness of solutions to the martingale problem. We do not give the details here since a full proof of uniqueness is hoped for at a later date. ■

Chapter 4

Super-Brownian motion is a scaling limit of bursting processes

In Section 4.1.2 we define a discrete model, which we call a bursting process, for spatial competition between different types of individuals. The process is essentially the Spatial Λ-Fleming-Viot process with typespace $K = \{0, 1\}$, uniform killing (that is, $\nu_i = \delta_{\{1\}}$) and discretized space. It can equally well be thought of as a generalization of the Voter model in which sites reproduce by imposing their type on some collection of the other sites (rather than just a single site). In this chapter we show that suitably rescaled, in dimension $d \geqslant 3$, bursting processes converge to super-Brownian motion.

4.1 Introduction

4.1.1 Some background

It is now well known that in dimensions $d \geqslant 2$ the Voter model can be rescaled to super-Brownian motion. This was proved initially in Cox et al. (2000) and was later generalised (to a stochastic Lotka-Volterra model) and refined in Cox and Perkins (2008) and Cox et al. (2010). Many other processes are also known to rescale to super-Brownian motion, for example see Cox and Klenke (2003), Durret and Perkins (1999) or van der Hofstad and Slade (2003). A common feature (which is almost a necessity) in processes which rescale to super-Brownian motion is the ability to rescale an ancestral lineage to a Brownian motion; thus ancestral lineages in the pre-limiting processes must have finite variance. For the SΛFV process, a case where they do not is explored in Berestycki et al. (2010) and gives rise to a quite different scaling limit.

All of Cox et al. (2000), Cox and Perkins (2008) and Cox et al. (2010) (and our own proof) operate via deriving a superprocess type martingale problem and taking a suitable scaling limit. In both Cox and Perkins (2008) and Cox et al. (2010) the main differences to the original paper Cox et al. (2000) are the extra drift terms. In our case the argument for this term is only cosmetically different to Cox et al. (2000) but the argument for convergence of the square bracket term (Sections 4.6 and 4.7) takes more effort.

We will work in dimensions $d \geqslant 3$, and combine the strategies of Cox et al. (2000) and Cox et al. (2010). Our argument is in spirit closest to that of Cox et al. (2000) but we will use the much improved method of Cox et al. (2010) to approach what in Cox et al. (2000) was the upgrading of L^1 estimates to L^2 ones (Section 4 of Cox et al. (2000)). We require some new ideas to obtain the asymptotic properties of the dual in Section 4.6. With these in hand we can approach the difficult part of the proof (the mean field simplification in Section 4.7) using a combination of techniques from Cox et al. (2000) and Cox et al. (2010).

This chapter presents a neater template for convergence to super-Brownian motion than can be found anywhere in the literature. We will give a heuristic explanation of the proof in Section 4.1.4, which includes precise details of where the extra complications lie.

Remark 4.1.1 *As for the Voter model, we expect that in dimension $d = 2$ a rescaling (with different particle mass to account for the clustering which occurs in two dimensions) that takes bursting processes*

to super-Brownian motion. Dimension $d = 1$ will not work for the same reasons as the $d = 1$ Voter model does not; see the introduction of Cox et al. (2000). The case $d = 2$ is more delicate than $d \geqslant 3$ but we do not expect the necessary modifications to the argument given here to be huge.

We will work on a lattice $\alpha\mathbb{Z}^d$ where $\alpha \in (0, 1]$ instead of just $\mathbb{Z}^d$. Our rescaling to super-Brownian motion will be indexed by N.

Remark 4.1.2 *We thus obtain some dependency on α of the limit obtained as $N \to \infty$. Some versions of the Spatial Λ-Fleming-Viot process are expected to rescale to super-Brownian motion and these versions can be constructed as a (measure valued) limit of bursting processes as $\alpha \to 0$. If in our limiting super-Brownian motion we let $\alpha \to 0$, and the parameters converge to something finite and non-zero, we produce strong evidence (but not a proof) that suitable versions of the Spatial Λ-Fleming-Viot process will rescale to super-Brownian motion.*

4.1.2 Definition of bursting processes

A bursting process is parametrized by

1. a mesh size $\alpha > 0$ for the lattice $\alpha\mathbb{Z}^d$,
2. a sequence $(\phi_n)_{n\in\mathbb{N}} \subseteq [0, \infty)$,
3. a sequence $(\Phi_n)_{n\in\mathbb{N}}$ of bounded subsets of $\alpha\mathbb{Z}^d$ such that $0 \notin \Phi_n$.

and the resulting process ξ we will call a (α, ϕ, Φ)-bursting process. At each time $t \in [0, \infty)$ we assign a random type $\xi_t(x) \in \{0, 1\}$ to each site $x \in \alpha\mathbb{Z}^d$. The corresponding measure valued process is

$$X_t = \alpha^d \sum_{x\in\alpha\mathbb{Z}^d} \xi_t(x)\delta_x \tag{4.1.1}$$

and we denote integration of φ against this measure by $X_t(\varphi)$. The factor α^d is the volume of $x + [-\alpha/2, \alpha/2)^d$, which is the section of $\mathbb{R}^d$ to which we notionally assign the type of $\xi_t(x)$. Set

$$\Phi_n(x) = x + \Phi_n(x)$$

and define the dynamics as follows. Let Λ be a Poisson point process with points (t, x, n) in $[0, \infty)\times\alpha\mathbb{Z}^d\times\mathbb{N}$ of rate

$$dt \otimes \alpha^d dx \otimes d\phi(n). \tag{4.1.2}$$

Here dt corresponds to Lebesgue measure on $[0, \infty)$, dx corresponds to the measure giving unit mass to each point of αZ^d and $d\phi(n)$ to the measure on $\mathbb{N}$ where $\{n\}$ has weight $\phi(n)$. Then

- if (t, x, n) is a point of Λ then at time t the sites $x + \Phi_n$ adopt the type $\xi_{t-}(x)$.
- in between the times $\{t\,;\,(t, x, n) \in \Lambda, y \in \Phi_n(x)\}$, $\xi_\cdot(y)$ does not change.

This is known as a reproduction event (of Φ_n about x at t).

We require the parameters of the bursting process to satisfy the following.

($\mathscr{C}1$)
- $\sum_n \phi_n < \infty$
- There exists $L \in (0, \infty)$ such that for all $n \in \mathbb{N}$, $\Phi_n \subseteq B_L(0)$.

($\mathscr{C}2$) For all $x, y \in \alpha\mathbb{Z}^d$,

$$\sum_n \mathbb{1}\{x \in \Phi_n(y)\}\phi_n = \sum_n \mathbb{1}\{y \in \Phi_n(x)\}\phi_n.$$

($\mathscr{C}3$) There exists $\sigma > 0$ such that

$$\left(\frac{1}{\sum_n \#(\Phi_n)\phi_n}\right) \sum_n \phi_n \sum_{x\in\alpha\mathbb{Z}^d} x_i x_j 1(x \in \Phi_n) = \delta_{ij}\sigma^2$$

where $x = (x_k)_1^d \in \mathbb{R}^d$, $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

Two immediate consequences of ($\mathscr{C}1$) are that $\sum_n \#(\Phi_n)\phi_n < \infty$ and $\sum_n \#(\Phi_n)^2\phi_n < \infty$. When we wish to use these facts we will simply refer to ($\mathscr{C}1$).

Remark 4.1.3 *We also assume that $\sum_n \#(\Phi_n)\phi_n > 0$ (if it is zero the bursting process is constant and rescales to the constant process of super-Brownian motion with zero diffusion and zero branching).*

Remark 4.1.4 *In fact the condition $\sum_n \#(\Phi_n)^2\phi_n < \infty$ is all that is necessary for the process to exist. This can be checked using the construction of infinite particle systems given in Ligget (1985).*

As we will see in Section 4.4, bursting processes exhibit the same sort of ancestral lineage duality as the Voter model does. In the Voter model case this originates in Harris (1975) and is usually known as the Harris decomposition. It can be found, for example, in Durret (1995). ($\mathscr{C}2$) and ($\mathscr{C}3$) are best seen as statements about the behaviour of the ancestral lineages.

Let B_t be a random walk with the same distribution as an ancestral lineage of ξ. The ancestral lineages of a bursting process are discussed in Section 4.3 but by now the reader should be familiar with the concept. Then there is some constant R and a $\alpha\mathbb{Z}^d$ valued random variable W such that B_t has exponential holding times with parameter R and at its jumps B_t increments its position by successive independent copies of W. In Lemma 4.3.3 we show that ($\mathscr{C}2$) means that W and $-W$ have the same distribution, whilst ($\mathscr{C}3$) means the axial components of W are uncorrelated $\alpha\mathbb{Z}$ valued random variables with common variance. (In other words, the covariance matrix of an ancestral lineage is a multiple of the identity matrix.)

However, our ancestral lineages will not be as well behaved as those of the Voter model. In the Voter model two lineages move independently before they coalesce but this is not so for general bursting processes. This is easily seen since if we have two individuals at, say, 0 and z then there could be an individual nearby capable of having a reproduction event killing the pair at 0 and z in the same instant.

4.1.3 Characterization of super-Brownian motion

We use the following martingale characterization of super-Brownian motion. Let $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ denote the space of càdlàg paths indexed by $[0,\infty)$ on $\mathcal{M}_F(\mathbb{R}^d)$ with the Skorokhod topology. Let $C_0^\infty(\mathbb{R}^d)$ denote the set of functions $\varphi : \mathbb{R}^d \to \mathbb{R}$ which have compact support and continuous partial derivatives of all orders.

Definition 4.1.5 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{Q})$ be a complete filtered space. An adapted $\mathbb{Q}$-a.s. continuous $\mathcal{M}_F(\mathbb{R}^d)$-valued process (X_t) is a super-Brownian motion with initial measure $X_0 \in \mathcal{M}_F(\mathbb{R}^d)$, branching rate $b \in (0,\infty)$ and diffusion rate $a^2 \in (0,\infty)$ if for all $\varphi \in C_0^\infty(\mathbb{R}^d)$ the process*

$$M_t(\varphi) = X_t(\varphi) - X_0(\varphi) - \int_0^t X_s\left(\frac{a^2}{2}\Delta(\varphi)\right) ds$$

is a continuous $\mathcal{F}_t$-martingale and

$$\langle M(\varphi)\rangle_t = \int_0^t X_s(b\varphi^2) ds.$$

Existence and uniqueness of the solution to this martingale problem are well known. A proof of uniqueness with test functions $\varphi \in C_0^\infty(\mathbb{R}^d)$ can be found in the appendix of Cox et al. (2000).

4.1.4 A heuristic explanation of the proof

A bursting process satisfying conditions ($\mathscr{C}1$)-($\mathscr{C}3$) (and ($\mathscr{C}4$) and ($\mathscr{C}5$) below on the initial states) can be rescaled to super-Brownian motion. We give a precise statement of the results in the next Section, but for now let us give a heuristic explanation of the steps of our proof.

Let us begin by examining the martingale problem of Definition 4.1.5. It is of the form

$$M_t(\varphi) = Z_t(\varphi) - Z_0(\varphi) - \int_0^t Z_s \overbrace{(\ldots)}^{\mathscr{A}} ds, \qquad \langle M(\varphi)\rangle_t = \int_0^t Z_s \overbrace{(\ldots)}^{\mathscr{B}} ds.$$

The way to interpret this martingale problems is as follows. The term labelled $\mathscr{A}$ carries the information describing the behaviour of a single ancestral lineage. In Definition 4.1.5 this is (morally, at least, since the process is measure valued) a Brownian motion with generator $\frac{a}{2}\Delta$.

The term labelled $\mathscr{B}$ relates to the potency of a particle, that is if we notionally pick one of the particles making up the support of the super-Brownian motion, the term labelled $\mathscr{B}$ tells us the rate at which this particle spreads its own genetic type. To be precise, the idea is that a particle 'infects' a site if it reproduces in such a way as changes the type at the site of its child. The potency of a particle is its ability to infect the rest of the system. The test function φ weights the spatial locations according to which sites we want to see the potency of. In Definition 4.1.5 the term $\mathscr{B}$ is a constant multiple of the test function, which corresponds to saying each particle reproduces independently of all other particles.

Remark 4.1.6 *Note that the two properties of Brownian motion we claim $\mathscr{A}$ and $\mathscr{B}$ represent are properties one would expect from the construction of super-Brownian motion as a limit of branching Brownian motions, see chapter 1 of Etheridge (2000).*

The fact that the martingale problem of Definition 4.1.5 has a unique solution says these two characteristics identify super-Brownian motion uniquely amongst càdlàg $\mathcal{M}_F(\mathbb{R}^d)$ valued processes. Thus in order to prove convergence of some process with ancestral lineages to super-Brownian motion one must expect to show that a single lineage rescales to Brownian motion and that in the limit the potency of a particle collapses to being constant. This is precisely what we will show.

We will derive the terms of the corresponding martingale problem for bursting processes in Theorem 4.4.3. The term in position $\mathscr{A}$ will correspond to a finite variance, finite rate random walk (with uncorrelated axial components thanks to ($\mathscr{C}3$)) and by our choice of rescaling it is a simple matter to see that a single lineages goes to a Brownian motion. We see this in Lemma 4.4.4. The term in position $\mathscr{B}$ will require substantially more effort. The convergence result corresponding to the term $\mathscr{B}$ is Lemma 4.5.2 and we will require all of Sections 4.6 and 4.7 to prove it.

Understanding the potency of the particles is essentially the same job as understanding the interaction between ancestral lineages. If, as in Definition 4.1.5 each particle reproduces independently then the ancestral lineages must move independently up until they occupy the same point in space, at which time they coalesce. In the Voter model case in order to understand the interaction between ancestral lineages we need only know about the behaviour of pairs of lineages (since multiple coalescence events do not occur). In a bursting process multiple coalescence events can occur and we need to work with triplets of lineages and a non-local coalescence mechanism. This is essentially the cause of the extra complications by comparison to Cox et al. (2000).

One result of these extra complications is that the constant which comes out in the square bracket term splits into two parts, γ_1 and γ_2 (see Theorem 4.2.3). The quantity γ_1 is about triplets of particles where one pair has a common ancestor but the other does not share it, whereas γ_2 is about triplets of particles with no common ancestor. They correspond to fixing one of the particles and looking at the rate at which it infects *pairs* of other particles. In γ_1 we count up the rates where it infects only one half of the pair, and in γ_2 we count up the rates where it infects neither. This comprises all the information available since taking γ_1 and γ_2 away from the total infection rate would give the rate of infecting both.

We refer to the result that the potency of a particle (to be precise, all particles simultaneously) converges to a constant (the same constant) as the *mean field simplification*. The argument for this proceeds as follows.

We will denote our N^{th} rescaled bursting process by X_t^N. Let ϵ_N^* be some sequence such that $\epsilon_N^* \to 0$ but $N\epsilon_N^* \to \infty$. On the time scale of the N^{th} process, as $N \to \infty$ an interval of length ϵ_N^* collapses into being (literally) no time at all. Thinking of our N^{th} stage rescaling as divided into infinitely many time intervals of length ϵ_N^*, we want the limiting behaviour of our N^{th} bursting processes over a time interval of length ϵ_N^* to look like the infinitesimal behaviour of super-Brownian motion. Inside each interval of length ϵ_N^* this corresponds to looking at the unscaled process over time $N\epsilon_N^* \to \infty$. So we want the unscaled process, run for all time, and then collapsed into a point to look like the infinitesimal local dynamics of super-Brownian motion. If we have got our mass rescaling correct (the space and time rescaling are fixed by $\mathscr{A}$) then what we see is that just enough mass survives long enough in the N^{th} rescaled process to appear in the limit, producing the intensity at which a (notional) single particle in the limiting super-Brownian motion reproduces. At this point the reader might like to examine the limits on the integrals in (4.7.3) which we (eventually) use to prove Lemma 4.5.2, where this idea comes into light.

In fact, as we see in (part 4 of the proof of) Lemma 4.7.5 it important that $\epsilon_N^* \to 0$ at a speed within the right scaling window to allows the bursting processes to even out sufficiently. It is the case that $X^N_{\epsilon_N^*} \to X_0$, and if we let $N\epsilon_N^* \to \infty$ too fast this does not occur and we do not give the pre-limiting processes enough time to even out and start looking like the infinitesimal local dynamics of super-Brownian motion.

Although our argument is not about $d = 2$, if one were to prove convergence of bursting processes to super-Brownian motion for dimension $d = 2$, one requires a slower choice of ϵ_N^* (and a different mass rescaling) to deal with the recurrent behaviour of the ancestral lineages. The care with which ϵ_N^* needs to be chosen corresponds loosely to the difficulty of getting estimates on asymptotic behaviour of the dual process. For us $\epsilon_N^* = N^{-1/4}$ will suffice. In Cox et al. (2010), for example, they use the exotic $(\log N)^{-19}$.

4.2 Results

In this section give a precise statement of the main results in this chapter.

4.2.1 The appropriate rescaling.

We will be simultaneously rescaling the lattice mesh, the particle mass and the speed of the process. Our rescaled bursting processes $\xi_t^N(\cdot)$, indexed by $N \in \mathbb{N}$ will be on the lattices

$$S_N = \frac{\alpha}{\sqrt{N}}\mathbb{Z}^d.$$

Our mass rescaling will be a factor $1/N$ and we will run time faster by a factor N. Thus our rescaled measure valued processes are

$$X_t^N = \frac{\alpha^d}{N}\sum_{x\in\alpha\mathbb{Z}^d}\xi_{Nt}\left(\frac{x}{\sqrt{N}}\right)\delta_x.$$

The dependence on α and d is permitted to disappear into the background. We stress that our rescaling to super-Brownian motion is indexed by N and α is kept constant throughout.

4.2.2 Initial conditions

Recall $\mathcal{M}_F(\mathbb{R}^d)$ denotes the space of finite measures on $\mathbb{R}^d$. We need conditions on the initial states, ξ_0^N.

($\mathscr{C}4$) For each N, $\sum_x \xi_0^N(x) < \infty$.

($\mathscr{C}5$) There exists some $X_0 \in \mathcal{M}_F(\mathbb{R}^d)$ such that $X_0^N \to X_0$ in $\mathcal{M}_F(\mathbb{R}^d)$.

Condition ($\mathscr{C}4$) corresponds to each initial state having only finitely many 1s, whilst ($\mathscr{C}5$) is obviously going to be necessary in order to obtain a scaling limit.

Remark 4.2.1 *Note that we do not take some set $A \subseteq \mathbb{R}^d$ and set $\xi_0^N = S_N \cap A$. The total mass of ξ_0^N under this definition would not remain finite under the rescaling (in dimensions $d \geqslant 3$). The combined effect of ($\mathscr{C}4$) and ($\mathscr{C}5$) is that the initial state X_0 of the limit process will have 2 dimensional support.*

For the SΛFV process, the case where one does not impose sparse initial conditions is covered in Berestycki et al. (2010).

Lemma 4.2.2 *($\mathscr{C}4$) and ($\mathscr{C}5$) imply $\sup_N X_0^N(1) < \infty$.*

From this point on we proceed from a bursting process satisfying ($\mathscr{C}1$)-($\mathscr{C}5$).

4.2.3 Results

In order to define the parameters of the limiting super-Brownian motion we need to use the ancestral lineages. The reader unfamilir with this type of duality might like to read Section 2 where a full explanation of what it means to be an ancestral lineage is given.

Let (B^x, B^y) be a pair of ancestral lineages of ξ, tracing back the succession of ancestors of the types at $x, y \in \alpha\mathbb{Z}^d$ from some time t_0. The distribution of this pair of random walks is independent of the time at which we start tracing back from so we can think of these walks as run for all time. Define

$$\tau(x,y) = \inf\{t \in [0,\infty)\,;\, B^x_t = B^y_t\}$$

to be the first time at which these walks meet (that is, when the types at x and y has their most recent common ancestor). Since $d \geqslant 3$ there is the possibility that they never meet and $\tau(x,y) = \infty$. In fact, since our two walkers are dependent it takes a small amount of work to show $\mathbb{P}[\tau(x,y) = \infty] > 0$ for $x \neq y$ from the standard result about random walks in $d \geqslant 3$.

Let $\mathbb{Q}^{2b,a^2}$ be the law of the super-Brownian motion of Definition 4.1.5 with initial state X_0. Let $\mathbb{Q}_N$ denote the law of $(X^N_t)_{t=0}^{\infty}$, which is also a probability measure on $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$. The main theorem of this chapter is as follows.

Theorem 4.2.3 *In dimensions $d \geqslant 3$, as $N \to \infty$, $\mathbb{Q}_N$ converges weakly to $\mathbb{Q}^{Q(\gamma_1+\gamma_2),R\sigma^2}$, in the space of probability measures on $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$. Here*

$$R = \alpha^d \sum_n \#(\Phi_n)\phi_n, \qquad Q = \alpha^{2d}\sum_n \#(\Phi_n)^2\phi_n$$

and $\gamma_1, \gamma_2 \in (0,\infty)$ are given by

$$\gamma_1 = \sum_{e,f\in\alpha\mathbb{Z}^d}\left(\sum_n \mathbb{1}\{0, e \in \Phi_n(f)\}\,\phi_n\right)\mathbb{P}[\tau(0,e) < \infty, \tau(0,f) = \infty]$$

$$\gamma_2 = \sum_{e,f\in\alpha\mathbb{Z}^d}\left(\sum_n \mathbb{1}\{e, f \in \Phi_n\}\,\phi_n\right)\mathbb{P}[\tau(0,e) = \infty, \tau(0,f) = \infty]$$

The quantities γ_1 and γ_2 should not be thought of as a pair which mirror each other. They correspond to genuinely different parts of the evolution of the pre-limiting processes, which we have already discussed in the course of Section 4.1.4.

Remark 4.2.4 *To recover the corresponding result in Cox et al. (2000) set $\alpha = 1$, choose the Φ_n to be one point sets and choose ϕ_n so as $\sum_n \phi_n = 1$. Then $\sum_n \mathbb{1}\{f \in \Phi_n\}\phi_n$ is the rate at which the site 0 reproduces to the site f, and*

$$\gamma_1 = \gamma_2 = \sum_{f\in\alpha\mathbb{Z}^d}\left(\sum_{n\in\mathbb{N}}\mathbb{1}\{f \in \Phi_n\}\phi_n\right)\mathbb{P}[\tau(0,f) = \infty].$$

4.3 The ancestral lineages

We set up the random walk which (with appropriate coupling) will turn out to be the path followed backwards in time by an ancestral lineage of ξ^N.

Define the rescaled sets

$$\Phi^N_n(x) = x + \frac{1}{\sqrt{N}}\Phi_n \subseteq S_N$$

and write also $\Phi^N_n = \Phi^N_n(0)$. Let W be an αZ^d valued random variable with law p given by

$$p(x) = \frac{\alpha^d}{R}\sum_n \mathbb{1}\{x \in \Phi_n\}\phi_n$$

where $R = \alpha^d \sum_n \#(\Phi_n)\phi_n$. Note that R is finite by ($\mathscr{C}1$) and $p(0) = 0$. Let $B^{x,N}$ be a random walk on S_N starting from $x \in S_N$, jumping at rate $R_N = NR$ and incrementing its position by successive independent copies of $W_N = \frac{1}{\sqrt{N}}W$ at each jump. Denote the distribution of W_N by

$$p_N(x) = p(x\sqrt{n}) = \frac{\sum_n \mathbb{1}\{x \in \Phi_n^N\}\phi_n}{\sum_n \#(\Phi_n^N)\phi_n},\ x \in S_N$$

and define $p_N(x, y) = p_N(y - x)$.

Lemma 4.3.1 *Let $x \in S_N$ and $t > 0$. If we trace the source of the type $\xi_t^N(x)$ backwards in time from time t until time 0, we follow a random walk with distribution $(B_s^{x,N})_{s=0}^t$ (where $s = 0$ corresponds to time t and $s = t$ to time 0).*

Remark 4.3.2 *Recall our convention that we always deal with the right continuous version of the ancestral lineages.*

Proof: The total rate of reproduction events hitting $y \in S_N$ is

$$\begin{aligned}\sum_z \sum_n \mathbb{1}\{y \in \Phi_n^N(z)\}\alpha^d N\phi_n &= \alpha^d N \sum_n \sum_z \mathbb{1}\{-z \in \Phi_N^n(-y)\}\phi_n \\ &= \alpha^d N \sum_n \#(\varphi_n)\phi_n = R_N\end{aligned}$$

where we translate by $-y - z$ in the first line. The total rate of the subset of such reproduction events which would hit y from a fixed z is $\alpha^d N \sum_n \mathbb{1}\{y \in \Phi_n^N(z)\}\phi_n$ and thus the probability of y inheriting the type of x in this fashion is

$$\frac{\sum_n \mathbb{1}\{y \in \Phi_n^N(z)\}\alpha^d N\phi_n}{\sum_z \sum_n \mathbb{1}\{y \in \Phi_n^N(z)\}\alpha^d N\phi_n} = p_N(z, y)$$

as required. ■

Specifically, by looking for the source of $\xi_t^N(x)$, we look back in time until x was last hit by a bursting event, and then move to the point y from which this burst originated (regardless of whether or not the type at x was actually changed by this event). We continue moving in this manner until (parameterizing time backwards) we reach time 0.

The following lemma explains the purpose of our conditions ($\mathscr{C}2$) and ($\mathscr{C}3$). For convenience we write $p = p_1$, $W = W_1$, etc.

Lemma 4.3.3 *The following hold:*

1. *$p_N(x) = p_N(-x)$ (and therefore $p_N(x, y) = p_N(y, x)$).*
2. *$\mathbb{E}[W_N] = 0$ and $|W|$ has moments of all orders.*
3. *Writing $W = (W(i))_{i=1}^d$, $\mathbb{E}[W(i)W(j)] = 0$ if $i \neq j$ and $\mathbb{E}\left[W(i)^2\right] = \sigma^2$.*

Proof: ($\mathscr{C}1$) is required for p to be defined. ($\mathscr{C}2$) says precisely that $p(x) = p(-x)$ which gives the first statement and implies $\mathbb{E}[W] = 0$. $|W|$ has moments of all orders because it is bounded (by L) as a consequence of ($\mathscr{C}1$). ($\mathscr{C}3$) is precisely the final statement. ■

The generator of the random walk $B^{x,N}$ is given by

$$\mathcal{A}^N\varphi(x) = R_N \sum_y p_N(x, y)(\varphi(y) - \varphi(x))$$

and the transition semigroup is given by

$$\mathcal{P}_t^N f(x) = \mathbb{E}\left[f(B_t^{x,N})\right].$$

We will usually take some $\varphi : \mathbb{R}^d \to \mathbb{R}$ as a test function, although strictly speaking in order to work with the generator one should define $\mathcal{A}^N\varphi$ for $\varphi : S_N \to \mathbb{R}$.

We will frequently need to study the interaction of two ancestral lineages and to do so we define

$$q(y,x,z) = \frac{\alpha^{2d}}{Q}\sum_n \mathbb{1}\{x,z \in \Phi_n(y)\}\phi_n.$$

for $y,x,z \in \alpha\mathbb{Z}^d$ where $Q = \alpha^{2d}\sum_n \#(\Phi_n)^2\phi_n$. The rescaled versions are $Q_N = NQ$ and for $(x,y,z) \in (S_N)^3$,

$$q_N(y,x,z) = \frac{\alpha^{2d}}{Q}\sum_n \mathbb{1}\{x,z \in \Phi_n^N(y)\}\phi_n.$$

We write $q = q_1$. The function $(x,z) \mapsto q(0,x,z)$ defines the law of a $(\alpha\mathbb{Z}^d)^2$ random variable. For now we have set up enough notation to proceed.

4.4 Decomposition of bursting processes

We use a stochastic integral definition of bursting processes which leads naturally to a martingale decomposition. In this respect the approach originates from Mueller and Tribe (1995). We set up the N-stage rescaling of the Poisson point processes as follows. Let

$$\{\Lambda(x,n)\,;\, x \in \alpha\mathbb{Z}^d, n \in \mathbb{N}\}$$

be a family of independent Poisson processes (defined on a common complete probability space) with $\Lambda(x,n)$ having rate $\alpha^d\phi_n$. Define

$$\Lambda_t^N(x,n) = \Lambda_{Nt}(x/\sqrt{N},n).$$

Thus for $x \in S_N$ and $n \in \mathbb{N}$, $\Lambda^N(x,n)$ is a Poisson process of rate $\alpha^d N\phi_n$. This induces the natural dependence between our N-stage rescaled processes. Define

$$\hat{\Lambda}^N(x,n)_t = \Lambda(x,n)_t - \alpha^d N\phi_n t$$

and note this is a martingale.

The bursting process ξ^N will turn out to be the unique solution to the system of equations

$$\zeta_t^N(x) = \zeta_0^N(x) + \sum_y\sum_n \int_0^t \left[\zeta_{s-}^N(y) - \zeta_{s-}^N(x)\right]\mathbb{1}\{x \in \Phi_n^N(y)\}d\Lambda^N(y,n) \tag{4.4.1}$$

where the first sum is over $y \in S_N$ and the second over $n \in \mathbb{N}$. It is easy to see that the equations of (4.4.1) correspond to a process with the dynamics discussed in Section 4.1.2.

Definition 4.4.1 *A solution ζ^N of (4.4.1) is a càdlàg $\{0,1\}^{S_N}$ valued process for which*

$$\forall x, \forall t > 0, \sum_y\sum_n \int_0^t \left|\zeta_{s-}^N(y) - \zeta_{s-}^N(x)\right| \mathbb{1}\{x \in \Phi_n^N(y)\}d\Lambda(y,n) < \infty$$

almost surely, with the intial condition ζ_0^N satisfying ($\mathscr{C}4$)-($\mathscr{C}5$).

Lemma 4.4.2 (4.4.1) *has the $(\alpha/\sqrt{N}, \Phi_n^N, \phi_n^N)$ bursting process ξ^N as its unique solution over $t \geqslant 0$. For each $T < \infty$*

$$\mathbb{E}\left[\sup_{t\leqslant T}\sum_x \xi_t^N(x)\right] < \infty$$

which, in particular, implies that $X_t^N(\varphi)$ is a.s. finite if φ is bounded.

PROOF: This is essentially the same as the proof of Lemma 2.1 in Cox et al. (2000). ■

As in Section 4.2.1 the corresponding rescaled measure valued processes are $X_t^N = \frac{\alpha^d}{N}\sum_x \xi_t^N(x)\delta_x$ and the integrals of functions against these measures are

$$X_t^N(\varphi) = \frac{\alpha^d}{N}\sum_x \xi_t^N(x)\varphi(x)$$

where $\varphi : S_N \to \mathbb{R}$ or $\varphi : \mathbb{R}^d \to \mathbb{R}$. We use the symbol 1 for the constant function $1(x) = 1$.

We write ξ_t^N for the set of sites of type 1 at time t (and as we already are, $\xi_t^N(x)$ for the value of the site x at time t). Let

$$\hat{\xi}_t^N(x) = 1 - \xi_t^N(x)$$

be the process ξ^N with the roles of the types 0 and 1 exchanged.

Let

$$m_s^{N,1}(\varphi) = \frac{\alpha^d}{N^2}\sum_{x,y,z} \varphi(x)\varphi(z)Q_N q_N(y,x,z)\xi_s^N(x)\xi_s^N(z)\hat{\xi}_s^N(y)$$

$$m_s^{N,2}(\varphi) = \frac{\alpha^d}{N^2}\sum_{x,y,z} \varphi(x)\varphi(z)Q_N q_N(y,x,z)\hat{\xi}_s^N(x)\hat{\xi}_s^N(z)\xi_s^N(y).$$

The Doob decomposition of $X_t^N(\varphi)$ is as follows.

Lemma 4.4.3 *Let $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ be bounded and measurable. Then*

$$X_t^N(\varphi) = X_0^N(\varphi) + M_t^N(\varphi) + \int_0^t X_s^N\left(\mathcal{A}^N\varphi\right) ds \tag{4.4.2}$$

where M^N is a càdlàg square integrable martingale given by

$$M_t^N(\varphi) = \sum_{y,n}\int_0^t \frac{\alpha^d}{N}\sum_x \varphi(x)\left[\xi_{s-}^N(y) - \xi_{s-}^N(x)\right]\mathbb{1}\{x \in \Phi_n^N(y)\}d\hat{\Lambda}_s^N(y,n)$$

and predictable square function given by

$$\langle M^N(\varphi)\rangle_s = \int_0^t \left(m_s^{N,1}(\varphi) + m_s^{N,2}(\varphi)\right) ds. \tag{4.4.3}$$

PROOF: Recall (4.4.1):

$$\xi_t^N(x) = \xi_0^N(x)\varphi(x) + \sum_y\sum_n \int_0^t \varphi(x)\left[\xi_{s-}^N(y) - \xi_{s-}^N(x)\right]\mathbb{1}\{x \in \Phi_n^N(y)\}d\Lambda_s^N(y,n).$$

We may split up $d\Lambda_s^N(y,n) = d\hat{\Lambda}_s^N(y,n) + \alpha^d N\phi_n ds$ and sum over $x \in S_N$ to get

$$X_t^N(\varphi_t) = X_0^N(\varphi_0) + M_t^N(\varphi) + \int_0^t L_s^N(\varphi)ds$$

where

$$L_s^N(\varphi_s) = \frac{\alpha^d}{N}\sum_{x,y,n} \varphi(x)\left[\xi_s^N(y) - \xi_s^N(x)\right]\mathbb{1}\{x \in \Phi_n^N(y)\}\alpha^d N\phi_n$$

Note that the left limits $s-$ have been dropped in the definition of L, since the integrator ds is continuous. The proof that $M^N(\varphi)$ is indeed a square integrable càdlàg martingale follows similarly to the proof of

Lemma 2.3 in Cox et al. (2000). By the same token we obtain that

$$\begin{aligned}\langle M^N(\varphi)\rangle_t &= \sum_{y,n}\int_0^t\left(\frac{\alpha^d}{N}\sum_x\varphi(x)\left[\xi^N_{s-}(y)-\xi^N_{s-}(x)\right]\mathbb{1}\{x\in\Phi^N_n(y)\}\right)^2\alpha^dN\phi_nds\\ &= \sum_{y,n}\int_0^t\frac{\alpha^{2d}}{N^2}\sum_{x,z}\varphi(x)\varphi(z)\left[\xi^N_s(y)-\xi^N_s(x)\right]\left[\xi^N_s(y)-\xi^N_s(z)\right]\mathbb{1}\{x,z\in\Phi^N_n(y)\}\alpha^dN\phi_nds\\ &= \int_0^t\frac{\alpha^d}{N^2}\sum_{y,x,z}\varphi(x)\varphi(z)\left[\xi^N_s(y)-\xi^N_s(x)\right]\left[\xi^N_s(y)-\xi^N_s(z)\right]Q_Nq_N(y,x,z)ds\end{aligned}$$

Noting that

$$\left[\xi^N_s(y)-\xi^N_s(x)\right]\left[\xi^N_s(y)-\xi^N_s(z)\right]=\xi^N_s(x)\xi^N_s(z)\hat{\xi}^N_s(y)+\hat{\xi}^N_s(x)\hat{\xi}^N_s(z)\hat{\xi}^N_s(y)$$

we obtain the form of (4.4.3). It remains to rearrange $L^N(\varphi)$. Note that

$$L^N_s(\varphi_s)=\frac{\alpha^d}{N}\sum_{x,y}\varphi(x)\left[\xi^N_s(y)-\xi^N_s(x)\right]R_Np_N(x,y)$$

and since $p_N(x,y)=p_N(y,x)$,

$$\begin{aligned}L^N_s(\varphi_s)&=\frac{\alpha^d}{N}\sum_{x,y}\xi^N_s(x)\left[\varphi(y)-\varphi(x)\right]R_Np_N(x,y)\\ &=\frac{\alpha^d}{N}\sum_x\xi^N_s(x)\sum_y\left[\varphi(y)-\varphi(x)\right]R_Np_N(x,y)\\ &=X^N_s(\mathcal{A}^N\varphi)\end{aligned}$$

which gives the form claimed for the final term of (4.4.2). ■

We now show that a single ancestral lineage rescales to a Brownian motion. This takes care of the convergence of the diffusion term in the decomposition of Lemma 4.4.2.

Lemma 4.4.4 *Let $\varphi:\mathbb{R}^d\to\mathbb{R}$ have bounded continuous third order derivatives. Then*

$$\lim_{N\to\infty}\left|\left|\mathcal{A}^N\varphi-\frac{R\sigma^2}{2}\Delta(\varphi)\right|\right|_\infty=0.$$

PROOF: Let us write $x=(x^i)_{i=1}^d\in\mathbb{R}^d$, and denote partial differentiation in the direction of the i^{th} coordinate as $\frac{\partial\varphi}{\partial i}=\varphi_i$. We write the Euclidean norm as $|\cdot|$. By Taylor's Theorem there is a random $Y_N\in[x,x+W_N]$ (the line segment between x and $x+W_N$) such that

$$\varphi(x+W_N)-\varphi(x)=\sum_{i=1}^d\varphi_i(x)W^i_N+\frac{1}{2}\sum_{i,j=1}^d\varphi_{ij}(Y_N)W^i_NW^j_N.$$

Hence, recalling that Lemma 4.3.3 gives $\mathbb{E}(W_N)=0$,

$$\begin{aligned}\left|\mathcal{A}^N\varphi-\frac{R\sigma^2}{2}\Delta(\varphi)\right|&=\left|R_N\mathbb{E}\left[\varphi(Y_N)-\varphi(x)-\frac{1}{2N}\sigma^2\Delta\varphi(x)\right]\right|\\ &\leqslant\frac{R}{2}\sum_{i,j=1}^d\mathbb{E}\left|\big(\varphi_{ij}(Y_N)-\varphi_{ij}(x)\big)NW^i_NW^j_N\right|\\ &\quad+\frac{R_N}{2}\sum_{i,j=1}^d\left|\varphi_{ij}(x)\mathbb{E}\left[W^i_NW^j_N-\frac{\delta_{ij}\sigma^2}{N}\right]\right|.\end{aligned}\tag{4.4.4}$$

By Lemma 4.3.3 $\mathbb{E}[W_N^i W_N^j] = \delta_{ij}\frac{\sigma^2}{\sqrt{N}}$ so the second term tends to zero. Since φ is Lipschitz we have $|\varphi(Y_n) - \varphi(x) \leqslant ||\varphi||_{Lip}|x - Y_N|$ and note also $|x - Y_n| \leqslant |W_N| = |W|/\sqrt{N}$. Thus the first term of (4.4.4) is bounded above by

$$\frac{R}{2}\sum_{i,j=1}^{d}\mathbb{E}\left[||\varphi||_{Lip}\frac{|W|}{\sqrt{N}}W^iW^j\right].$$

Since ($\mathscr{C}1$) implies W has moments of all order, this tends to zero as $N \to \infty$. ■

The following Lemma is our first use of duality.

Lemma 4.4.5 *If $\varphi : \mathbb{R}^d \to \mathbb{R}$ is bounded,*

$$\mathbb{E}\left[X_t^N(\varphi)\right] = X_0^N\left(\mathcal{P}_t^N(\varphi)\right).$$

In particular $\mathbb{E}\left[X_t^N(1)\right] = X_0^N(1)$ and $\sup_N \mathbb{E}\left[X_t^N(1)\right] < \infty$.

Proof: Note that

$$\mathbb{P}\left[B_t^{w,N} = x\right] = \mathbb{P}\left[B_t^{x,N} = w\right]$$

by Lemma 4.3.3 part (1) and the definition of $B^{\cdot,N}$. So

$$\begin{aligned}
\mathbb{E}\left[X_t^N(\varphi)\right] &= \frac{\alpha^d}{N}\sum_x \mathbb{E}\left[\xi_t^N(x)\right]\varphi(x)\\
&= \frac{\alpha^d}{N}\sum_x \varphi(x)\mathbb{P}\left[B_t^{x,N} \in \xi_0^N\right]\\
&= \frac{\alpha^d}{N}\sum_x \varphi(x)\sum_w \xi_0^N(w)\mathbb{P}\left[B_t^{x,N} = w\right]\\
&= \frac{\alpha^d}{N}\sum_w \xi_0^N(w)\sum_x \varphi(x)\mathbb{P}\left[B_t^{w,N} = x\right]\\
&= \frac{\alpha^d}{N}\sum_w \xi_0^N(w)\mathbb{E}\left[\varphi(B_t^{w,N})\right]\\
&= X_0^N(\mathcal{P}_t^N(\varphi))
\end{aligned}$$

as required. The bound on $X_t^N(1)$ now follows by Lemma 4.2.2. ■

Lemma 4.4.6 *Let $p \geqslant 1$ and $T > 0$. There is a constant $C_{p,T} < \infty$ such that*

$$\mathbb{E}\left[\sup_{t\leqslant T} X_t^N(1)^p\right] \leqslant C_{p,T}\left(X_0^N(1)^p + 1\right)$$

Proof: For $p = 1$ this follows from Lemma 4.4.5 and the Burkholder-Davis inequality. The proof for the case $p > 1$ (using Martingale inequalities from Burkholder (1973)) is essentially that of Lemma 2.4(b) from Cox et al. (2000).) ■

We now examine the square bracket of the martingale term.

Lemma 4.4.7 *Let $s \in (0, \infty)$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$ be bounded and measurable. Then for $p = 1, 2$,*

$$\left|m_s^{N,p}(\varphi)\right| \leqslant Q||\varphi||_\infty^2 X_s^N(1)$$

Proof: Using that $\hat{\xi}_s^N(\cdot) \leqslant 1$,

$$\left|m_s^{N,2}(\varphi)\right| \leqslant \frac{\alpha^d}{N}||\varphi||_\infty^2\sum_y \xi_s^N(y)Q\sum_{x,z} q_N(y,x,z)$$

and noting that

$$\begin{aligned}\sum_{x,z} q_N(y,x,z) &= \frac{\alpha^{2d}}{Q}\sum_n\sum_z \mathbb{1}\{z\in\phi_n^N(y)\}\phi_n\sum_x \mathbb{1}\{x\in\Phi_n^N(y)\}\\ &= \frac{\alpha^{2d}}{Q}\sum_n\sum_z \mathbb{1}\{y\in\phi_n^N(z)\}\#(\Phi_n)\phi_n\\ &= \frac{\alpha^{2d}}{Q}\sum_n \#(\Phi_n)^2\phi_n = 1,\end{aligned} \tag{4.4.5}$$

we have the result for $i=2$. Similarly, using $\hat{\xi}_s^N(y), \xi_s^N(z) \leqslant 1$,

$$\left|m_s^{N,1}\right| \leqslant \frac{\alpha^d}{N}||\varphi||_\infty^2 \sum_x \xi_s^N(x) Q \sum_{y,z} q_N(y,x,z).$$

We note

$$\begin{aligned}\sum_{y,z} q_N(y,x,z) &= \frac{\alpha^{2d}}{Q}\sum_n\sum_z \mathbb{1}\{z\in\Phi_n^N(y)\}\sum_y \mathbb{1}\{x\in\Phi_n^N(y)\}\phi_n\\ &= \frac{\alpha^{2d}}{Q}\sum_n\sum_z \mathbb{1}\{z\in\Phi_n^N(y)\}\sum_y \mathbb{1}\{-y\in\Phi_n^N(-x)\}\phi_n\\ &= \frac{\alpha^{2d}}{Q}\sum_n \#(\Phi_n)^2\phi_n = 1\end{aligned}$$

where we use a translation by $-x-y$ to get from the first line to the second and then proceed as in (4.4.5). The result for $i=1$ follows. ■

Lemma 4.4.8 *Let $p \geqslant 1$ and let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be bounded. Then for any $t \in (0,\infty)$*

$$\sup_N \mathbb{E}\left[\sup_{s\leqslant t}\left|M_s^N(\varphi)\right|^p\right] < \infty$$

Proof: By Lemma 4.4.7,

$$\mathbb{E}\left[\sup_{s\leqslant t}\left|M_s^N(\varphi)\right|^p\right] \leqslant Ct\mathbb{E}\left[\sup_{s\leqslant t} X_s^N(\varphi)^p\right]$$

and the result now follows from Lemma 4.4.6. ■

4.5 Tightness and convergence

Recall that $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ denotes the space of càdlàg paths indexed by $[0,\infty)$ on $\mathcal{M}_F(\mathbb{R}^d)$ with the Skorokhod topology. Let $C_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ be the subspace of such paths which are continuous, endowed with the topology of uniform convergence on compact sets.

Lemma 4.5.1 *$(\mathbb{Q}_N)$ is a tight sequence of probability measures on $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ and all of its limit points are supported by $C_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$.*

Proof: The argument is essentially the same as that given in Cox et al. (2000) (or, for that matter, in Cox and Perkins (2008) or Cox et al. (2010)) and we omit it. One could also prove this theorem using Theorem D.2 and the results from Section 4.4. ■

At this point we require a statement to the effect that the mean field simplication takes place in our situation. The proof of this comes from estimates obtained via the dual process and these can be found in Sections 4.6 and 4.7. To achieve the proper chronology one should insert Sections 4.6 and 4.7 here, but to prove Theorem 4.2.3 we require only the final result of Section 4.7:

Lemma 4.5.2 *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be bounded and Lipschitz. Then for each $t > 0$,*

$$\mathbb{E}\left[\left|\langle M^N(\varphi)\rangle_t - Q(\gamma_1+\gamma_2)\int_0^t X_s^N(\varphi^2)ds\right|\right] \to 0$$

as $N \to \infty$.

We can now give

PROOF: [Of Theorem 4.2.3.] Let $\varphi \in C_0^\infty(\mathbb{R}^d)$. Let $\mathbb{Q}$ be a limit point of $(\mathbb{Q}_N)$. By Lemma 4.5.1 and Skorokhods theorem (3.1.8 in Ethier and Kurtz (1986)) we may assume (via a change of probability space) that there is a process X with law $\mathbb{Q}$ and a subsequence N_k such that

$$X^{N_k} \to X \text{ a.s. in } D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty). \tag{4.5.1}$$

Let $\varphi \in C_0^\infty(\mathbb{R}^d)$. We note

$$\begin{aligned}
&\sup_{t\leqslant T}\left|\int_0^t X_s^{N_k}\left(\mathcal{A}^{N_k}\varphi\right)ds - \int_0^t X_s\left(\frac{\sigma^2 R}{2}\Delta(\varphi)\right)ds\right| \qquad (4.5.2)\\
&\quad\leqslant \sup_{t\leqslant T}\left|\int_0^t X_s^{N_k}\left(\mathcal{A}^{N_k}\varphi - \frac{\sigma^2 R}{2}\Delta(\varphi)\right)ds\right| + \sup_{t\leqslant T}\left|\int_0^t X_s^{N_k}\left(\frac{\sigma^2 R}{2}\Delta(\varphi)\right) - X_s\left(\frac{\sigma^2 R}{2}\Delta(\varphi)\right)ds\right|\\
&\qquad\to 0 \text{ as } k\to\infty \text{ a.s.}
\end{aligned}$$

To see why, use Lemma 4.4.4 to show that the first term on the right hand side tends a.s. to zero, whilst the dominated convergence theorem (using Lemma 4.4.6 to get the dominating function) and (4.5.1) show the second term tends a.s. to zero.

Define

$$M_t(\varphi) = X_t(\varphi) - X_0(\varphi) - \int_0^t X_s\left(\frac{\sigma^2 R}{2}\Delta(\varphi)\right)ds.$$

Equations (4.5.1), (4.5.2) and ($\mathscr{C}5$) show that $|M_t^{N_k}(\varphi) - M_t(\varphi)| \to 0$ a.s. and we can use a.s. continuity of $t \mapsto M_t$ (from Lemma 4.5.1) to deduce a.s. uniform convergence on $[0,t]$:

$$\sup_{s\leqslant t}\left|M_s^{N_k}(\varphi) - M_s(\varphi)\right| \to 0 \tag{4.5.3}$$

For $i = 1,\ldots,m$ let $h_i : \mathcal{M}_F(\mathbb{R}^d) \to \mathbb{R}$ be bounded and continuous and let $0 \leqslant t_1 < \ldots < t_m \leqslant s < t$. Equations (4.5.1), (4.5.3) and Lemma 4.4.8 imply that

$$\begin{aligned}
&\mathbb{E}\Big[\Big(M_t(\varphi)^2 - M_s(\varphi)^2 - Q(\gamma_1+\gamma_2)\int_s^t X_r(\varphi^2)dr\Big)\prod_1^n h_i(X_{t_i})\Big]\\
&\quad= \lim_{k\to\infty}\mathbb{E}\Big[\Big(M_t^{N_k}(\varphi)^2 - M_s^{N_k}(\varphi)^2 - Q(\gamma_1+\gamma_2)\int_s^t X_r^{N_k}(\varphi^2)dr\Big)\prod_1^n h_i(X_{t_i}^{N_k})\Big]. \qquad (4.5.4)
\end{aligned}$$

Lemma 4.5.2 implies that (4.5.4) is equal to

$$\lim_{k\to\infty}\mathbb{E}\Big[\Big(M_t^{N_k}(\varphi)^2 - M_s^{N_k}(\varphi)^2 - \langle M^{N_k}(\varphi)\rangle_t + \langle M^{N_k}(\varphi)\rangle_s\Big)\prod_1^n h_i(X_{t_i}^{N_k})\Big]$$

which is zero by the martingale properties of M stated in Lemma 4.4.3. Thus X satisfies the martingale problem posed in Definition 4.1.5. Therefore $\mathbb{Q} = \mathbb{Q}^{Q(\gamma_1+\gamma_2),R\sigma^2}$. Since our limit point $\mathbb{Q}$ was arbitrary, it follows that $\mathbb{Q}_N \to \mathbb{Q}^{Q(\gamma_1+\gamma_2),R\sigma^2}$ as claimed. ■

4.6 The dual particle system.

The remaining two sections of this chapter will be concerned solely with proving Lemma 4.5.2.

We now consider tracing back the types of more than one site at once. In the Voter model when two ancestral lineages coalesce only one of the lineages jumps. For bursting processes there is positive probability of two lineages both moving from different sites to the same parent in a single reproduction event. They do not have to coalesce in this way since the reproduction event causing coalescence may be centered about the position of one of the walks. This leads us to the following realisation.

Remark 4.6.1 *Let $(B_s^{x,N}))_{s=0}^t$ denote the random walk tracing back the sources of opinions of the site x from time t. Then unless all the Φ_n are one point sets, for $x \neq y$ the lineages $(B_s^{x,N}))_{s=0}^t$ and $(B_s^{y,N}))_{s=0}^t$ are not independent, even before they coalesce.*

To see this, suppose $x, y, z \in S_N$ and n are such that $x, y \in z + \Phi_n^N$ and $x \neq y$. Then the information as to whether the Poisson process $\Lambda^N(z,n)$ has jumped in $(s, s+\epsilon]$ affects the motion of both the walks $(B_s^{x,N}))_{s=0}^t$ and $(B_s^{y,N}))_{s=0}^t$.

We already denote the random walks of our coalescing dual process by $B^{\cdot,N}$ and we now go on to specify the precise coupling between the movement of the lineages. The time reversibility and spatial homogeneity of the underlying Poisson point process Λ permits us to describe our dual process as follows.

Definition 4.6.2 *The dual of our bursting process over time $[0,t]$ is a system $\{B^{x,N}\}$ of coalescing random walks which behave as follows. For $s \in [0,t]$*

- *If $\Lambda^N(x,n)$ jumps at time s then for all y such that $B_{s-}^{y,N} \in x + \Phi_n^N$, $B_s^{y,N} = x$.*
- *No other movement occurs.*

That is, a reproduction event causes all affected random walks to move to the origin of the burst.

For $x, y \in S_N$ define

$$\tau^N(x,y) = \inf\{s \in [0,t]\,;\, B_s^{x,N} = B_s^{y,N}\}$$

where as usual $\inf \varnothing = \infty$. We suppress the dependence of $B^{x,N}$ and $\tau^N(x,y)$ on t since it will almost always be obvious from the context. On the rare occasion that it is necessary to be particular (e.g. the proof of Lemma 4.7.5) we will say explicitly which time we are tracing back from.

Note that Definition 4.6.2 defines a right continuous sytem of walks, whereas the ancestral lineages are really left continuous. Recall Remark 4.3.2 where we noted each lineage is stochastically continuous and thus we can ignore this technicality.

Remark 4.6.3 *The system $\{B^{x,N}\,;\, x \in A\}$ is translation invariant in the sense that*

$$\mathbb{P}[\forall a \in A,\ B_s^{a,N} = y_a] = \mathbb{P}[\forall a \in A,\ B_s^{a+z,N} = y_a + z]$$

where $y_a \in S_N$, symmetric (by Lemma 4.3.3) in the sense that

$$\mathbb{P}[\forall a \in A,\ B_s^{a,N} = 0] = \mathbb{P}[\forall a \in A,\ B_s^{-a,N} = 0]$$

and time reversible in the senses that

$$\mathbb{P}\left[B_s^{z,N} = y\right] = \mathbb{P}\left[B_s^{y,N} = z\right]$$

and

$$\mathbb{P}\left[B_s^{y,N} = u, B_s^{z,N} = v, \tau^N(y,z) > s\right] = \mathbb{P}\left[B_s^{u} = y, B_s^{v,N} = z, \tau^N(u,v) > s\right].$$

This completes our description of the dual process. Recall that when we outlined the heuristics of the proof in Section 4.1.4 the following quantity played a major role.

Definition 4.6.4 *Let $\epsilon_N^* = N^{-1/4}$.*

Our first big step towards Lemma 4.5.2 is the following, and we will spend the rest of this section proving:

Lemma 4.6.5 *It holds that*

$$\sup_{y\neq z, u\neq 0} N\mathbb{P}\left[B^{y,N}_{\epsilon^*_N} - B^{z,N}_{\epsilon^*_N} = u\right] \to 0$$

where the supremum is over $y, z, u \in S_N$ such that $y \neq z$ and $u \neq 0$.

Remark 4.6.6 *By taking $\epsilon^*_N = t$ in the proof of this Lemma, one obtains $\mathbb{P}\left[B^y_{Nt} - B^z_{Nt} = u\right] \leqslant CN^{-3/2}$ for some constant C. We require only the stated result.*

PROOF: First we unravel the time and spatial rescaling. Let $v = \sqrt{N}(y - z) \neq 0$ and let

$$\hat{C}^v_t = \sqrt{N}\left(B^y_t - B^z_t\right).$$

Then we must prove that for given $u \in \alpha\mathbb{Z}^d$,

$$N\mathbb{P}\left[\hat{C}^v_{N\epsilon^*_N} = u\right] \to 0 \tag{4.6.1}$$

uniformly over $v \neq 0$.

The Markov property and spatial homogeneity of bursting processes imply that for sake of investigating a single jump of $\hat{C}^v$, we need only be concened with its current value; that is to say $\hat{C}^v$ is Markov and has independent increments. However the evolution of C^v is not spatially homogeneous. For as long as $|\hat{C}^v_t| > 2L$ the bursting events controlling our two walkers are independent, and $\hat{C}^v$ moves as a rate $2R\sum_{n=1}^{\infty}\#(\Phi_n)\phi_n$ random walk with jump distribution W. For $|\hat{C}^v_t| \leqslant 2L$ the situation is not so simple. When our two walkers are close enough together that a single bursting event could affect them both, $\hat{C}^v$ is jumping at a slower rate with a different jump distribution and some probability of absorption at 0.

Definition 4.6.7 *We say a random walk on $\alpha\mathbb{Z}^d$ is a standard motion if it jumps at rate $2R$ with jump distribution W.*

We can describe the transitions of a walker D in standard motion as follows. To each pair (b, n) such that $0 \in b + \Phi_n$ and each $j \in \{1, 2\}$ we associate a Poisson process $\overline{\Lambda}(b, n, j)$ of rate $\alpha^d\phi_n$. When t is a jump time of $\overline{\Lambda}(b, n, j)$, the transition

$$D_t = D_{t-} + b$$

is made. The reader may readily verify that this describes a standard motion (note that it requires ($\mathscr{C}2$)). We intend to couple $\hat{C}^v$ to a system $\mathcal{D}$ of branching standard motions.

Remark 4.6.8 *In what follows we have walks which are part of the ancestral lineages and walks which are related in some way to the difference between two ancestral lineages. To distinguish verbally between the two we will refer to the ancestral lineages as particles and the differences between ancestral lineages as walkers.*

Crucially, if we took two independent particles $\hat{B}(1)$ and $\hat{B}(2)$ moving at rate $\sum_n \#(\Phi_n)\phi_n$ with jump distribution W, a standard motion describes the walker $\hat{B}(1)_t - \hat{B}(2)_t$. We want to suppress jumps of $\hat{B}(1) - \hat{B}(2)$ in such a way as constructs a walker moving according to the dependent dynamics of $B^y - B^z$. Take a standard motion in state a, and specify that the state 0 is absorbing. Suppress the jumps corresponding to

$$\{(b, n, j)\,;\, j = 1,\ a \in \Phi_n(b)\},$$

(that is, half the jumps which would cause the dependent system $B^y - B^z$ to both move and coalesce), and when a jump corresponding to

$$\{(b, n, j)\,;\, j = 2,\ 0 \in \Phi_n(b + a)\}$$

occurs (the other half of such jumps) we move straight to the site 0. (Of course, this is not the only way to hit 0). The case $j = 1$ corresponds to the first particle jumping and $j = 2$ to the second.

We now construct a system $\mathcal{D}^v$ of random walkers in standard motion. The general principle is that whenever we have to suppress a jump we introduce a new walker, also in standard motion, to compensate. At each time exactly one walker in the system $\mathcal{D}^v$ will carry the title of *youngest walker*. Initially (at time 0) there is a single walker $D(1)^v$ who is the first youngest walker, starts at v, and moves according to a standard motion. At all times our system will involves a finite number of standard motions, $D(\cdot)$, one of which is the youngest walker. The motion of the youngest walker is especially important and we decompose it as above. As the name might suggest, at any fixed time the youngest walker is the particle which was born most recently.

We now describe $\mathcal{D}$ from initial state v.

- Given the current positions of the currently existing walkers $D(1)^v, D(2), \ldots, D(m)$, they move as a system of independent standard motions.
- When the youngest walker makes a transition out of state a, we use our decomposition to look at which of the processes $\overline{\Lambda}(b, n, j)$ caused the transition to take place, and correspondingly the following occurs:
  1. If $a \neq 0$, $j = 1$ and $a \notin \Phi_n(b)$, the transition occurs without additional consequences and the youngest walker is unchanged.
  2. If $a \neq 0$, $j = 2$ and $-a \notin \Phi_n(b)$, the transition occurs without additional consequences and the youngest walker is unchanged.
  3. If $a \neq 0$, $j = 1$ and $a \in \Phi_n(b)$, the youngest walker makes the transition and instantaneously leaves behind a new walker $D(m+1)$ at a. This new walker takes over the title of youngest walker.
  4. If $a \neq 0$, $j = 2$ and $-a \in \Phi_n(b)$, the youngest walker makes the transition and instantaneously leave behind a walker $D(m+1)$ at 0. This new walker holds the title of youngest walker.
  5. As soon as the position of the youngest walker is 0 the system is frozen. The walkers remain constant for all remaining time.

There is very little to prove in showing the process

$$\mathcal{D}^v = \{D(1)^v, D(2), \ldots, D(n_t)\}$$

exists and has a right continuous version which is strongly Markov. We claim that the youngest walker follows the random walk $\hat{C}^v$; Thus we may couple our two systems together in such a way as

$$\hat{C}^v_t = D(n_t)_t$$

(where $D(1) = D(1)^v$). We will prove this claim below, but first let us prove (4.6.1).

Remark 4.6.9 *Cases 1 and 2 correspond to reproduction events which only move one of the particles. In case 3, the position of the youngest walker does not change, only which walker holds the title. This corresponds to the suppressed jumps of Remark 4.6.8. In case 4 we create the jumps causing coalescence (i.e. when the youngest walks hits 0) when both B^y and B^z move. The cases when coalescence is caused by only one of the two particles moving occur as part of 1 and 2. Case 5 corresponds to the behaviour after the particles have coalesced.*

Let n_t be (càdlàg process of) the index at time t of the youngest walker (which, if it is non-zero, is also the number of walkers at t).

$$\begin{aligned}
\mathbb{P}\left[\hat{C}^v_{N\epsilon^*_N} = u\right] &\leqslant \mathbb{P}\left[\exists m \in \{1, \ldots, n_t\},\ D(m)_{N\epsilon^*_N} = u\right] \\
&\leqslant \sum_{n=1}^{\infty} \mathbb{P}\left[\exists m \in \{1, \ldots, n\},\ D(m)_{N\epsilon^*_N} = u\right] \mathbb{P}\left[n_{Nt} = n\right] \\
&\leqslant \sum_{n=1}^{\infty} \sum_{m=1}^{n} \mathbb{P}\left[D(m)_{N\epsilon^*_N} = u\right] \mathbb{P}\left[n_{N\epsilon^*_N} = n\right] \qquad (4.6.2)
\end{aligned}$$

If we set $T(m)$ to be the birth time of $D(m)$, then by Lemma A.3 in Cox et al. (2000),

$$\mathbb{P}\left[D(m)_{N\epsilon_N^*} = u\right] \leqslant \mathbb{P}[D(m)_{N\epsilon_N^*} = D(m)_{T(m)}] = \mathbb{P}[D^0_{N\epsilon_N^*} = 0]$$

where D^0 is a standard motion started at 0. Continuing from (4.6.2),

$$\begin{aligned}\mathbb{P}\left[\hat{C}^v_{N\epsilon_N^*} = u\right] &\leqslant \mathbb{P}\left[D^0_{N\epsilon_N^*} = 0\right] \sum_{n=1}^{\infty} n\mathbb{P}\left[n_{N\epsilon_N^*} = n\right] \\ &= \mathbb{P}\left[D^0_{N\epsilon_N^*} = 0\right] \mathbb{E}[n_{N\epsilon_N^*}] \end{aligned} \tag{4.6.3}$$

which is a bound independent of u (but note n_t depends on v!). $\mathbb{E}[n_{N\epsilon_N^*}]$ is defined for the positive random variable n_t and may (until we prove otherwise) be infinite. So we must estimate $n_{N\epsilon_N^*}$, the number of (moving) walkers in the system $\mathcal{D}^v$. Let $n_\infty = \lim_{t\to\infty} n_t$ denote the total number of walkers of $\mathcal{D}^v$ over all time. Clearly

$$\mathbb{E}[n_{N\epsilon_N^*}] \leqslant \mathbb{E}[n_\infty] \leqslant \mathbb{E}[\widetilde{n}_\infty]$$

where $\widetilde{n}_\infty$ is the total number of jumps initiated by the youngest walker from a point within $\alpha\mathbb{Z}^d \cap B_{2L}(0)$. We will shortly show that $\mathbb{E}[\widetilde{n}_\infty]$ is bounded uniformly over $v \neq 0$. The standard local limit theorem (see, for example, A.3 of Cox et al. (2000)) for simple random walks in $d \geqslant 3$ gives us a constant C such that $\mathbb{P}[D^0_t = 0] \leqslant Ct^{-d/2}$ and we then have that

$$N\mathbb{P}[\hat{C}^v_{N\epsilon_N^*} = u] \leqslant CN(N\epsilon_N^*)^{-3/2} = CN^{-1/8} \to 0.$$

This proves (4.6.1).

PROOF: [that $\mathbb{E}[\widetilde{n}_\infty] < \infty$.] Let us consider briefly the situation at time t when the youngest walker $D(n_t)$ is at $a \in \alpha\mathbb{Z}^d \cap B_{2L}(0)\backslash\{0\}$. Let

$$\mathcal{K} = \{a \in \alpha\mathbb{Z}^d\backslash\{0\}\,;\, \mathbb{P}\left[D^a_T = 0\right]\}$$

where T is the first jump of D^a, a standard motion started from a. A new walker is created precisely when the youngest walker jumps from within $\mathcal{K}$. Of course by ($\mathscr{C}1$) $\mathcal{K} \subseteq \alpha\mathbb{Z}^d \cap B_{2L}(0)$, and hence

$$r_0 = \min\{\mathbb{P}\left[D^a_T = 0\right]\,;\, a \in \mathcal{K}\} > 0.$$

Thus if we are looking to wait until either the youngest walker has left $\mathcal{K}$ into $\alpha\mathbb{Z}^d\backslash\{0\}$ or has hit 0, we wait at most a geometric number of jumps with success probability r_0. After leaving $\mathcal{K}$ up until it re-enters or hits 0, the youngest walker moves as a standard motion. In particular, it moves as a random walk with a bounded symmetric jump distribution. From well known facts concerning transience of random walks in dimensions $d \geqslant 3$,

$$r_1 := \sup\left\{\mathbb{P}\left[\exists t \geqslant 0, D^a_t \in \mathcal{K} \cup \{0\}\right]\,;\, a \in \alpha\mathbb{Z}^d\backslash(\mathcal{K} \cup \{0\})\right\} < 1.$$

Thus the number of returns a standard motion may make to $\mathcal{K}\backslash\{0\}$ is bounded above by a geometric random variable with success probability $1 - r_1 > 0$.

Combining these observations with the strong Markov property of $\mathcal{D}^v$, there are independent random variables R_1 (geometric with parameter $1 - r_1$) and $R_0(k)$ (geometric with parameter r_0) such that

$$\widetilde{n}_\infty \leqslant \sum_{k=0}^{R_0} R_1(k)$$

and thus $\mathbb{E}[\widetilde{n}_\infty] \leqslant \frac{1}{r_0(1-r_1)}$. Note that this does not depend on a. ■

PROOF: [that the youngest walker in $\mathcal{D}^v$ has the distribution of $\hat{C}^v$.] We show the two processes have the same jump distribution and jump rate. Since the process $\mathcal{D}^v$ is strongly Markov it suffices to consider a

single jump of the youngest walker $D(m)$ from some arbitrary $a \in \alpha\mathbb{Z}^d$. We first consider the total jump rate out from a. Since this depends on a, let us write it $J(a)$. Then $J(0) = 0$, and for $a \neq 0$,

$$\begin{aligned}
J(a) &= \alpha^d \Bigg((\sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b), a \notin \Phi_n(b)\}\phi_n + \sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b), -a \notin \Phi_n(b)\}\phi_n \\
&\qquad + \sum_{b,n} \mathbb{1}\{0 \in \Phi(b), -a \in \Phi_n(b)\}\phi_n \Bigg) \\
&= \alpha^d \Bigg(\sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b), a \notin \Phi_n(b)\}\phi_n + \sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b-a), -a \notin \Phi_n(b-a)\}\phi_n \\
&\qquad + \sum_{b,n} \mathbb{1}\{0 \in \Phi(b-a), -a \in \Phi_n(b-a)\}\phi_n \Bigg) \\
&= \alpha^d \sum_{b,n} \mathbb{1}\{0 \in \Phi(b) \text{ or } a \in \Phi(b)\}\phi_n
\end{aligned}$$

which is the jump rate of $\hat{C}^a$. The terms in the first line come from the movement in cases 1,2 and 4 respectively. For $a \neq 0$, if T is the time of the next jump of $D(m)$,

$$\mathbb{P}[D(n_T)^a_T = 0] = \frac{\alpha^d}{J(a)} \left(\sum_n \mathbb{1}\{0 \in \Phi_n(a)\}\phi_n + \sum_n \mathbb{1}\{-a \in \Phi(0)\}\phi_n + \sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b), a \in \Phi_n(b))\phi_n \right)$$

which is the same as for $\hat{C}^a$. Again the terms come from cases 1, 2 and 4 respectively. The first two terms are the coalescence where only one of the two particles jumps. For $c \neq 0$,

$$\mathbb{P}[D(n_T)^a_T = c] = \frac{\alpha^d}{J(a)} \left(\sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b), a \notin \Phi_n(b)\}\phi_n + \sum_{b,n} \mathbb{1}\{0 \in \Phi_n(b), -a \notin \Phi_n(b)\}\phi_n \right)$$

where again matches up to the transitions of $\hat{C}^v$. Here we obtain only terms from cases 1 and 2. ■

This completes the proof of Lemma 4.6.5. ■

4.7 The mean field simplification

In this section we use the estimates of Section 4.6 to take the limit of the square bracket term $M^N(\varphi)$ from our martingale decomposition. There are two main results we need in order to do this, which are given as Lemma 4.7.1 and Lemma 4.7.5. Then finally we give the proof of Lemma 4.5.2 and thus complete the argument leading to our main result. Lemma 4.7.1 is relatively easy to prove and is very similar to an argument from Cox et al. (2010). Lemma 4.7.5 is proof of the mean field simplification and will take us much longer.

Let $\mathcal{F}^N_s$ be the filtration generated by ξ^N_s. With mild abuse of notation let ξ^N_0 also refer to the initial set of 1s (i.e. the set $\{x \in S_N\,;\, \xi^\alpha_0(x) = 1\}$).

Lemma 4.7.1 *Let $\epsilon^*_N \leqslant t_1 \leqslant t_2 \leqslant T$. Then there exists $C \in (0,\infty)$ such that for all $N \in \mathbb{N}$*

$$\mathbb{E}\left[\left(\int_{t_1}^{t_2} \mathbb{E}\left[m^{N,i}_s(\varphi) \middle| \mathcal{F}^N_{s_1-\epsilon^*_N} \right] - m^{N,i}_{s-\epsilon^*_N}(\varphi)\, ds \right)^2 \right] \leqslant C\epsilon^*_N (t_2 - t_1) T \left(1 + X^N_0(1)^2\right) \tag{4.7.1}$$

for $i = 1, 2$.

Proof: Note that whenever $s_2 > s_1 + \epsilon^*_N$, $m^{N,i}_{s_1}(\varphi) - \mathbb{E}\left[m^{N,i}_{s_1}(\varphi) \middle| \mathcal{F}_{s_1-\epsilon^*_N} \right]$ is $\mathcal{F}^N_{s_2-\epsilon^*_N}$ measurable, and hence

$$\mathbb{E}\left[\left(m^{N,i}_{s_1}(\varphi) - \mathbb{E}\left[m^{N,i}_{s_1}(\varphi) \middle| \mathcal{F}^N_{s_1-\epsilon^*_N} \right] \right) \left(m^{N,i}_{s_2}(\varphi) - \mathbb{E}\left[m^{N,i}_{s_2}(\varphi) \middle| \mathcal{F}^N_{s_2-\epsilon^*_N} \right] \right) \right] = 0.$$

Hence (4.7.1) is equal to

$$\int_{t_1}^{t_2}\int_{t_1}^{(s_1+\epsilon_N^*)\wedge t_2}\mathbb{E}\left[\left(m_{s_1}^{N,i}(\varphi)-\mathbb{E}\left[m_{s_1}^{N,i}(\varphi)\middle|\mathcal{F}^N_{s_1-\epsilon_N^*}\right]\right)\left(m_{s_2}^{N,i}(\varphi)-\mathbb{E}\left[m_{s_2}^{N,i}(\varphi)\middle|\mathcal{F}^N_{s_2-\epsilon_N^*}\right]\right)\right]ds_2ds_1$$
$$\leqslant 2\int_{t_1}^{t_2}\int_{s_1}^{(s_1+\epsilon_N^*)\wedge t_2}\mathbb{E}\left[\left(m_{s_1}^{N,i}(\varphi)-\mathbb{E}\left[m_{s_1}^{N,i}(\varphi)\middle|\mathcal{F}^N_{s_1-\epsilon_N^*}\right]\right)\left(m_{s_2}^{N,i}(\varphi)-\mathbb{E}\left[m_{s_2}^{N,i}(\varphi)\middle|\mathcal{F}^N_{s_2-\epsilon_N^*}\right]\right)\right]ds_2ds_1$$

Applying Lemma 4.4.7 the above is less than or equal to

$$2||\varphi||_\infty^2 Q\mathbb{E}\left[\iint_{t_1\leqslant s_1\leqslant s_2\leqslant(s_1+\epsilon_N^*)\wedge t_2}X_{s_1}^N(1)X_{s_2}^N(1)ds_2ds_1\right].$$

By Lemma 4.4.5 $\mathbb{E}\left[X_{s_2}^N(1)|\mathcal{F}_{s_1}^N\right]=X_{s_1}^N(1)$ and thus

$$\mathbb{E}\left[X_{s_1}^N(1)X_{s_2}^N(1)\right]=\mathbb{E}\left[\mathbb{E}\left[X_{s_2}^N(1)|\mathcal{F}_{s_1}^N\right]X_{s_1}^N(1)\right]=\mathbb{E}\left[\left(X_{s_1}^N(1)\right)^2\right].$$

which gives us that (4.7.1) is less than or equal to

$$2||\varphi||_\infty^2 Q\int_{t_1}^{t_2}\mathbb{E}\left[X_{s_1}^N(1)^2\right](\epsilon_N^*\wedge(t_2-s_1))\,ds_1$$
$$\leqslant 2||\varphi||_\infty^2 Q\epsilon_N^*\int_{t_1}^{t_2}\mathbb{E}\left[X_{s_1}^N(1)^2\right]ds_1.$$

Using Lemma 4.4.6 completes the proof. ■

Let

$$\gamma_1^N=\sum_{e,f\in S_N}\left(\sum_n\mathbb{1}\left\{0,e\in\Phi_n^N(f)\right\}\phi_n\right)\mathbb{P}\left[\tau^N(0,e)\leqslant\epsilon_N^*,\tau^N(0,f)>\epsilon_N^*\right]$$
$$\gamma_2^N=\sum_{e,f\in S_N}\left(\sum_n\mathbb{1}\left\{e,f\in\Phi_n^N\right\}\phi_n\right)\mathbb{P}\left[\tau^N(0,e)>\epsilon_N^*,\tau^N(0,f)>\epsilon_N^*\right]$$

Lemma 4.7.2 *For $p=1,2$, as $N\to\infty$, $\gamma_p^N\to\gamma_p$.*

Proof: Since $N\epsilon_N^*\to\infty$ we simply note

$$\gamma_1^N=\sum_{e,f\in\alpha\mathbb{Z}^d}\left(\sum_n\mathbb{1}\left\{e,f\in\Phi_n\right\}\phi_n\right)\mathbb{P}\left[\tau(0,e)\leqslant N\epsilon_N^*,\tau(0,f)>N\epsilon_N^*\right]$$

and the result follows. The argument for γ_2 is no different. ■

Let

$$r_N(x,z)=\frac{\alpha^{2d}}{Q}\sum_n\#\left(\Phi_n^N(-x)\cap\Phi_n^N(-z)\right)\phi_n\qquad\text{if } x\neq z$$

and $r_N(0,0)=0$. Let $\mathcal{R}$ be an $\alpha\mathbb{Z}^d$ valued random variable such that $\mathbb{P}[\mathcal{R}=w]=r(0,w)$ with the convention that $r=r_1$.

We next give a result which will be used to prove the crucial Lemma 4.7.5. This result essentially says that as far as our limiting process is concerned, if a pair lineages moves for time $[0,\epsilon_N^*]$ and are conditioned to finish in the same (finite) set of sites then that pair of lineages will coalesce.

Lemma 4.7.3 *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be bounded and measurable. Then there exists a constant $C \in (0,\infty)$ such that*

$$\frac{\alpha^d}{N}\sum_{x,y,z}\varphi(x)\varphi(z)\mathbb{P}\left[B^{x,N}_{\epsilon^*_N}\in\xi^N_0, B^{z,N}_{\epsilon^*_N}\in\xi^N_0, \tau^N(x,z)>\epsilon^*_N\right]Qq_N(y,x,z) \tag{4.7.2}$$
$$\leqslant C||\varphi||^2_\infty X^N_0(1)^2\left(N\sup_{u\neq v, w\neq 0}\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}-B^{v,N}_{\epsilon^*_N}=w\right]\right)$$

Remark 4.7.4 *Lemma 4.7.3 is designed to pair up with Lemma 4.6.5.*

Proof: Let us first note that

$$\sum_w r_N(0,w)=\frac{\alpha^{2d}}{Q}\sum_n\sum_{y,w}\mathbb{1}\{0,w\in\Phi^N_n(y)\}\phi_n=1$$

by the same calculation as (4.4.5). Now, $\sum_y q_N(y,x,z)=r_N(x,z)$ so (4.7.2) is less than or equal to

$$\begin{aligned}
&||\varphi||^2_\infty\frac{Q\alpha^d}{N}\sum_{x,z}\mathbb{P}\left[B^{x,N}_{\epsilon^*_N}\in\xi^N_0, B^{z,N}_{\epsilon^*_N}\in\xi^N_0,\tau^N(x,z)>\epsilon^*_N\right]r_N(x,z)\\
&=||\varphi||^2_\infty\frac{Q\alpha^d}{N}\sum_{x,z,u,v}\xi^N_0(u)\xi^N_0(v)\mathbb{P}\left[B^{x,N}_{\epsilon^*_N}=u, B^{z,N}_{\epsilon^*_N}=v,\tau^N(x,z)>\epsilon^*_N\right]r_N(x,z)\\
&=||\varphi||^2_\infty\frac{Q\alpha^d}{N}\sum_{x,z,u,v}\xi^N_0(u)\xi^N_0(v)\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}=x, B^{v,N}_{\epsilon^*_N}=z,\tau^N(u,v)>\epsilon^*_N\right]r_N(x,z)\\
&\leqslant||\varphi||^2_\infty Q\alpha^{-d}\sum_{u,v}\frac{\alpha^d\xi^N_0(u)}{N}\frac{\alpha^d\xi^N_0(v)}{N}N\sum_{x,z}\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}=x, B^{v,N}_{\epsilon^*_N}=z\right]r_N(x,z)\\
&\leqslant||\varphi||^2_\infty Q\alpha^{-d}\sum_{u,v}\frac{\alpha^d\xi^N_0(u)}{N}\frac{\alpha^d\xi^N_0(v)}{N}N\sum_w r_N(0,w)\sum_x\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}=x, B^{v,N}_{\epsilon^*_N}=x+w\right]\\
&=||\varphi||^2_\infty Q\alpha^{-d}\sum_{u,v}\frac{\alpha^d\xi^N_0(u)}{N}\frac{\alpha^d\xi^N_0(v)}{N}N\sum_w r_N(0,w)\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}-B^{v,N}_{\epsilon^*_N}=w\right]\\
&\leqslant||\varphi||^2_\infty Q\alpha^{-d}X^N_0(1)^2\left(N\sup_{u\neq v,w\neq 0}\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}-B^{v,N}_{\epsilon^*_N}=w\right]\right)
\end{aligned}$$

where we used Remark 4.6.3 to get from the second to third lines, and set $z=x+w$ to get from the fourth to fifth. ■

We now have the crucial Lemma.

Lemma 4.7.5 *Let $\varphi:\mathbb{R}^d\to\mathbb{R}$ be bounded and Lipschitz. There exists $C\in(0,\infty)$ such that for $p=1,2$ and all $s\in(\epsilon^*_N,\infty)$,*

$$\begin{aligned}
&\left|\mathbb{E}\left[m^{N,p}_s(\varphi)\,\middle|\,\mathcal{F}^N_{s-\epsilon^*_N}\right]-Q\gamma_pX_{s-\epsilon^*_N}(\varphi^2)\right|\\
&\qquad\leqslant C||\varphi||\left(\frac{1}{N^{1/4}}+\frac{1}{N^{1/2}}+N\sup_{u\neq v,w\neq 0}\mathbb{P}\left[B^{u,N}_{\epsilon^*_N}-B^{v,N}_{\epsilon^*_N}=w\right]+|\gamma^N_p-\gamma_p|\right)X^N_{s-\epsilon^*_N}(1)
\end{aligned}$$

where $||\varphi||=||\varphi||_\infty\left(||\varphi||_{\text{lip}}+||\varphi||_\infty\right)+||\varphi^2||_{\text{lip}}$.

Remark 4.7.6 *It will be important to keep track of different times and as such we need to consider the dependence of the system of walks $\{B^{x,N}\}_{x\in S_N}$ on the time at which we trace the lineage back from. For the proof of Lemma 4.7.5 we adopt the convention that $B^{x,N}$ refers to tracing lineages back from the fixed time $s\in(0,\infty)$. This convention also affects the coalescence times*

$$\tau^N(x,y)=\inf\{t\geqslant 0\,;\,B^{x,N}_t=B^{y,N}_t\}.$$

PROOF: For now let us concentrate on the case $p = 1$. Note that

$$\mathbb{E}\left[\xi_s^N(x)\xi_S^N(z)\hat{\xi}_s^N(y) \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right] = \mathbb{P}\left[\xi^N_{s-\epsilon_N^*}(B^{x,N}_{\epsilon_N^*}) = 1, \xi^N_{s-\epsilon_N^*}(B^{z,N}_{\epsilon_N^*}) = 1, \xi^N_{s-\epsilon_N^*}(B^{y,N}_{\epsilon_N^*}) = 0 \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right].$$

Thanks to the Markov property the lineages $B^{w,N}$ in the above expression, which run back over time $[s, s-\epsilon_N^*]$, are independent of $\mathcal{F}^N_{s-\epsilon_N^*}$. Of course also $\xi^N_{s-\epsilon_N^*} \in \mathcal{F}^N_{s-\epsilon_N^*}$, and therefore the conditioning $\mathbb{P}[\cdots|\mathcal{F}^N_{s-\epsilon_N^*}]$ acts as normal expectation with respect to *only* the random variables $\{B_t^{w,N}\}_{t=0}^{\epsilon_N^*}$.

We thus have

$$\begin{aligned}
&\mathbb{E}\left[m_s^{N,1}(\varphi)) \;\middle|\; \mathcal{F}^N_{s-s\epsilon_N^*}\right] \\
&\quad = \frac{\alpha^d}{N}\sum_{x,y,z}\varphi(x)\varphi(z)Qq_N(y,x,z)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*}, B^{z,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*}, B^{y,N}_{\epsilon_N^*} \notin \xi^N_{s-\epsilon_N^*} \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right]
\end{aligned}$$

The proof will come in four stages. First we will use smoothness of φ to change the $\varphi(x)\varphi(z)$ to $\varphi(x)^2$. Then we will apply Lemma 4.7.3 twice to change the $\mathbb{P}[\cdots|\mathcal{F}^N_{s-\epsilon_N^*}]$ into

$$\begin{aligned}
&\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*}, \tau^N(x,z) \leqslant \epsilon_N^*, \tau^N(x,y) > \epsilon_N^* \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right] \\
&\quad = \sum_u \xi^N_{s-\epsilon_N^*}(u)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} = u, \tau^N(x,z) \leqslant \epsilon_N^*, \tau^N(x,y) > \epsilon_N^* \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right] \\
&\quad = \sum_u \xi^N_{s-\epsilon_N^*}(u)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} = u, \tau^N(x,z) \leqslant \epsilon_N^*, \tau^N(x,y) > \epsilon_N^*\right].
\end{aligned}$$

Our third stage will be a change of variables and rearrangement using translation and symmetry. This will leave us with a term which as our fourth and final stage we can take a limit of. In essence, the $\sum_u \xi^N_{s-\epsilon_N^*}(u)$ will become $X^N_{s-\epsilon_N^*}$ and the remaining $\sum_{x,y,z}\varphi(x)^2Qq(y,x,z)\mathbb{P}[\cdots]$ will become $Q\gamma_1\varphi(u)^2$.

Along the way we will gradually accumulate error terms which we will record as $\mathcal{E}_1^N, \mathcal{E}_2^N \ldots$. The term which gives a non-zero contribution we keep track of as $\Sigma_1^N, \Sigma_2^N \ldots$.

PART ONE: Let

$$\Sigma_1^N = \frac{\alpha^d}{N}\sum_{x,y,z}\varphi(x)^2Qq_N(y,x,z)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*}, B^{z,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*}, B^{y,N}_{\epsilon_N^*} \notin \xi^N_{s-\epsilon_N^*} \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right]$$

and then $\mathbb{E}\left[m_s^{N,1}(\varphi)) \;\middle|\; \mathcal{F}^N_{s-s\epsilon_N^*}\right] - \Sigma_1^N = \mathcal{E}_1^N$ where

$$\begin{aligned}
|\mathcal{E}_1^n| &\leqslant Q||\varphi||_\infty\frac{\alpha^d}{N}\sum_{x,y,z}|\varphi(x)-\varphi(z)|\,Qq_N(y,x,z)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*}, B^{z,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*} \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right] \\
&\leqslant Q||\varphi||_\infty||\varphi||_{\text{lip}}\frac{\alpha^d}{N}\sum_x\sum_u\xi^N_{s-\epsilon_N^*}(u)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} = u \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right]\sum_{y,z}q_N(x,y,z)|x-z| \\
&= Q||\varphi||_\infty||\varphi||_{\text{lip}}\frac{\alpha^d}{N}\sum_u\xi^N_{s-\epsilon_N^*}(u)\sum_x\mathbb{P}\left[B^{x,N}_{\epsilon_N^*} = u\right]\sum_z r_N(x,z)|x-z| \\
&= Q||\varphi||_\infty||\varphi||_{\text{lip}}\sum_u\frac{\alpha^d\xi^N_{s-\epsilon_N^*}(u)}{N}\sum_x\mathbb{P}\left[B^{u,N}_{\epsilon_N^*} = x\right]\sum_w r_N(0,w)|w| \\
&= Q||\varphi||_\infty||\varphi||_{\text{lip}}X^N_{s-\epsilon_N^*}(1)\mathbb{E}\left[\frac{|\mathcal{R}|}{\sqrt{N}}\right] \\
&\leqslant \frac{C}{\sqrt{N}}||\varphi||_\infty||\varphi||_{\text{lip}}X^N_{s-\epsilon_N^*}(1)
\end{aligned}$$

where $w = z - x$. $\mathbb{E}[|\mathcal{R}|] < \infty$ by $(\mathscr{C}4)$.

Part Two: Let

$$\Sigma_2^N = \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(x)^2 Q q_N(y,x,z) \mathbb{P}\left[B^{x,N}_{\epsilon^*_N} \in \xi^N_{s-\epsilon^*_N}, \tau^N(x,z) \leqslant \epsilon^*_N, B^{y,N}_{\epsilon^*_N} \notin \xi^N_{s-\epsilon^*_N} \;\middle|\; \mathcal{F}^N_{s-\epsilon^*_N} \right]$$

and then $\Sigma_2^N + \mathcal{E}_2^N = \Sigma_1^N$ where

$$\mathcal{E}_2^N = \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(x)^2 Q q_N(y,x,z) \mathbb{P}\left[B^{x,N}_{\epsilon^*_N} \in \xi^N_{s-\epsilon^*_N}, B^{z,N}_{\epsilon^*_N} \in \xi^N_{s-\epsilon^*_N}, \tau^N(x,z) > \epsilon^*_N, B^{y,N}_{\epsilon^*_N} \notin \xi^N_{s-\epsilon^*_N} \;\middle|\; \mathcal{F}^N_{s-\epsilon^*_N} \right].$$

Dropping $B^{y,N}_{\epsilon^*_N} \notin \xi^N_{s-\epsilon^*_N}$ and using Lemma 4.7.3 (applied with the zero time of Lemma 4.7.3 taken to be what is currently time $s - \epsilon^*_N$),

$$|\mathcal{E}_2^N| \leqslant C||\varphi||^2_\infty X^N_{s-\epsilon^*_N}(1)^2 \left(N \sup_{u \neq v, w \neq 0} \mathbb{P}\left[B^{u,N}_{\epsilon^*_N} - B^{v,N}_{\epsilon^*_N} = w \right] \right).$$

We cut away one last error term. Let

$$\Sigma_3^N = \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(x)^2 Q q_N(y,x,z) \mathbb{P}\left[B^{x,N}_{\epsilon^*_N} \in \xi^N_{s-\epsilon^*_N}, \tau^N(x,z) \leqslant \epsilon^*_N, \tau^N(x,y) > \epsilon^*_N \;\middle|\; \mathcal{F}^N_{s-\epsilon^*_N} \right]$$

and then $\Sigma_3^N = \Sigma_2^N + \mathcal{E}_3^N$ where

$$\mathcal{E}_3^N = \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(x)^2 Q q_N(y,x,z) \mathbb{P}\left[B^{x,N}_{\epsilon^*_N} \in \xi^N_{s-\epsilon^*_N}, \tau^N(x,z) \leqslant \epsilon^*_N, B^{y,N}_{\epsilon^*_N} \in \xi^N_{s-\epsilon^*_N}, \tau^N(x,y) > \epsilon^*_N \;\middle|\; \mathcal{F}^N_{s-\epsilon^*_N} \right].$$

By essentially the same application of Lemma 4.7.3 we obtain

$$|\mathcal{E}_3^N| \leqslant C||\varphi||^2_\infty X^N_{s-\epsilon^*_N}(1)^2 \left(N \sup_{u \neq v, w \neq 0} \mathbb{P}\left[B^{u,N}_{\epsilon^*_N} - B^{v,N}_{\epsilon^*_N} = w \right] \right).$$

Part Three: Note that

$$\Sigma_3^N = \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(x)^2 Q q_N(y,x,z) \sum_u \xi^N_{s-\epsilon^*_N}(u) \mathbb{P}\left[B^{x,N}_{\epsilon^*_N} = u, \tau^N(x,z) \leqslant \epsilon^*_N, \tau^N(x,y) > \epsilon^*_N \;\middle|\; \mathcal{F}^N_{s-\epsilon^*_N} \right].$$

By the Markov property at time $s - \epsilon^*_N$, translation, symmetry and translation again

$$\begin{aligned}
&\mathbb{P}\left[B^{x,N}_{\epsilon^*_N} = u, \tau^N(x,z) \leqslant \epsilon^*_N, \tau^N(x,y) \geqslant \epsilon^*_N \;\middle|\; \mathcal{F}^N_{s-\epsilon^*_N} \right] \\
&\quad = \mathbb{P}\left[B^{x,N}_{\epsilon^*_N} = u, \tau^N(x,z) \leqslant \epsilon^*_N, \tau^N(x,y) > \epsilon^*_N \right] \\
&\quad = \mathbb{P}\left[B^{0,N}_{\epsilon^*_N} = u - x, \tau^N(0,z-x) \leqslant \epsilon^*_N, \tau^N(0,y-x) > \epsilon^*_N \right] \\
&\quad = \mathbb{P}\left[B^{0,N}_{\epsilon^*_N} = x - u, \tau^N(0,x-z) \leqslant \epsilon^*_N, \tau^N(0,x-y) > \epsilon^*_N \right] \\
&\quad = \mathbb{P}\left[B^{u,N}_{\epsilon^*_N} = x, \tau^N(u,x-z+u) \leqslant \epsilon^*_N, \tau^N(u,x-y+u) > \epsilon^*_N \right] \\
&\quad = \mathbb{P}\left[B^{u,N}_{\epsilon^*_N} = x, \tau^N(u,e+u) \leqslant \epsilon^*_N, \tau^N(u,f+u) > \epsilon^*_N \right]
\end{aligned}$$

where $e = x - z$ and $f = x - y$. Note also

$$\begin{aligned}
\mathbb{1}\{x, z \in \Phi^N_n(y)\} &= \mathbb{1}\{x, x - e \in \Phi^N_n(x - f)\} \\
&= \mathbb{1}\{0, -e \in \Phi^N_n(-f)\}.
\end{aligned}$$

In our new variables,

$$\Sigma_3^N = \frac{\alpha^d}{N}\sum_{x,e,f,u}\varphi(x)^2\xi^N_{s-\epsilon_N^*}(u)Qq_N(-f,0,-e)\mathbb{P}\left[B^{u,N}_{\epsilon_N^*}=x,\tau^N(u,e+u)\leqslant\epsilon_N^*,\tau^N(u,f+u)>\epsilon_N^*\right]$$
$$=\sum_u\frac{\alpha^d\xi^N_{s-\epsilon_N^*}(u)}{N}Q\sum_{e,f}q_N(-f,0,-e)\mathbb{E}\left[\varphi\left(B^{u,N}_{\epsilon_N^*}\right)^2\mathbb{1}\{\tau^N(u,e+u)\leqslant\epsilon_N^*,\tau^N(u,f+u)>\epsilon_N^*\}\right].$$

Applying the Markov property at time $s-\epsilon_N^*$,

$$\begin{aligned}\Sigma_3^N &= \sum_u\frac{\alpha^d\xi^N_{s-\epsilon_N^*}(u)}{N}Q\sum_{e,f}q_N(-f,0,-e)\mathcal{P}^N_{\epsilon_N^*}\varphi(u)^2\,\mathbb{P}\left[\tau^N(u,e+u)\leqslant\epsilon_N^*,\tau^N(u,f+u)\geqslant\epsilon_N^*\right]\\
&= X^N_{s-\epsilon_N^*}\left(\mathcal{P}^N_{\epsilon_N^*}\varphi^2\right)Q\sum_{e,f}q_N(-f,0,-e)\mathbb{P}\left[\tau^N(0,e)\leqslant\epsilon_N^*,\tau^N(0,f)>\epsilon_N^*\right]\\
&= X^N_{s-\epsilon_N^*}\left(\mathcal{P}^N_{\epsilon_N^*}\varphi^2\right)Q\sum_{e,f}q_N(-f,0,-e)\mathbb{P}\left[\tau^N(0,-e)\leqslant\epsilon_N^*,\tau^N(0,-f)>\epsilon_N^*\right]\\
&= X^N_{s-\epsilon_N^*}\left(\mathcal{P}^N_{\epsilon_N^*}\varphi^2\right)Q\gamma_1^N\end{aligned}$$

It is this expression that we deal with.

Part Four: For the remainder of the proof let us write $\varphi^2(u)$ instead of $\varphi(u)^2$. The idea is that thanks to the spatial rescaling $P^N_{\epsilon_N^*}\left(\varphi^2\right)\approx\varphi^2$. We note

$$\begin{aligned}\left|\mathcal{P}^N_{\epsilon_N^*}\varphi^2(u)-\varphi^2(u)\right| &= \mathbb{E}\left[\varphi^2\left(B^{u,N}_{\epsilon_N^*}\right)-\varphi^2(u)\right]\\
&\leqslant||\varphi^2||_{\text{lip}}\mathbb{E}\left[\left|B^{0,N}_{\epsilon_N^*}\right|\right]\\
&\leqslant||\varphi^2||_{\text{lip}}\left(\mathbb{E}\left[\left|B^{0,N}_{\epsilon_N^*}\right|^2\right]\right)^{1/2}\\
&= ||\varphi^2||_{\text{lip}}\left(\mathbb{E}\left[\left|\frac{B^{0,1}_{N\epsilon_N^*}}{\sqrt{N}}\right|^2\right]\right)^{1/2}\\
&\leqslant C||\varphi^2||_{\text{lip}}\frac{1}{\sqrt{N}}(N\epsilon_N^*)^{1/2}\\
&= C||\varphi^2||_{\text{lip}}N^{-1/8}\end{aligned}$$

since recall we have $\epsilon_N^*=N^{-1/4}$. Note C does not depend on u. Therefore,

$$\begin{aligned}\left|\Sigma_3^N-\gamma_1^NQX^N_{s-\epsilon_N^*}(\varphi^2)\right| &= Q\gamma_1^N\frac{\alpha^d}{N}\sum_x\xi^N_{s-\epsilon_N^*}(u)\left|\mathcal{P}^N_{\epsilon_N^*}\varphi^2(u)-\varphi^2(u)\right|\\
&\leqslant C||\varphi^2||_{\text{lip}}X^N_{s-\epsilon_N^*}(1)N^{-1/4}\end{aligned}$$

Finally,

$$\left|X^N_{s-\epsilon_N^*}(\varphi^2)Q\gamma_1^N-X^N_{s-\epsilon_N^*}(\varphi^2)Q\gamma_1^N\right|\leqslant C||\varphi||_\infty^2\left|\gamma_1^N-\gamma_1\right|X^N_{s-\epsilon*N}(1).$$

Collecting all the error terms together completes the proof for $p=1$.

The $p=2$ case: The argument is the same strategy as the $p=1$ case and we will only give an outline. Let us recycle our notation $\mathcal{E}_1^N,\Sigma_1^N,\ldots$. First obtain

$$\mathbb{E}\left[m_s^{N,2}(\varphi)\,\middle|\,\mathcal{F}^N_{s-\epsilon_N^*}\right]$$
$$=\frac{\alpha^d}{N}\sum_{x,y,z}\varphi(x)\varphi(z)q_N(y,x,z)\mathbb{P}\left[B^{x,N}_{\epsilon_N^*}\notin\xi^N_{s-\epsilon_N^*},B^{z,N}_{\epsilon_N^*}\notin\xi^N_{s-\epsilon_N^*},B^{y,N}_{\epsilon_N^*}\in\xi^N_{s-\epsilon_N^*}\,\middle|\,\mathcal{F}^N_{s-\epsilon_N^*}\right].$$

We then change $\varphi(x)\varphi(z)$ to $\varphi(y)^2$. This is best done in two steps, and if one makes the first step the transition from $\varphi(x)\varphi(z)$ to $\varphi(x)\varphi(y)$ then the error term $\mathcal{E}_1^N$ is controlled essentially as before. Note that by $(\mathscr{C}2)$ (as in (4.4.5)), $\sum_z q_N(y,x,z) = \sum_n \#(\Phi_n)\mathbb{1}\{x \in \Phi_n^N(y)\}\phi_n$ and by $(\mathscr{C}1)$ $\#(\Phi_n) \leqslant C(L/\alpha)^d$ for some constant $C \in (0,\infty)$. Thus

$$\begin{aligned}
|\mathcal{E}_1^N| &\leqslant Q||\varphi||_\infty ||\varphi||_{\text{lip}} \frac{\alpha^d}{N} \sum_y \mathbb{P}\left[B^{y,N} \in \xi^N_{s-\epsilon_N^*} \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right] \sum_{x,z} q_N(y,x,z)|y-z| \\
&\leqslant Q||\varphi||_\infty ||\varphi||_{\text{lip}} \sum_{u,y} \frac{\alpha^d \xi^N_{s-\epsilon_N^*}(u)}{N} \mathbb{P}\left[B^{y,N} = u\right] \sum_z |y-z| C L^d p_N(y,x) \\
&\leqslant CL^d Q||\varphi||_\infty ||\varphi||_{\text{lip}} \sum_u \frac{\alpha^d \xi^N_{s-\epsilon_N^*}(u)}{N} \sum_y \mathbb{P}\left[B^{u,N} = y\right] \sum_w |w| C L^d p_N(w) \\
&\leqslant CL^d Q||\varphi||_\infty ||\varphi||_{\text{lip}} X^N_{s-\epsilon_N^*}(1) \frac{1}{\sqrt{N}} \mathbb{E}\left[|W|\right].
\end{aligned}$$

The error term $\mathcal{E}_2^N$ arising from changing $\varphi(x)\varphi(y)$ to $\varphi(y)^2$ can be controlled in exactly the same way. This leaves us dealing with

$$\Sigma_2^N = \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(y)^2 q_N(y,x,z) \mathbb{P}\left[B^{x,N}_{\epsilon_N^*} \notin \xi^N_{s-\epsilon_N^*}, B^{z,N}_{\epsilon_N^*} \notin \xi^N_{s-\epsilon_N^*}, B^{y,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*} \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right].$$

In the same way as in the case $p = 1$ we can transform this with two applications of Lemma 4.7.3 to

$$\begin{aligned}
\Sigma_4^N &= \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(y)^2 q_N(y,x,z) \mathbb{P}\left[\tau^N(y,x) > \epsilon_N^*, \tau^N(y,z) > \epsilon_N^*, B^{y,N}_{\epsilon_N^*} \in \xi^N_{s-\epsilon_N^*} \;\middle|\; \mathcal{F}^N_{s-\epsilon_N^*}\right] \\
&= \frac{\alpha^d}{N} \sum_{x,y,z} \varphi(y)^2 q_N(y,x,z) \sum_u \xi^N_{s-\epsilon_N^*}(u) \mathbb{P}\left[B^{y,N}_{\epsilon_N^*} = u, \tau^N(y,x) > \epsilon_N^*, \tau^N(y,z) > \epsilon_N^*\right]
\end{aligned}$$

We then note that

$$\mathbb{P}\left[B^{y,N}_{\epsilon_N^*} = u, \tau^N(y,x) > \epsilon_N^*, \tau^N(y,z) > \epsilon_N^*\right] = \mathbb{P}\left[B^{u,N}_{\epsilon_N^*} = y, \tau^N(u,u+e) > \epsilon_N^*, \tau^N(u,u+f) > \epsilon_N^*\right]$$
$$\mathbb{1}\{x,z \in \Phi_n^N(y)\} = \mathbb{1}\{-e,-f \in \Phi_n^N\}$$

where $e = y - x$ and $f = y - z$. Proceeding in the same fashion as the case $p = 1$ we reach

$$\begin{aligned}
\Sigma_4^N &= \frac{\alpha^d}{N} \sum_{u,y,e,f} \varphi(y)^2 Q q_N(-e,-f,0) \mathbb{P}\left[B^{u,N}_{\epsilon_N^*} = y, \tau^N(u,u+e) > \epsilon_N^*, \tau^N(u,u+f) > \epsilon_N^*\right] \\
&= \sum_u \frac{\alpha^d \xi^N_{s-\epsilon_N^*}(u)}{N} Q \sum_{e,f} q_N(-e,-f,0) \mathbb{E}\left[\varphi(B^{u,N}_{\epsilon_N^*})^2\right] \mathbb{P}\left[\tau^N(0,e) > \epsilon_N^*, \tau^N(0,f) > \epsilon_N^*\right] \\
&= \sum_u \frac{\alpha^d \xi^N_{s-\epsilon_N^*}(u)}{N} \mathcal{P}^N_{\epsilon_N^*} \varphi^2(u) Q \sum_{e,f} q_N(-e,-f,0) \mathbb{P}\left[\tau^N(0,-e) > \epsilon_N^*, \tau^N(0,-f) > \epsilon_N^*\right] \\
&= X^N_{s-\epsilon_N^*}\left(\mathcal{P}^N_{\epsilon_N^*} \varphi^2\right) Q \gamma_2^N
\end{aligned}$$

The remainder of the argument proceeds exactly as in part 4 of the $p = 1$ case. ∎

Combining our results from Sections 4.6 and 4.7 allows us to finally prove Lemma 4.5.2.

PROOF: [of Lemma 4.5.2] Recall $\langle M^N(\varphi)\rangle_s = \int_0^t m_s^{N,1}(\varphi) + m_s^{N,2}(\varphi)ds$. We use the decomposition

$$\begin{aligned}
&\left|\int_0^t m_s^{N,1}(\varphi) + m_s^{N,2}(\varphi)ds - Q(\gamma_1+\gamma_2)\int_0^t X_s^N(\varphi^2)ds\right| \qquad (4.7.3)\\
&\qquad \leqslant \sum_{p=1,2}\left(\int_0^{\epsilon_N^* \wedge t}\left|m_s^{N,p}(\varphi)\right|ds + \left|\int_{\epsilon_N^*}^{\epsilon_N^* \vee t} m_s^{N,p}(\varphi) - \mathbb{E}\left[m_s^{N,p}(\varphi)\;\middle|\;\mathcal{F}^N_{s-\epsilon_N^*}\right]ds\right|\right.\\
&\qquad \left. + \int_{\epsilon_N^*}^{\epsilon_N^* \vee t}\left|\mathbb{E}\left[m_s^{N,p}(\varphi)\;\middle|\;\mathcal{F}^N_{s-\epsilon_N^*}\right] - Q\gamma_p X^N_{s-\epsilon_N^*}(\varphi)\right|ds + Q\gamma_p\int_{(t-\epsilon_N^*)^+}^t X_s(\varphi^2)ds\right)
\end{aligned}$$

We estimate the first term (inside the sum) with Lemma 4.4.7 and the third with Lemma 4.7.5. The fourth we estimate with the bound $X_s^N(\varphi^2) \leqslant ||\varphi||_\infty^2 X_s^N(1)$. Taking expectations and estimating the second term using $\mathbb{E}[|Z|] \leqslant (\mathbb{E}[Z^2])^{1/2}$ and Lemma 4.7.1 gives us

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^t m_s^{N,1}(\varphi) + m_s^{N,2}(\varphi)ds - Q(\gamma_1+\gamma_2)\int_0^t X_s^N(\varphi^2)ds\right|\right]\\
&\qquad \leqslant \epsilon_N^* C_1||\varphi||_\infty^2\,\mathbb{E}\left[\sup_{s\leqslant t} X_s^N(1)\right] + \left(\epsilon_N^*(t\wedge\epsilon_N^*)C_2||\varphi||_\infty^2 t\left(1+X_0^N(1)^2\right)\right)^{1/2}\\
&\qquad + \left(\frac{1}{N^{1/4}} + \frac{1}{N^{1/2}} + N\sup_{\substack{u\neq v\\ w\neq 0}}\mathbb{P}\left[B^{u,N}_{\epsilon_N^*} - B^{v,N}_{\epsilon_N^*} = w\right] + \left|\gamma_p^N - \gamma_p\right|\right)tC_3||\varphi||\,\mathbb{E}\left[\sup_{s\in(\epsilon_N^*,\epsilon_N^*\vee t)} X^N_{s-\epsilon_N^*}(1)\right]\\
&\qquad + \epsilon_N^* C_4\,\mathbb{E}\left[\sup_{s\leqslant t} X_s^N(1)\right].
\end{aligned}$$

In the above, using Lemma 4.4.6 on the first, third and fourth terms, Lemma 4.6.5 and Lemma 4.7.2 on the third, and finally the fact that $\epsilon_N^* \to 0$ on the first, second and fourth, we obtain that the whole expression tends to zero as $N \to \infty$. ■

Chapter 5

Future work

We outline some ideas for future work.

5.1 Directly related to this thesis

5.1.1 Related to Chapter 3

- Address the question of uniqueness to the martingale problem for the SΛFV process (see Definition 3.1.5). In the absence of any obvious duality it is hard to gauge the difficulty of this problem. One approach might be to use a lookdown construction similar to Donnelly and Kurtz (1999a) in order to get duality; another idea would be to use a Dawson-Girsanov transform of the neutral process (although it seems unlikely this would cover all cases).
- Look for a way to understand the versions of the SΛFV process which cannot be characterized using Poisson point processes (see Remark 2.1.11).

5.1.2 Related to Chapter 4

- Remove the condition that size the reproduction events of the bursting process must be uniformly bounded above by L (part of $(\mathscr{C}1)$).
- Do the necessary adaptations to prove bursting processes also rescale to super-Brownian motion in dimension 2.

5.2 Further ideas

5.2.1 A deposition-type model

Let Λ be a Poisson point process with points $(t, x, r, p) \in [0, \infty) \times \mathbb{R}^d \times \{-1, 0, 1\}$ of rate $dt \otimes dx \otimes \mu(di) \otimes \nu(dn)$. One could study the process

$$H_t(x) = \sum_{(t,y,r,p)\in\Lambda} p\mathbb{1}\{s \leqslant t, x \in B_r(y)\}.$$

When $\int_0^\infty r^d \mu(dr) < \infty$ and $\nu \in \mathscr{P}\{-1, 0, 1\}$ this defines (almost everywhere) a function which can be thought of as a d-dimensional surface. If μ is an infinite measure and p is not a point mass the surface will be rough. For example if ν puts mass $1/2$ on both -1 and 1, for each $c \in \mathbb{Z}$ the set of $x \in \mathbb{R}^d$ such that $H_t(x) = c$ will be a totally disconnected set.

This process might be interesting to rescale, for example if we scale in such a way as $t \to H_t(x)$ becomes a Brownian motion then we obtain a Gaussian field.

5.2.2 On the hierachical group

In Dawson and Greven (1993) a model of interacting diffusions is considered with the geographical space as the hierarchical group (in place of the usual $\mathbb{R}^d$). We will not include references to this area in our own bibliography and refer the interested reader to Dawson et al. (1996) and Dawson and Greven (1999), amongst others.

The Hierarchical group is best thought of as a tree like structure which describes the spatial structure of an infinite collection of colonies of individuals. The interactions occurs on different spatial scales which are set up in such a way as on each spatial scale the interaction is of a particular asymptotic rate and (as the system size tends to infinity) can be isolated on a unique time-scale. Dawson and Greven (1993) obtained results about duality, clustering, the nature of the non-trivial equilibria, and a highly technical result concerning a fixed point of the rescaling procedure. Dawson and Greven (1993) worked with a system of interacting (via migration) Wright-Fisher diffusions. In the later papers (Dawson et al. (1996) and Dawson and Greven (1999)) it was shown that the system can be analysed in more detail using a general type space and Fleming-Viot processes in place of Wright-Fisher diffusions[1].

Most population models are restricted to dealing with local effects (i.e. selection, mutation) superimposed on a particle system which on its own would exhibit ancestral lineage duality. One advantage of the multi-scale approach is that one can analyse superimposed effects in which individuals interact over a large spatial scale, providing the interaction occurs in such a way as it depends only on (i) local things and (ii) the global average of types from the large spatial scale.

Therefore, it seems likely that one could use these tools to provide a mathematical formulation for individuals within competing populations cooperating. The idea one wishes to capture is that if, in some large region, there are disproportionately many type 1 individuals and very few type 2 individuals, the type 1s are able to kill off the type 2s at an increased rate. Let us give an example of what this means on a single spatial scale, in the style of Dawson and Greven (1993).

Fix $N \in \mathbb{N}$ and let $c, b \in [0, \infty)$. Let $t \to x_j^N(t)$ for $j \in \{1, \ldots, \mathbb{N}\}$ be a system of interacting diffusions defined by the equations

$$dx_j^N(t) = c\left[\overline{x}^N(t) - x_j^N(t)\right] + b\left[\overline{x}_j^N(t)(1 - x_j^N(t))\right] + \sqrt{x_j^N(t)(1 - x_j^N(t))}dW_j^N(t) \tag{5.2.1}$$

$$\overline{x}^N(t) = \frac{1}{N}\sum_{j=1}^{n} x_j^N(t)$$

where $(W_j^N)_{j,N}$ are independent Brownian motions. The interpretation is that at each site j we have individuals with type taken from $\{0, 1\}$, and $x_j^N(t)$ specifies the proportion (so $0 \leqslant x_j^N(t) \leqslant 1$) of type 1s at time t in site j.

The final term in (5.2.1) is Wright-Fisher noise. The first term of (5.2.1) corresponds to migration; individuals migrate out of the global average into site j at rate c. The middle term is intended to model cooperation. It has a high value (corresponding to a force increasing the number of type 1s at j) when $x_j^N(t)$ is low (i.e. when there are few type 1s at j) but the global average of type 1s is high.

In the limit as $N \to \infty$ the global average $\overline{x}^N(t)$ of type 1s is expected to converge (as a consequence of strong law of large number effects) to some constant $\theta \in [0, 1]$. Thus the dynamics of the limiting diffusions are expected to be

$$dx_j(t) = c(\theta - x_j(t)) + b\theta(1 - x_j(t)) + \sqrt{x_j(t)(1 - x_j(t))}dW_j(t).$$

where $W_j(t)$ are independent Brownian motions. Note that the processes at different sites have decoupled in the limit, leaving us with a system of (countably many) independent identically distributed diffusions.

It might also be interesting to analyse the system (5.2.1) where the type 0s are given a selective advantage (which introduces another term into (5.2.1)) to compensate their cooperative disadvantage.

5.2.3 Connections to the Brownian web

Again, we will not include references to this area in our own bibliography. The connection is that, suitably represented, the dual of the one dimensional SΛFV process rescales to the Brownian web. This

[1] Recall that the two type Fleming-Viot process has the Wright-Fisher diffusion describing the total mass of a single type

was essentially proved in Ferrari et al. (2005). However, connections between the SΛFV process and the Brownian web remains unexplored and it is possible that useful techniques could be transferred between the two.

Appendix A

The topology on Ξ

We go about placing a topology on Ξ and proving the results in Section 2.1.1. Let

$$C(K) = \{f : K \to \mathbb{R}\,;\, f \text{ is continuous}\}\,,$$

equipped with the supremum norm $||\cdot||_\infty$. Let

$$\mathcal{L}^1((\mathbb{R}^d)^n, C(K^n)) = \left\{\Phi : (\mathbb{R}^d)^n \to C(K^n)\,;\, \Phi \text{ is measurable and } \int_{(\mathbb{R}^d)^n} ||\Phi(z)||_\infty dz < \infty\right\}$$

and let

$$\Phi_1 \sim_2 \Phi_2 \text{ iff } \{z \in (\mathbb{R}^d)^n\,;\, \Phi_1(z) = \Phi_2(z)\} \text{ is Lebesgue null.}$$

Let $L^1((\mathbb{R}^d)^n, C(K^n))$ be the quotient of $\mathcal{L}^1((\mathbb{R}^d)^n, C(K^n))$ by $\sim_2$. Then $L^1((\mathbb{R}^d)^n, C(K^n))$ is a Banach space equipped with the norm $||\Phi|| = \int_{(\mathbb{R}^d)^n} ||\Phi(z)||_\infty dz$.

For ease of notation let us write

$$L^\infty = L^\infty(\mathbb{R}^d, \mathcal{M}_F(K)) \text{ and } L^1[n] = L^1((\mathbb{R}^d)^n, C(K^n)).$$

Write also $L^1 = L^1[1]$.

Proposition A.1 (Evans (1998)) *L^∞ is isometrically isomorphic to a closed subspace of the dual of L^1 via the action*

$$(\rho, \Phi) = \int_{\mathbb{R}^d} \langle \Phi(x), \rho(x)\rangle\, dx,$$

where $\langle \Phi(x), \rho(x)\rangle = \int_K (\Phi(x)(k))\ (\rho(x)(dk))$.

We induce the weak-star topology on $\Xi \subseteq L^\infty$ from this identification.

Remark A.2 *It is in general not true that L^∞ is isomorphic to the whole dual of L^1. By Theorem IV.1.1 of Diestel and Uhl (1977), isomorphism holds if and only if $\mathcal{M}_F(K)$ has the Radon-Nikodym property with respect to $(\mathbb{R}^d, \mathcal{L})$ (see the remarks following Definition III.1.3 of (Diestel and Uhl, 1977)). This fails, for example, if $K = \{0,1\}^{\mathbb{N}}$ (with the usual σ-algebra generated by cylinder sets).*

From the separability of L^1 and the Banach-Alaoghu theorem we obtain

Proposition A.3 (Evans (1998)) *Ξ is a compact metrizable space.*

Let $C(\Xi)$ denote the continuous functions from $\Xi \to \mathbb{R}$ equipped with supremum norm $||\cdot||_\infty$. Then $C(\Xi)$ is a Banach space. An application of the Stone-Weierstrass theorem identifies a suitable class of test functions:

Proposition A.4 (Evans (1998)) *$\mathscr{I}$ is dense in $C(\Xi)$.*

PROOF: Evans proves this without the restriction that for Ψ of the form $(\star, n)$, ψ has compact support. The modification required to prove our stated result is minimal. ■

Proposition A.5 *$\mathscr{I}$ is a seperating, convergence determining class of Ξ.*

PROOF: Evans proves that $\mathscr{I}$ separates points of Ξ. By Lemma 4.3 of Ethier and Kurtz (1986) the compactness of Ξ implies that $M \subseteq C(\Xi)$ is separating iff M is convergence determining. ■

Proposition A.6 *If $\rho^m, \rho \in \Xi$ and for all Ψ of the form $(1, \star)$ (in other words, $\psi : \mathbb{R}^d \to \mathbb{R}$ continuous with compact support and $\chi \in C(K)$) we have*

$$\int_{\mathbb{R}^d} \psi(x) \langle \chi, \rho^m(x) \rangle \, dx \to \int_{\mathbb{R}^d} \psi(x) \langle \chi, \rho(x) \rangle \, dx$$

as $m \to \infty$, then $\rho^m \to \rho$ in Ξ.

PROOF: We write $||\cdot||$ for the norm on L^1. Lemma A.2 of Evans (1998) shows that functions of the form $x \to \psi(x)\chi(\cdot)$ where $\psi : \mathbb{R}^d \to \mathbb{R}$ is integrable and $\chi \in C(K)$ are dense in L^1. A straightforward modification of the argument allows us to assume that ψ is continuous with compact support.

Let $\widetilde{\Phi} \in L^1$. Let $\epsilon > 0$ and and choose ψ, χ as above so as $||\widetilde{\Phi}(\cdot) - \psi(\cdot)\chi|| < \epsilon/3$. By our assumptions we may choose M such that for all $m > M$, $|\int_{\mathbb{R}^d} \psi(x) \langle \chi, \rho^m(x) \rangle \, dx - \int_{\mathbb{R}^d} \psi(x) \langle \chi, \rho(x) \rangle \, dx| < \epsilon/3$. A straightforward use of the triangle inequality now shows that for all $m > M$

$$\left| \int_{\mathbb{R}^d} \left\langle \widetilde{\Phi}(x), \rho^m(x) \right\rangle dx - \int_{\mathbb{R}^d} \left\langle \widetilde{\Phi}(x), \rho(x) \right\rangle dx \right| < \epsilon.$$

Thus, in the terminology of Proposition A.1, $(\rho^m, \widetilde{\Psi}) \to (\rho, \widetilde{\Psi})$ for all $\widetilde{\Psi} \in L^1$ and we have the result. ■

We have now proved all the statements of Proposition 2.1.7.

Appendix B

$\mathcal{I}$-Measurability

Lemma B.1 *The set I_0^α is a measurable subset of I. The function*

$$(x, i) \to \mathbb{1}\{x \in E_i\}$$

defined on $\mathbb{R}^d \times I$ is $(\mathbb{B}^d \otimes \mathcal{I})/\mathbb{B}$ measurable. The functions

$$i \to \alpha^d \#(\alpha\mathbb{Z}^d \cap E_i) \qquad i \to \mathcal{L}(E_i) \qquad i \to \mathscr{D}(i)$$

are $\mathcal{I}/\mathbb{B}^d$ measurable.

Proof:

Since E_i is Borel the function $x, \to \mathbb{1}\{x \in E_i\}$ is measurable. By ($\mathscr{H}2$), for each $x \in \mathbb{R}^d$ the function $i \to \mathbb{1}\{x \in \mathbb{E}_i\}$ is measurable. Measurability of $(x, i) \to \mathbb{1}\{x, \in E_i\}$ now follows (by Lemma 8.1(d) of Williams (1991), for example).

Let us note also that

$$\alpha^d \#(E_i \cap \alpha\mathbb{Z}^d) = \alpha^d \sum_{y \in \alpha\mathbb{Z}^d} \mathbb{1}\{y \in E_i\} = \int_{\mathbb{R}^d} \mathbb{1}\{[y]_\alpha \in E_i\} dy$$

and by ($\mathscr{H}2$) this is a measurable function of i. By dominated convergence (domination is easily achieved since by ($\mathscr{H}1$) E_i is bounded),

$$\int_{\mathbb{R}^d} \mathbb{1}\{[y]_\alpha \in E_i\} dy \to \int_{\mathbb{R}^d} \mathbb{1}\{y \in E_i\} dy = \mathcal{L}(E_i).$$

This expresses $i \to \mathcal{L}(E_i)$ as a pointwise limit of $\mathcal{I}/\mathbb{B}^d$ measurable functions, which implies $i \to \mathcal{L}(E_i)$ is measurable.

Note that

$$\mathscr{D}(i) = \lim_{\alpha \downarrow 0} \sup_{x \in \alpha\mathbb{Z}^d} |x| \mathbb{1}\{x \in E_i\}$$

because E_i is open. Thus, again by the algebra of measurable functions, $i \to \mathscr{D}_i$ is measurable. It follows from this that $I_0^\alpha \in \mathcal{I}$. ■

Appendix C

Proof of Lemma 3.2.4

Lemma C.1 *For each $\alpha \in (0,1]$,*

$$\int_I \mu^\alpha(di) \int_0^1 \nu_i(du) u \mathscr{D}(i)^{2d} < \infty.$$

PROOF: In fact we will prove something a bit stronger. Let $0 < \beta_1 < \beta_2 < \infty$. Note

$$\begin{aligned}\int_I \mu(di) \int_0^1 \nu_i(du)\, u\beta_1^d \mathbb{1}\{\beta_1 \leqslant \mathscr{D}(1) \leqslant \beta_2\} &\leqslant \int_I \mu(di) \int_0^1 \nu_i(du)\, u\mathscr{D}(i)^d \mathbb{1}\{\beta_1 \leqslant \mathscr{D}(i) \leqslant \beta_2\} \\ &\leqslant \int_I \mu(di) \int_0^1 \nu_i(du)\, u\mathscr{D}(i)^d. \qquad \text{(C.1)}\end{aligned}$$

By ($\mathscr{H}3$),

$$\int_I \mu(di) \int_0^1 \nu_i(du)\, u \mathbb{1}\{\beta_1 \leqslant \mathscr{D}(1) \leqslant \beta_2\} < \infty.$$

The stated result follows by taking $\beta_1 = \alpha, \beta_2 = 1/\alpha$ and noting that $\alpha \leqslant \mathscr{D}(i) \leqslant 1/\alpha$ if $i \notin I_0^\alpha$. ■

PROOF: [Of Lemma 3.2.4] We seek to use Theorem 3.9 of Ligget (1985), from which the stated results will follow, and to do so we must check equations (3.3) and (3.8) of Ligget (1985). There are some other minor conditions required to check that our setting really satisfies the setup of Ligget (1985) but we will omit those details. In order to check (3.3) and (3.8) of Ligget (1985) we must express our system in the notation of Ligget (1985). Let

$$\mathscr{H} = \big\{H \subseteq \alpha\mathbb{Z}^d\,;\, H \neq \varnothing \text{ and } \exists i \in I, x \in \alpha\mathbb{Z}^d \text{ such that } H \subseteq E_i(x)\big\}$$

be the set of all possible combinations of sites that a single reproduction event could overwrite. Let us write

$$\Phi_i^\alpha(y) = \alpha\mathbb{Z}^d \cap E_i(y).$$

For $A \subseteq K^H$ set

$$\begin{aligned}c_H(\zeta, A) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \\ \left\{\mathbb{1}\big\{H \subseteq \Phi_i^\alpha([y]_\alpha)\big\}\left(\prod_{x \in \Phi_i^\alpha([y]_\alpha)\setminus H} \hat{T}(x;)\right)\left(\prod_{x\in H} T(x;)\right)\right. \\ \left.\times \prod_{x\in H}\Big[T(x;)\mathbb{1}\{\exists \eta \in A, \eta(x) = \zeta(y)\} + \hat{T}(x;)\mathbb{1}\{\exists \eta \in A, \eta(x) = \zeta(x)\}\Big]\right\}\end{aligned}$$

which is the rate at which the sites in H change type to a state $\eta \in A$ if the process is in state ζ. Note that if $A = K^H$ the bottom line of the above expression is just 1.

Remark C.2 *We do not 'double count' events; The term*

$$\mathbb{1}\{H \subseteq E_i([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H}\hat{T}(x;)\right)\left(\prod_{x\in H}T(x)\right)$$

checks that $H \subseteq \alpha\mathbb{Z}^d \cap E_i([y]_\alpha)$ and the sites which are overwritten by the event (y, i, u, f, g) are precisely H. But note that we do not care whether or not the type of the overwritten site was 'changed' to the type it already had. This is in keeping with the notation of Ligget (1985).

We note the slightly shorter form,

$$c_H(\zeta, A) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg)$$
$$\left\{\mathbb{1}\{H \subseteq \Phi_i^\alpha([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H}\hat{T}(x;)\right)\left(\prod_{x\in H}T(x;)\mathbb{1}\{\exists\eta\in A, \eta(x)=\zeta(y)\}\right)\right\}.$$

Equation (3.3) of Ligget (1985) is the statement that a single site must only change type at finite rate. Since the action of Definition 3.2.2 is spatially homogeneous we need only check the following.

$$\begin{aligned}
&\sum_{\{H\in\mathscr{H}\,;\,H\ni 0\}} \sup\left\{c_H\left(\zeta, K^H\right)\,;\,\zeta\in K^{\alpha\mathbb{Z}^d}\right\} \\
&\qquad\leqslant \sum_{\{H\in\mathscr{H}\,;\,H\ni 0\}} \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}} U(df) \\
&\qquad\qquad \times \mathbb{1}\{H \subseteq \Phi_i^\alpha([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in H}\mathbb{1}\{u>f(x)\}\right) \\
&\qquad= \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}} U(df)\,\mathbb{1}\{0\subseteq E_i([y]_\alpha)\} \\
&\qquad\qquad \times \sum_{\{H\in\mathscr{H}\,;\,H\ni 0\}} \mathbb{1}\{H \subseteq \Phi_i^\alpha([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in H}\mathbb{1}\{u>f(x)\}\right) \\
&\qquad= \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}} U(df)\,\mathbb{1}\{0\subseteq \Phi_i^\alpha([y]_\alpha)\} \\
&\qquad\qquad \times \sum_{G\subseteq\Phi_i^\alpha([y]_\alpha)\setminus\{0\}} \mathbb{1}\{u\geqslant f(0)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus(G\cup\{0\})}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in G}\mathbb{1}\{u>f(x)\}\right)
\end{aligned} \tag{C.2}$$

We note that by Lemma 3.2.1,

$$\begin{aligned}
&\sum_{G\subseteq\Phi_i^\alpha([y]_\alpha)\setminus\{0\}} \int_{\mathcal{U}} U(df)\,\mathbb{1}\{u\geqslant f(0)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus(G\cap\{0\})}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in G}\mathbb{1}\{u>f(x)\}\right) \\
&\quad= u \sum_{G\subseteq\Phi_i^\alpha([y]_\alpha)\setminus\{0\}} u^{\#(G)}(1-u)^{\#(\Phi_i^\alpha([y]_\alpha)\setminus(G\cap\{0\}))} \\
&\quad= u \sum_{m=0}^{\#(\Phi_i^\alpha([y]_\alpha))-1} \sum_{\substack{G\subseteq\Phi_i^\alpha([y]_\alpha)\setminus\{0\} \\ \#(G)=m}} u^{\#(G)}(1-u)^{\#(\Phi_i^\alpha([y]_\alpha))-1-\#(G)} \\
&\quad= u \sum_{m=0}^{\#(\Phi_i^\alpha([y]_\alpha))-1} \binom{\#(\Phi_i^\alpha([y]_\alpha))-1}{m} u^m(1-u)^{[\#(\Phi_i^\alpha([y]_\alpha))-1]-m} \\
&\quad= u.
\end{aligned} \tag{C.3}$$

Putting (C.3) into (C.2) gives

$$\begin{aligned}\sum_{\{H\in\mathscr{H}\,;\,H\ni 0\}}\sup\left\{c_H\left(\zeta,K^H\right)\,;\,\zeta\in K^{\alpha\mathbb{Z}^d}\right\} &\leqslant \int_{\mathbb{R}^d}dy\int_I\mu^\alpha(di)\int_0^1\nu_i(du)\,\mathbb{1}\{0\subseteq\Phi_i^\alpha([y]_\alpha)u\\ &\leqslant \int_I\mu^\alpha(di)\int_0^1\nu_i(du)u\int_{\mathbb{R}^d}dy\,\mathbb{1}\{|y|\leqslant\mathscr{D}(i)+\alpha\}\\ &= \int_I\mu(di)\int_0^1\nu_i(du)u\int_{\mathbb{R}^d}dy\,\mathbb{1}\{|y|\leqslant\mathscr{D}(i)+\alpha, i\notin I_0^\alpha\}\\ &\leqslant \int_I\mu(di)\int_0^1\nu_i(du)u\int_{\mathbb{R}^d}dy\,\mathbb{1}\{|y|\leqslant 2\mathscr{D}(i)\}\\ &= C\int_I\mu(di)\int_0^1\nu_i(du)\,u\mathscr{D}(i)^d<\infty.\end{aligned}$$

where we used that $i\in I\backslash I_0^\alpha$ implies $\alpha\leqslant\mathscr{D}(i)$ to get from the third to fourth lines. Finiteness in the above follows from ($\mathscr{H}3$) and thus (3.3) of Ligget (1985) holds. Note that (up to a constant) this is precisely the bound we would expect from the dynamics of Definition 3.2.2.

We now approach (3.8) of Ligget (1985), which is a condition to the effect that long range dependence of the reproduction mechanism is small. Again, since the action of Definition 3.2.2 is spatially homogeneous we need only check that

$$\sum_{\{H\in\mathscr{H}\,;\,0\in H\}}\sum_{w\in\alpha\mathbb{Z}^d\backslash\{0\}}\sup\left\{||C_H(\zeta,\cdot)-C_H(\beta,\cdot)||_{TV}\,;\,\zeta(u)=\beta(u)\text{ for all }u\in\alpha\mathbb{Z}^d\backslash\{w\}\right\}\tag{C.4}$$

is finite. We begin with the estimate that if $\zeta(u)=\beta(u)$ for all $u\neq w$,

$$\begin{aligned}&|C_H(\zeta,A)-C_H(\beta,A)|\qquad\qquad(C.5)\\ &\quad\leqslant\int_{\mathbb{R}^d}dy\int_I\mu^\alpha(di)\int_0^1\nu_i(di)\int_{\mathcal{U}}U(df)\\ &\qquad\mathbb{1}\{H\subseteq\Phi_i^\alpha([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\backslash H}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in H}\mathbb{1}\{u>f(x)\}\right)\\ &\qquad\quad\times\left(\prod_{x\in H}\mathbb{1}\{\exists\eta\in A,\eta(x)=\zeta(y))\}-\prod_{x\in H}\mathbb{1}\{\exists\eta\in A,\eta(x)=\beta(y))\}\right)\\ &\quad\leqslant\int_{\mathbb{R}^d}dy\int_I\mu^\alpha(di)\int_0^1\nu_i(di)\int_{\mathcal{U}}U(df)\\ &\qquad\mathbb{1}\{H\subseteq\Phi_i^\alpha([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\backslash H}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in H}\mathbb{1}\{u>f(x)\}\right)\\ &\qquad\quad\times 2\,\mathbb{1}\{w\in H\}\\ &\quad\leqslant 2\int_{\mathbb{R}^d}dy\int_I\mu^\alpha(di)\int_0^1\nu_i(di)\int_{\mathcal{U}}U(df)\Bigg\{\mathbb{1}\{0,w\in\Phi_i^\alpha([y]_\alpha)\}\\ &\qquad\mathbb{1}\{H\subseteq\Phi_i^\alpha([y]_\alpha)\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\backslash H}\mathbb{1}\{u\leqslant f(x)\}\right)\left(\prod_{x\in H}\mathbb{1}\{u>f(x)\}\right).\qquad(C.6)\end{aligned}$$

Noting that the final line above does not depend on A we have

$$\sup\left\{||C_H(\zeta,\cdot)-C_H(\beta,\cdot)||_{TV}\,;\,\zeta(u)=\beta(u)\text{ for all }u\in\alpha\mathbb{Z}^d\backslash\{w\}\right\}\leqslant(C.6)$$

and putting this into (C.4) gives

$$
\begin{aligned}
(\text{C.4}) &\leqslant 2 \sum_{w\in\alpha\mathbb{Z}^d\setminus\{0\}} \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(di) \int_{\mathcal{U}} U(df) \Bigg\{ \mathbb{1}\{0, w \in \Phi_i^\alpha([y]_\alpha)\} \\
&\qquad \sum_{\{H\in\mathscr{H}\,;\,0\in H\}} \mathbb{1}\{H \subseteq \Phi_i^\alpha([y]_\alpha)\} \left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H} \mathbb{1}\{u \leqslant f(x)\} \right) \left(\prod_{x\in H} \mathbb{1}\{u > f(x)\} \right) \Bigg\} \\
&= 2 \sum_{w\in\alpha\mathbb{Z}^d\setminus\{0\}} \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(di)\, \mathbb{1}\{0, w \in \Phi_i^\alpha([y]_\alpha)\} \int_{\mathcal{U}} U(df) \\
&\qquad \sum_{G\subseteq\Phi_i^\alpha([y]_\alpha)\setminus\{0\}} \mathbb{1}\{u \geqslant f(0)\} \left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus(G\cup\{0\})} \mathbb{1}\{u \leqslant f(x)\} \right) \left(\prod_{x\in G} \mathbb{1}\{u > f(x)\} \right) \\
&= 2 \int_I \mu^\alpha(di) \int_0^1 \nu_i(di)\, u \int_{\mathbb{R}^d} dy\, \mathbb{1}\{0 \in \Phi_i^\alpha([y]_\alpha) \sum_{w\in\alpha\mathbb{Z}^d\setminus\{0\}} \mathbb{1}\{w \in \Phi_i^\alpha([y]_\alpha,)\} \\
&= 2 \int_I \mu^\alpha(di) \int_0^1 \nu_i(di)\, u \int_{\mathbb{R}^d} dy\, \mathbb{1}\{-[y]_\alpha \in \Phi_i^\alpha(0) \int_{\mathbb{R}^d} dz \mathbb{1}\{[z]_\alpha \in \Phi_i^\alpha([y]_\alpha)\} \\
&\leqslant 2 \int_I \mu^\alpha(di) \int_0^1 \nu_i(di)\, u \left(\int_{\mathbb{R}^d} dy \mathbb{1}\{|y| \leqslant \mathscr{D}(i) + \alpha, i \notin I_0^\alpha\} \right)^2 \\
&\leqslant 2 \int_I \mu^\alpha(di) \int_0^1 \nu_i(di)\, u \left(\int_{\mathbb{R}^d} dy \mathbb{1}\{|y| \leqslant 2\mathscr{D}(i)\} \right)^2 \\
&\leqslant C \int_I \mu^\alpha(di) \int_0^1 \nu_i(di)\, u \mathscr{D}(i)^{2d}
\end{aligned}
$$

where we used (C.3) to get from the second to third lines and then the fact that $i \in I^\alpha\backslash 0$ implies $\alpha \leqslant \mathscr{D}(i)$ to get from the fourth to the fifth. This is finite by Lemma C.1.

Thus we may use Theorem 3.9 of Ligget (1985). In particular the system of Definition 3.2.2 is well defined by its dynamics and has a Markov pre-generator Ω^α whose closure $\overline{\Omega}$ is a Markov generator. All that remains to do is show that the pre-generator Ω^α matches our expression in the statement of the Lemma. From Proposition 3.2(a) of Ligget (1985), the pre-generator for $J \in \Delta^\alpha$ as

$$
\Omega^\alpha J(\eta) = \sum_{H\in\mathscr{H}} \int_{K^H} c_H(\eta, d\beta) \Big[J(\eta^\beta) - J(\eta) \Big]
$$

where $\eta^\beta(x) = \eta(x)$ for $x \notin H$ and $\eta^\beta(x) = \beta(x)$ for $x \in H$. Considering the term relating to $J(\eta)$,

$$
\begin{aligned}
\sum_{H\in\mathscr{H}} \int_{K^H} c_H(\eta, d\beta) J(\eta) &= \sum_{H\in\mathscr{H}} J(\eta) c_H(\eta, K^H) \\
&= J(\eta) \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \\
&\qquad \sum_{H\in\mathscr{H}} \mathbb{1}\big\{H \subseteq \Phi_i^\alpha([y]_\alpha)\big\} \left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H} \hat{T}(x;) \right) \left(\prod_{x\in H} T(x;) \right) \\
&= J(\eta) \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg).
\end{aligned}
$$

The last line of the above follows since for each (y, i, u, f, g) there is precisely one $H \in \mathscr{H}$ such that

$$
\mathbb{1}\big\{H \subseteq \Phi_i^\alpha([y]_\alpha)\big\} \left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H} \hat{T}(x;) \right) \left(\prod_{x\in H} T(x;) \right) = 1,
$$

namely $H = \{x \in \Phi_i^\alpha([y]_\alpha)\,;\, T(x;) = 1\}$. Now the term relating to $J(\eta^\beta)$. Writing out and swapping some of the integrals gives us

$$\begin{aligned}
&\sum_{H\in\mathscr{H}} \int_{K^H} c_H(\eta, d\beta) J(\eta^\beta) \\
&\quad = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) \sum_{H\in\mathscr{H}} \mathbb{1}\big\{H \subseteq \Phi_i^\alpha([y]_\alpha)\big\} \\
&\qquad \left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H} \hat{T}(x;)\right)\left(\prod_{x\in H} T(x;)\right) \int_{K^H} \left\{ \left(\prod_{x\in H} \mathbb{1}\{\exists \eta \in d\beta, \eta(x) = \zeta(y)\}\right)\right\} J(\eta^\beta) \qquad \text{(C.7)}
\end{aligned}$$

As before, with (y, i, u, f, g) fixed there is precisely one $H \in \mathscr{H}$ such that

$$\mathbb{1}\big\{H \subseteq \Phi_i^\alpha([y]_\alpha)\big\}\left(\prod_{x\in\Phi_i^\alpha([y]_\alpha)\setminus H} \hat{T}(x;)\right)\left(\prod_{x\in H} T(x;)\right) = 1,$$

namely $H = \{x \in \Phi_i^\alpha([y]_\alpha)\,;\, T(x;) = 1\}$. For this H,

$$\int_{K^H} \left\{ \left(\prod_{x\in H} \mathbb{1}\{\exists \eta \in d\beta, \eta(x) = \zeta(y)\}\right)\right\} J(\eta^\beta) = J(\eta^{y,i,u,f,g})$$

since the measure we are integrating with becomes a point mass on $\beta \in K^H$ where $\beta(x) = \eta(y)$. Putting both these observations into (C.7) gives

$$\sum_{H\in\mathscr{H}} \int_{K^H} c_H(\eta, d\beta) J(\eta^\beta) = \int_{\mathbb{R}^d} dy \int_I \mu^\alpha(di) \int_0^1 \nu_i(du) \int_{\mathcal{U}^2} U(df)U(dg) J(\eta^{y,i,u,f,g})$$

which completes the proof. ■

Appendix D

The Aldous-Rebolledo criterion

We give a statement of a very useful tightness criterion which is not commonly found in the literature. Let A be a complete separable metric space. Recall that $C(A)$ denotes the space of real valued continuous functions of A.

Theorem D.1 (Theorem 9.1, Ethier and Kurtz (1986)) *Let $\{Y^m\,;\, m \in \mathbb{N}\}$ be a sequence of càdlàg A valued processes. Suppose the sequence Y^m satisfies the compact containment condition in A; for each $T < \infty$ and $\epsilon > 0$ there exists a compact set $\Gamma_{\epsilon,T}$ of A such that*

$$\sup_n \mathbb{P}\big[\textit{for all } t \leqslant T,\, Y^n_\cdot \notin \Gamma_{\epsilon,T}\big] < \epsilon.$$

Let Θ be a dense subset of $C(A)$ in the topology of uniform convergence on compact sets. Then $\{Y^m_\cdot\,;\, m \in \mathbb{N}\}$ is relatively compact if and only if for each $f \in \Theta$, $\{f(Y^m_\cdot)\,;\, m \in \mathbb{N}\}$ is relatively compact as a set of processes in $D_{\mathbb{R}}[0,\infty)$.

Theorem D.2 (Aldous-Rebolledo Criterion, Rebolledo (1980)) *Let $\{Y^m_\cdot\,;\, m \in \mathbb{N}\}$ be a sequence of real valued processes. Then $\{Y^m_\cdot\}$ is tight in $D_{\mathcal{M}_F(\mathbb{R})}[0,\infty)$ if the following conditions are satisfied.*

(1) For each fixed $t \in [0,\infty)$, $\{Y^m_t\}_{m\in\mathbb{N}}$ is tight (as a sequence of real valued random variables).

(2) Let $\epsilon > 0$ and $T < \infty$, and for each m let $\tau_m \leqslant T$ be a stopping time with respect to the filtration of Y^m. Then there exists $\theta > 0$ and $m_0 < \infty$ such that

$$\sup_{m \geqslant m_0} \sup_{\delta \in [0,\theta]} \mathbb{P}\left[|Y^n_{\tau_m+\delta} - Y^n_{\tau_m}| > \epsilon\right] < \epsilon.$$

Remark D.3 *There is a useful specialization of the criterion for the case where Y^m is a semimartingale. If $Y^m = M^m + A^m$ is a decomposition of the semimartingale Y^m, where M^m is a local martingale and A^m has locally finite variation, then it suffices to check (2) for both $(\langle M^m\rangle)_{m\in\mathbb{N}}$ and $(A^m)_{m\in\mathbb{N}}$, where $\langle M^m\rangle$ is the bracket process of M^m.*

Bibliography

N. H. Barton, A. M. Etheridge, and J. Kelleher. A new model for extinction and recolonisation in two dimensions: quantifying phylogeography. *Evolution*, In Press, 2010a.

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

N. Berestycki, A.M. Etheridge, and M. Hutzenthaler. Survival, extinction and ergodicity in a spatially continuous population model. *Markov Process. Rel. Fields*, 15(3):265–288, 2009.

N. Berestycki, A. Etheridge, and A. Véber. Scaling limits of the spatial lambda-fleming viot process. *In Preparation*, 2010.

J. Bertoin and J. F. Le Gall. Stochastic flows associated to coalescent processes. *Probab. Th. Rel. Fields*, 126:261–288, 2003.

M. Birkner, J. Blath, M. Möhle, M. Steinrücken, and J. Tams. A modified lookdown construction for the xi-fleming-viot process with mutation and populations with recurrent bottleneck. *Alea*, 6:25–61, 2009.

D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Probab.*, 1:19–42, 1973.

J. T. Cox and A. Klenke. Rescaled interacting diffusions converge to super-brownian motion. *Ann. Appl. Probab.*, 13:501–514, 2003.

J. T. Cox and E. A. Perkins. Renormalization of the two-dimensional lotka-volterra model. *Ann. Appl. Probab.*, 18(2):747–812, 2008.

J. T. Cox, R. Durret, and E. A. Perkins. Rescaled voter models converge to super-brownian motion. *Ann. Probab.*, 28(1):185–234, 2000.

J. T. Cox, M. Merle, and E. A. Perkins. Co-existence in a two-dimensional lotka-volterra model. *Electron. J. Probab.*, 15:1190–1266, 2010.

D. A. Dawson and A. Greven. Multiple time scale analysis of interacting diffusions. *Probab. Th. Rel. Fields*, 95:467–508, 1993.

D. A. Dawson and A. Greven. Hierarchically interacting fleming-viot processes with selection and mutation: Multiple space time scale analysis and quasi-equilibria. *Electron. J. Probab.*, 4:1–81, 1999.

D. A. Dawson and K. H. Hochberg. Wandering random measures in the fleming-viot model. *Ann. Probab.*, 10(3):554–580, 1982.

D. A. Dawson and T. G. Kurtz. Applications of duality to measure-valued diffusions,. *Springer Lecture Notes in Control and Inf. Sci.*, 42:177–191, 1982.

D. A. Dawson, A. Greven, and J. Vaillancourt. Equilibria and quasi-equilibria for infinite systems of fleming-viot processes. *Math. Society*, 347(7):2277–2360, 1996.

J. Diestel and J. J. Uhl. *Vector Measures*. Providence, 1977.

P. Donnelly and T. G. Kurtz. A countable representation of the fleming-viot measure valued diffusion. *Ann. Probab.*, 24(2):698–742, 1996.

P. Donnelly and T. G. Kurtz. Genealogical processes for fleming-viot models with selection and recombination. *Ann. Appl. Probab.*, 9(4):1091–1148, 1999a.

P. Donnelly and T. G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.*, 27(1):166–205, 1999b.

R. Durret. *Ten Lectures on Particle Systems (in St. Flour Lecture Notes, Lecture Notes in Math 1608).* Springer, New York, 1995.

R. Durret and E. A. Perkins. Rescaled contact processes converge to super-brownian motion in two or more dimensions. *Probab. Th. Rel. Fields*, 114:309–399, 1999.

A. M. Etheridge. *An Introduction To Superprocesses.* American Mathematical Society, 2000.

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

A. M. Etheridge and T. G. Kurtz. Personal communication. In Preparation.

A. M. Etheridge, R. C. Griffiths, and J. E. Taylor. A coalescent dual process in a moran model with genic selection, and the lambda coalescent limit. *Theoretical Population Biology*, 78(2):77–92, 2010.

S. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, 1986.

S. Ethier and T. G. Kurtz. Fleming-viot processes in population genetics. *SIAM Journal on Control and Optimization*, 31(2):345–386, 1993.

S. N. Evans. Coalescing markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincare*, B33:339–358, 1998.

P. A. Ferrari, L. R. G. Fontes, and X. Y. Wu. Two-dimensional poisson trees converge to the brownian web. *Ann. Inst. H. Poincare Probab. Statis.*, 41(5):851–858, 2005.

W. H. Fleming and M. Viot. Some measure-valued markov processes in population genetics theory. *Indiana University Mathematics Journal*, 28(5):817–843, 1979.

T. E. Harris. Additive set-valued markov process and graphical methods. *Ann. Probab.*, 3:643–663, 1975.

O. Kallenberg. Canonical representations and convergence criteria for processes with interchangeable increments. *Wahrsch. Verw. Gebiete*, 27:33–36, 1973.

M. Kimura. Stepping stone model of population. *Ann. Rept. Nat. Inst. Genetics, Japan*, 3:62–63, 1953.

J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.

S. M. Krone and C. Neuhauser. Ancestral processes with selection. *Theoretical Population Biology*, 51: 210–237, 1997.

T. M. Ligget. *Interacting Particle Systems.* Springer, 1985.

L. Limic, A. Véber, and A. Wakolbinger. New constructions of the spatial lambda-fleming-viot process. *In Preparation*, 2010.

P. Moran. Random processes in genetics. *Proc. Cambridge Philos. Soc.*, 54:60–71, 1958.

C. Mueller and R. Tribe. Stochastic pdes arising from the long range contact process and long range voter model. *Probab. Th. Rel. Fields*, 102:519–546, 1995.

J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

R. Rebolledo. Sur l'existence de solutions a certains problegravemes de semimartin-gales. *C.R. Acad. Sci., Paris*, 290:843–846, 1980.

S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability*, 36:1116–1125, 1999.

R. van der Hofstad and G. Slade. Convergence of critical oriented percolation to super-brownian motion above 4+1 dimensions. *Ann. Inst. H. Poincare Probab. Statis.*, 39:413–485, 2003.

D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.