Forwards in time approaches to understanding genealogies in Λ-Fleming-Viot type models



Bastian Wiederhold

Wadham College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

Declaration

I hereby declare that, expect where specific reference is made to the work of others the contents of this dissertation are original.

Chapter 2 contains a joint work with Julian Kern. By the time of my joining of the project, Julian Kern had already envisioned the model and first results such as the convergence to the Wright-Fisher diffusion. I contributed in particular to the sections involving the lookdown constructions and writing the manuscript in its current form.

Chapter 3 contains a joint work with Raphaël Forien. I wrote the majority of the document, whereas Raphaël Forien provided crucial ideas regarding the precise scaling regimes as well as more complex calculations and possible simplifications.

Chapter 4 contains another joint work with Raphaël Forien, generalizing the ideas of Chapter 3 to varying population size. I wrote most of the manuscript. Raphaël Forien contributed crucial elements of the proof strategy.

As this thesis is in an integrated format, additional statements of authorship can be found at the end of Chapter 2, Chapter 3 and Chapter 4.

Julian Kern will include the joint work in his submission for his doctoral degree at Technische Universität Berlin. Raphaël Forien will include a summary of the joint work in his submission for the habilitation universitaire, which will be awarded by Aix-Marseille Université.

Acknowledgements

I am very grateful to have had Alison Etheridge as a supervisor: her explanations and ideas, as well as cordial advice on any topic have guided this journey. It has been a pleasure to work with Raphaël Forien, Julian Kern and Apolline Louvet, all exceptional mathematicians and friends.

I wish to thank Kira, my sister Paulina, my parents Annett and Torsten and my grandparents Renate and Klaus for all their support in the past and the present to continue with my way. Andrea has been a great flatmate and friend.

I will remember the third floor of the Statistics Department as a remarkable community of friends and extraordinary characters. Gabriel, Daniel, Thomas, Adrian (Martini + Fischer), Ian, Terence, Kimberly, Dave, David, Leo, Peter, Michal, Rivka, Felix, Matthew, João - I was fortunate to have been in the Statistics Department at the same time as them.

My DPhil has been made possible by the support of the Engineering and Physical Sciences Research Council Grant [EP/V520202/1].

I spend the summer of 2022 in Bonn, at the Hausdorff Institute for Mathematics founded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2047/1 - 390685813. It has been great to embark on this trip together with Adrian and Terence. While the start of the DPhil was under pandemic circumstances, I have warm memories of meeting many researchers of the scientific community, including Raphaël and Julian, in Bonn for the first time.

The layout of this thesis was created by Keith A. Gillow, Sam Evans and John McManigle.

Contents

1 Introduction **1**

2 A Λ-Fleming-Viot type model with intrinsically varying population size **19**

3 Isolation by distance under various forms of power-law dispersal **60**

4 Isolation by distance under varying population size **110**

5 Conclusion and outlook **169**

Bibliography **175**

Appendices

A Appendix **180**

B Appendix **189**

Abstract

This thesis covers three projects in the area of population genetics based on models from the broad class of Λ-Fleming-Viot processes. The unifying idea is that reproduction occurs through a sequence of events in which a proportion of the population is replaced by a parental type sampled from the population prior to the event.

The first work, joint with Julian Kern, introduces a Λ-Fleming-Viot model with varying population size. In events, instead of replacing a proportion of the population, first a random proportion is removed and then a random, not necessarily equal, mass of the chosen parental type is added. Consequently, each reproduction event may result in a change of the population size. This leads to a change in behaviour for small and frequent events: a deterministic drift naturally arises and the diffusive behaviour of the classical Λ-Fleming-Viot model is not possible anymore. Additionally, the varying population size complicates the analysis of genealogies - without knowledge of the population size the backwards in time dynamic of ancestral lineages is not Markovian.

The second and third work, joint with Raphaël Forien, consider spatial Λ-Fleming-Viot models. Events are now associated with a region in space in which the sampling of a parental type and the replacement of a proportion of the population takes place. We are interested in describing patterns of isolation by distance, the decline of relatedness of individuals from a population as their spatial separation increases. As individuals at different locations are only related if they had a common ancestor, the patterns reveal the spatial behaviour of the ancestors of the samples. In the second work, the size distribution of reproduction events follows various power-law dispersal rules leading to ancestors originating from potentially far away. This leads to much slower decay of genetic relatedness as a function of the separation than in classical cases, in which the affected region was bounded. We show that a variety of different behaviours is possible, in particular, when choosing different sizes for the

parental sampling and the area in which a proportion of the population is replaced. The third project extends this approach to a model with varying population density. The patterns of isolation by distance reveal that ancestors are more likely to have originated from areas of high population density. Additionally, if the population density is higher, the time to the common ancestor of the samples will be longer.

The analysis of all three projects will have a common element. Analytically, to infer information about genealogies, so events in the past, it is normally useful to reverse the flow of time and consider the behaviour of samples through the model backwards in time. Here, however, we will use two forwards-in-time approaches to deduce genealogical behaviour. The first, lookdown constructions, is well-known and equips the model with exchangeable labels imposing a hierarchy on individuals during reproduction events. As the initial labels are exchangeable, the model without knowledge of the labels will still evolve by the same dynamics. The second approach, used by the second and third project, uses neutral genetic types under an infinite-alleles model to trace families. There, we will demonstrate a profound relationship between central limit theorems and isolation by distance and thus ancestral behaviour.

1 Introduction

1.1 Overview

Earth's biosphere is inconceivably complex. Billions of years of evolution have created an abundance of optimised organic machines, interacting on countless scales to create the world we are witnessing. The basis of this beauty is lain by a central property of life: different forms of imperfect self-replication allow for selection of, often surprising, adaptations to certain environmental situations. In order to self-replicate, information describing an organism needs to be stored; today we know that this storage is provided by the DNA, whose functional units are known as genes.

The field of population genetics studies this self-replication mechanism of genes with the aim of providing models, which elegantly describe the circumstances, while still capturing the gist of a biological scenario. Naturally, the aforementioned complexity needs to be reduced and individuals are often described by so called types - subsuming their genetic information into a few or only a single value - and interact in no or only an abstract space. Additionally, to account at least for some of the disregarded complexity most modern models are stochastic.

This thesis will be concerned with Λ-Fleming-Viot processes, a class of population genetics models, both with rich applications and mathematical ingenuity covered in a large body of literature. Their common characteristic are reproduction events in which a certain proportion of the population is replaced by a parental type. The

parental type is typically chosen uniformly at random from the population prior to the event or, in the spatial case, from the prior population in a neighbourhood.

In the analysis of population genetics models, one is often interested in a backwards in time, genealogical, coalescing perspective. In most cases, we only observe present genetic data, from which we would like to reconstruct past events. As far as we know, all life on earth can be traced back to single cell organisms which populated earth billions of years ago. Therefore, we can aim to relate genetic samples through potential pasts in a backwards in time coalescing process, whose behaviour we would like to describe. Fortunately for the analysis, the backwards in time processes are often conveniently connected to the forwards in time model through a mathematical relationship known as duality.

Incorporating new biological phenomena can pose challenges for backwards in time duality arguments. This work will focus on varying population size as the potentially easiest form of an ecological component. Already in this case, a prominent difficulty is the loss of the Markov property of the backwards in time lineage counting process. In addition, the population size might not be reversible in time (see Chapter 2). Equally, in space, the past motion of ancestral lineages can reveal information about the profile of the population and thus its forwards trajectory. It becomes necessary to describe the joint dynamics of the time-reversal of the population size and the contained lineages - interactions which are normally too involved for meaningful progress.

Classically, the mathematical relationship to the dual process is often a sampling duality, which relates the probability that a certain number of individuals sampled from the present are of a specific type with the probability that all their (mathematical) ancestors are of this type. This is only possible if parents essentially pass their type to their offspring and there are no complex interactions of types during reproduction events. One can imagine various reproductive mechanisms in which potential parents, chosen from the population prior to the event, compete for reproductive success based on their types ([CS20]). If the interaction of types

becomes too complicated, such as a non-monotone fitness relationship, sampling dualities are generally not available.

In this thesis, we will consider two approaches based on neutral 'markers' that circumvent classical backwards in time arguments. Lookdown constructions, which represent stochastic processes as an infinite ordered system of particles, are a well-known approach based on neutral labels and are able to cover complicated biological situations ([DK96], [EK19], [Eth+24]). Another approach, first studied in [For22], is to relate patterns of 'isolation by distance' to central limit theorems for spatial models. If individuals are of specific types, one can measure isolation by distance by the probability of two individuals sampled at a given separation being of the same type. This probability can be calculated from the fluctuations described by central limit theorems. The resulting expressions reveal information about the large scale behaviour of the ancestral lineages, as sampled individuals can only be of the same type if no mutation occurred along either ancestral lineage since their most recent common ancestor.

This thesis will present three works placed in this joint realm. Next, we will introduce the three projects in more detail, while elaborating on the main connecting themes.

1.2 A Λ-Fleming-Viot type model with intrinsically varying population size

1.2.1 Origin of the Λ-Fleming-Viot model

Even today the Wright-Fisher model is the most ubiquitous model in population genetics. The premise is as follows: the population size is fixed and from one generation to the next, each individual picks a parent uniformly at random from the previous generation and inherits its type. Backwards in time, if several ancestral lineages are offspring of the same ancestor in the previous generation, they coalesce.

The enormous time spans of evolution and the often high number of individuals in nature allow for a fundamental idea: the large-time and large-population size limit of the model will be a rather good approximation to the real-world scenario.

If we consider large population sizes, the uniform parental choice leads to a faster decay in the probability of three or more individuals of a sample choosing the same parental type in the previous generation than for two individuals. In the limit, only binary coalescence events are possible, leading to Kingman's coalescent ([Kin82]). Forwards in time, the proportion of a single type then follows the continuous Wright-Fisher diffusion allowing only infinitesimal changes over a single generation. The term Fleming-Viot process originates from the paper [FV79], in which Fleming and Viot introduced the measure-valued analogue of the real-valued Wright-Fisher diffusion, assuming values in the probability measures $\mathcal{M}_1(\mathbb{K})$ on a potentially infinite type space $\mathbb{K}$. At any point in time, a now random measure $\rho_t \in \mathcal{M}_1(\mathbb{K})$ describes the probability $\rho_t(K)$ of finding a certain set of types $K \subseteq \mathbb{K}$ in a sample from the population. Considering the proportion process $\rho_t(K)$ for set of types $K \subseteq \mathbb{K}$ allows to transition from the measure-valued process back to the Wright-Fisher diffusion. Thus, we will use the term Wright-Fisher diffusion even in the measure-valued context to achieve a clearer distinction from the term Λ-Fleming-Viot model, which is introduced next.

Clearly, in a real population, more or less often events such as catastrophes or invasions of a new type, will influence the genetic composition of a significant part of the population in a rather short time frame; potentially a single generation. In a Λ-Fleming-Viot model a non-zero fraction of individuals can be replaced at a single point in time even in the infinite population limit ([Pit99], [Sag99], [DK99b]). In an event of 'impact' $u \in (0, 1]$ at time t, a proportion u of the population dies and is replaced by an equal mass of a new type k_0, sampled uniformly at random from the distribution $\rho_{t-} \in \mathcal{M}_1(\mathbb{K})$ of the population just before the event at time t. We could describe this informally by a transition of form

$$\rho_t \mapsto (1 - u)\rho_t + u\delta_{k_0}, \tag{1.1}$$

where δ_{k_0} denotes the Dirac measure on k_0. A mixture of events of different impacts u could be described by a mixing measure on $(0, 1]$ giving ample flexibility for modelling scenarios for a population. Ideally, to view this model as a generalisation of the

Wright-Fisher model, we would like to incorporate the Wright-Fisher diffusion. How could we parameterise the mixing measure to include the Wright-Fisher diffusion? Let us first consider a process with only events of impact u. If we let $u \downarrow 0$ and simultaneously accelerate the process by u^{-2}, the sequence of processes will converge to the Wright-Fisher diffusion. Therefore, a natural choice for the mixing measure on $(0, 1]$ is $u^{-2}\Lambda(du)$, where Λ is a finite measure on $(0, 1]$. The Wright-Fisher diffusion can then be seen as the 'zero-impact' process in the Λ-Fleming-Viot model. If $\Lambda = \delta_0$, where δ_0 denotes the Dirac-measure on zero, we therefore define the Λ-Fleming-Viot process to be the Wright-Fisher diffusion. In the resulting Λ-Fleming-Viot process, Λ determines the mixture of events of different impacts from $(0, 1]$ and $\Lambda(\{0\})$ the strength of the Wright-Fisher component, which describes infinitesimal changes in the composition of the population per generation. If $\Lambda((0, 1]) \neq 0$, the corresponding genealogy of a sample from the population, the Λ-coalescent, contains again multi-mergers, in which several lineages come together at the same time as they originated from the same ancestor.

1.2.2 Project 1: a Λ-Fleming-Viot model with intrinsically varying population size

A project together with Julian Kern contained in Chapter 2 explores the consequences of dropping the assumption that the 'mass' of offspring added in a reproduction event equals the amount removed. This possibility was first mentioned in [EK19, Section 4.2]. To this end, let z_b^u be a non-constant random variable with expectation u describing a (now random) new-born mass. We then consider events of the form

$$\rho_t \mapsto (1 - u)\rho_t + z_b^u \delta_{k_0}. \tag{1.2}$$

The process $(\rho_t)_{t\geq 0}$ describes now the total mass of different types and assumes values in the finite measures $\mathcal{M}_F(\mathbb{K})$. In parallel to the classical Λ-Fleming-Viot model, we would like to answer the same two questions: what is the associated 'zero-impact' process, i.e. the limit we obtain for $u \downarrow 0$? and then how can we parameterize the

associated mixing measure to include this process? It turns out that, as $u \downarrow 0$, we need to accelerate the process corresponding to Equation (1.2) by u^{-1} instead of u^{-2} to achieve the convergence. The limiting 'zero-impact' process is characterized by two dynamics: upward jumps and a drift. In between jumps, the total population size experiences a deterministic drift towards a value $\gamma_b \geq 0$ of the form

$$\frac{d}{dt}\rho_t(\mathbb{K}) = \Big(\gamma_b - \rho_t(\mathbb{K})\Big)\rho_t(\mathbb{K}).$$

The drift of the population size influences all types in the population equally. Additionally, there are upwards jumps of the population size, in which we choose a parent and just add a mass of the parental type without removing any (see Section 2.2.3, Equation (2.11)). In parallel to the classical model, to achieve consistency with this 'zero-impact' process in a mixed impact process, we sample events according to a finite-expectation measure $\Pi(du) := u^{-1}\Lambda(du)$, where Λ is again a finite measure on $(0, 1]$.

A similar argument can be carried out for processes in which the added mass is always u, whereas the proportion removed z_d^u equals u only in expectation. Again, as $u \downarrow 0$, if we accelerate by u^{-1} the sequence of processes converges to a limiting process with deterministic drift and now jumps, in which we just remove a random proportion of the population without adding any mass. In this case, the deterministic drift is described by

$$\frac{d}{dt}\rho_t(\mathbb{K}) = \Big(1 - \gamma_d\rho_t(\mathbb{K})\Big)\rho_t(\mathbb{K})$$

for a value $\gamma_d \geq 0$.

Both observations combined motivated us to consider the following model, which we call $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process. It is characterized by drift parameters $\boldsymbol{\gamma} = (\gamma_d, \gamma_b) \geq 0$ and a finite-mean measure Π on $[0, 1] \times (0, \infty)$ satisfying (for each component)

$$\int_{[0,1]\times(0,\infty)} \boldsymbol{z}\, d\Pi(\boldsymbol{z}) < +\infty.$$

This assumption corresponds to the previous parametrization of $u^{-1}\Lambda(du)$. The arrival of events is governed by a Poisson process. In events, a pair $\boldsymbol{z} = (z_d, z_b) \in$

$[0,1] \times (0,\infty)$ is sampled according to Π. The value z_d denotes the random proportion dying in the event and z_b the mass of the newly born individuals. This amounts to the transition

$$\rho_t \mapsto (1 - z_d)\rho_t + z_b\delta_{k_0},$$

where k_0 is again a parental type sampled uniformly according to the normalized distribution $\rho_{t-}/\rho_{t-}(\mathbb{K})$. Resulting from the 'zero-impact' component, at any time, the population size experiences a deterministic drift according to

$$\frac{d}{dt}\rho_t(\mathbb{K}) = \Big(\gamma_b - \gamma_d\rho_t(\mathbb{K})\Big)\rho_t(\mathbb{K}).$$

Jumps and drift result in the existence of a carrying capacity K which, if $\rho_0(\mathbb{K}) = K$, is the expected population size at any time (see Section 2.2.1, Section 2.2.3).

The main challenge in the analysis of the model is the intrinsically varying population size. The population not only acts as an extrinsic parameter varying with time, but also each reproduction event can lead to a change in population size. We explore various properties of the model. For example, under a natural convergence criterion on the characterising parameters $(\boldsymbol{\gamma}, \Pi)$ the associated sequence of processes converges to the process associated to the limiting parameter (see Theorem 2.2.5). Notably, one might wonder whether the Wright-Fisher diffusion, the 'zero-impact' of the classical, constant population size, Λ-Fleming-Viot process can appear. Theorem 2.2.6 demonstrates that, if the convergence of the sequence of parameters $(\boldsymbol{\gamma}, \Pi)$ is such that the population size becomes constant, the Wright-Fisher diffusion reappears and we can recover any standard Λ-Fleming-Viot process in the limit (see Theorem 2.2.6). Combined, Theorem 2.2.5 and Theorem 2.2.6 illustrate that a constant population size leads to additional cancellations. They are necessary to consider the faster time scale leading to the emergence of a diffusion, which can be regarded as very rapid, small scale changes in the population (see Equation (2.12)).

1.2.3 Forwards in time approach 1: lookdown constructions

Even though this model might seem like only a slight deviation from the standard Λ-Fleming-Viot process, an important difference arises when considering duality. The special case of a population size with upwards jumps and a downward drift indicates that the population size is not necessarily reversible. The current state of the population will contain information about the sequence of past events. To consider the backwards in time genealogy, it is not sufficient to work with a Poisson process with the same intensities as the forwards in time process. Instead, one requires the time-reversed population size. Both processes are linked, and a sampling duality requires a conditioning on the whole evolution of the population size process.

In Section 2.7, we will illustrate that the lookdown construction of [DK99b] is a suitable tool, with a more forwards in time flavour, for the analysis of the coalescence behaviour. The duality conditional on the population size is embedded in the construction. The approach allows one to represent the process as an infinite sequence of exchangeable individuals. We assign to each label from the natural numbers a type chosen at random from the type distribution of the population. A change of fraction $f \in (0, 1]$ of the composition of the population occurs as follows. We first flip for each individual independently a biased coin which comes up heads with probability f. All individuals for which the coin flip delivers a head participate in the event. As the parent we choose the participating individual with the lowest label. In other words, the individuals with the higher labels 'look down' and adopt the type of the lowest-label individual. One can recover the model without labels from the infinite sequence as the limit of the empirical averages for an ever-increasing number of individuals. Due to the initial random assignment of individuals to labels and the specific reproduction events, the dynamics of the model without labels remain unchanged.

Through the lookdown framework we will be able to analyse properties such as 'fixation' of types (Theorem 2.2.7) and not coming down from infinity (Theorem 2.2.8) (the latter property effectively describing that 'fixation' will require an infinite amount of time). More of interest to readers already familiar with lookdown

constructions, we also deduce the well-posedness of the martingale problem for the process and show that one is able to reconstruct the driving noise from the martingale problem (see Section 2.6).

1.3 Spatial central limit theorems describing isolation by distance

1.3.1 Isolation by distance and the Wright-Malécot formula

Due to the limited time between birth and maturity, in most species, individuals found the next generation at a position close to where they were born, and the likelihood of having offspring at a certain location decreases as a function of the distance to the birthplace. Naturally, the same heuristic holds true for the genes carried by the individuals. With each replication of the genetic information from one generation to the next, small deviations occur, most of which might not significantly alter the fitness of individuals, so called neutral mutations. Combining both observations, individuals further apart in space have, on average, also a larger genetic distance. This phenomenon is known as isolation by distance.

To measure this isolation, let us assume that individuals carry specific genetic types, and due to the vastness of the information on the DNA, any mutation leads to new types, not previously observed. Then the probability to sample the same type at two different locations provides a measure for the isolation, determined by the movement of individuals relative to the time-frame provided by the mutation rate. To find the same type, tracing back their ancestral lines, neither must have seen a mutation since their coalescence. This is known as identity by descent. Wright and Malécot ([Wri43], [Mal48]) calculated this probability of identity by descent (POI) and discovered the formula

$$F(x) \approx \int_0^\infty e^{-2\mu t} G_{2t}(x) dt, \tag{1.3}$$

where x is the vector describing the spatial separation of the sampled individuals, μ the mutation rate along the ancestral lines and G_t the transition kernel of a Brownian motion. As the transition kernel is invariant under rotation, so is the above formula

under rotation of x. The exponential $e^{-2\mu t}$ captures the probability of not observing a mutation along either ancestral lineage. The quantity $G_{2t}(x)$ describes the density of the probability of finding ancestral lineages started at separation x at the same place after time t. Coalescence is taken to occur instantaneously in the limit. Over the large temporal and spatial scales of nature, spatial Brownian motion provides a good approximation of the jump process described by the birth places of the ancestors in the ancestral lineages due to the central limit theorem. Integrating the probability of not seeing a mutation and the probability of coalescence over the complete past from the present time (zero) to the distant past (at infinity), one arrives at the probability of identity.

Two projects, contained in Chapter 3 and Chapter 4, which are joint work with Raphaël Forien, will uncover generalisations of the Wright-Malécot formula (1.3) to a wide class of spatial population models.

1.3.2 Spatial Λ-Fleming-Viot model

The spatial Λ-Fleming-Viot model (SLFV) offers an elegant extension of the Λ-Fleming-Viot paradigm to space ([Eth08]). Effectively, a standard Λ-Fleming-Viot process $\rho_t(z)$ is associated to every point in space, say $z \in \mathbb{R}^d$, with coordination between the reproduction events at close positions. Instead of individual Poisson point processes for each location, a Poisson process determines the space-time point (x, t), event-region (typically ball-shaped, in which case events come along with a radius $r > 0$) and impact $u \in (0, 1]$. In our ball-shaped case, a parent k_0 is sampled from the average type distribution

$$\frac{1}{V_r} \int_{B(x,r)} \rho_{t-}(z, dk) dz,$$

where $B(x, r)$ denotes the ball of radius r and centre x and V_r its volume. The previous replacement mechanism (1.1) of the standard Λ-Fleming-Viot model now occurs for every position inside the affected ball

$$\rho_t(z) \mapsto \begin{cases} (1-u)\rho_t(z) + u\delta_{k_0} & z \in B(x, r), \\ \rho_t(z) & z \notin B(x, r). \end{cases}$$

This model has found widespread applications such as explaining different genealogies and effective population sizes [BEV10], hybrid zones [EFP17], and varying selection [BEK21].

1.3.3 Forwards in time approach 2: central limit theorem approach

Based on SLFV models, in [For22] a fascinating connection between central limit theorems and the POI/Wright-Malécot formula was found.

If we sample individuals at time t according to two probability densities ϕ, ψ on $\mathbb{R}^d$, the probability to find the same type is given by

$$\begin{aligned} P_t(\phi,\psi) &= \int_{(\mathbb{R}^d\times[0,1])^2} \mathbb{1}_\Delta(k_1,k_2)\phi(x)\psi(y)\rho_t(x,dk_1)\rho_t(y,dk_2)dxdy \\ &= \mathbb{E}\Big[\Big\langle \rho_t\otimes\rho_t, \mathbb{1}_\Delta(\phi\otimes\psi)\Big\rangle\Big], \end{aligned}$$

where $\mathbb{1}_\Delta(k_1,k_2) = \mathbb{1}_{k_1=k_2}$ denotes the diagonal. The mass two copies $\rho_t(x), \rho_t(y)$ of the process place on the diagonal is precisely the probability of sampling the same type from both positions x, y. The probability densities ϕ, ψ are necessary as point-wise evaluation is not possible in the SLFV.

Suppose $\mathbb{K} = [0,1]$ and that in addition to standard reproduction events, in between consecutive events that affect the point x at times t_1 and t_2, mutation occurs according to

$$\rho_t(x,dk) = e^{-\mu(t-t_1)}\rho_{t_1}(x,dk) + \Big(1-e^{-\mu(t-t_1)}\Big)dk,$$

where dk denotes the uniform distribution on $[0,1]$. This exponential decay towards a uniform distribution can be interpreted as individuals assuming new uniform types on $[0,1]$ at rate μ and was first introduced in [VW15]. This mutation mechanism captures the idea of the POI: at rate μ, as new types are created, previous information vanishes and becomes irrelevant for the POI.

As mentioned previously, evolutionary time-scales and spatial scales are often large compared to the life-span and age of individuals. Thus, we will 'zoom out' both in space and time and consider a diffusive scaling. This scaling is controlled by

a sequence $\delta_N \to 0$ as $N \to \infty$. For example, we rescale space by setting $x' = x/\delta_N$. At the same time, we let the impact $u \downarrow 0$ emulating an increase in population size corresponding to a law of large numbers. This is achieved by considering $u' = u/N$ as $N \to \infty$. All other parameters will be rescaled in terms of N and δ_N. Having two parameters determining the scaling gives us more flexibility in identifying interesting regimes. We also let the mutation rate $\mu \downarrow 0$, in such a way that the probability of mutation stays of order one on a time period one of accelerated time. This assumption is justified by the low mutation rate observed in nature. Under this scaling and assuming $\rho_0^N = \lambda$, the sequence of rescaled processes ρ_t^N converges to Lebesgue measure λ on $\mathbb{R}^d \times [0,1]$. We can view λ as the element of the statespace of ρ_t^N, which is uniformly distributed on the type space $[0,1]$ for all positions $x \in \mathbb{R}^d$. The uniform distribution on types results from the mutation mechanism, which dominates as the impact converges to zero. The deterministic limit allows one to consider the fluctuations

$$Z_t^N = \sqrt{N\delta_N^{2-d}}(\rho_t^N - \lambda),$$

where we assume $\sqrt{N\delta_N^{2-d}} \to \infty$ as $N \to \infty$, with more precise information given in Chapter 2 and Chapter 3. The existence of the limit of the sequence of fluctuation processes Z_t^N can be seen as a central limit theorem. The limit is given by a stochastic partial differential equation (SPDE) of the form

$$dZ_t = (\Delta Z_t - \mu Z_t)dt + dW_t, \tag{1.4}$$

for the standard Laplace operator Δ and a Gaussian noise W_t on $\mathbb{R}_+ \times \mathbb{R}^d \times [0,1]$, uncorrelated in space and time.

Returning now to the above POI calculation, sampling from the N-th rescaled process, we can write

$$\begin{aligned} P_t^N(\phi,\psi) &= \mathbb{E}\Big[\Big\langle \rho_t^N \otimes \rho_t^N, \mathbb{1}_\Delta(\phi\otimes\psi)\Big\rangle\Big] \\ &= \mathbb{E}\Big[\Big\langle (\rho_t^N - \lambda) \otimes (\rho_t^N - \lambda), \mathbb{1}_\Delta(\phi\otimes\psi)\Big\rangle\Big] \\ &= (N\delta_N^{2-d})^{-1}\mathbb{E}\Big[\Big\langle Z_t^N \otimes Z_t^N, \mathbb{1}_\Delta(\phi\otimes\psi)\Big\rangle\Big]. \end{aligned}$$

Here, we used that as λ is uniform on the type space $[0,1]$, it does not contribute if evaluated against the diagonal function $\mathbb{1}_\Delta = \mathbb{1}_{k_1=k_2}$. This establishes a connection between the POI and the fluctuations. Taking limits on both sides

$$\lim_{t\to\infty}\lim_{N\to\infty} N\delta_N^{2-d} P_t^N(\phi,\psi) = \mathbb{E}\Big[\Big\langle Z_\infty \otimes Z_\infty, \mathbb{1}_\Delta(\phi\otimes\psi)\Big\rangle\Big], \tag{1.5}$$

where Z_∞ denotes the stationary distribution of (1.4). In case of events having a fixed radius $R>0$ and impact $u\in(0,1]$, [For22] recovers the classical Wright-Malécot formula (1.3), integrated with respect to the sampling densities ϕ,ψ, for the right-hand side of Equation (1.5). However, if we choose the radius in events according to

$$\nu(dr) = \mathbb{1}_{[1,\infty)}(r)\frac{dr}{r^{1+d+\alpha}}, \alpha\in(0,2),$$

the formula for sampling individuals according to two locations x,y becomes (ignoring the integration with respect to the densities ϕ,ψ)

$$F(x,y) = \int_0^\infty e^{-2\mu t} G_t^\alpha(x-z_1) G_t^\alpha(y-z_2) K(z_1,z_2) dz_1 dz_2.$$

Compared to (1.3), the transition kernel G_t of a Brownian motion has been replaced by the transition kernel of an α-stable process. This reflects an altered movement of the ancestral lineages due to the long-range jumps. Additionally, the term

$$K(z_1,z_2) := \frac{1}{\|z_1-z_2\|^{d+\alpha}}$$

captures global coalescence through large jumps if the two ancestral lineages are at two separate locations z_1,z_2. In the classical Wright-Malécot formula instead only local coalescence $K(z_1,z_2)=\mathbb{1}_{z_1=z_2}$ with both lineages at the same position was possible. Accordingly, in the SPDE (1.4) the operator $\mathcal{D}$ is replaced by a fractional Laplacian and the Gaussian noise has spatial correlations determined by $K(z_1,z_2)$.

The two joint projects with Raphaël Forien will illustrate that this correspondence of central limit theorems to the asymptotic motion and coalescence behaviour of ancestral lineages is a more general phenomenon.

1.3.4 Project 2: power-law dispersal

In Chapter 3, we generalise this approach to a much broader range of models. Many works in biology have observed that species movement includes occasional long-range displacement. In plants this may most prominently occur through seed or pollen dispersal ([CMS00], [Val+17]). Similarly, in marine species eggs can cross large distances ([Kel+14]) or, also in terrestrial animals, mature individuals may carry out long-range migration ([FK07]). In any of these cases, exceptionally large displacements might provide an advantage as they allow individuals to cross less habitable regions.

In addition to choosing the radius of events according to heavy-tailed distributions, we introduce two modifications. First, we allow the impact to decay as a function of the radius and, second, we allow different radii r_1, r_2 for the ball $B(x, r_1)$, in which the parental type is sampled, and the ball $B(x, r_2)$, in which a proportion of the population is replaced. For example, in what we call the one-tail regime, the intensity measure ν governing the reproduction events is chosen as

$$\nu_N(du, dr_1, dr_2) = \mathbb{1}_{r_2 \geq 1}\, \delta_{u_N r_2^{-c}}(du)\, \delta_{r_2^b}(dr_1)\, \frac{dr_2}{r_2^{1+d+a-c}},$$

where $u_N \in (0, 1]$, $a \in (0, \infty)$ and $b, c \in [0, \infty)$ are parameters. The name 'one-tail' is chosen to reflect that r_1 is a deterministic function of r_2. In another regime, the two-tail regime, r_1 and r_2 will carry independent heavy-tails (see Section 3.3.1).

In [For22] regimes were obtained in which either ancestral lineages approximated Brownian motions (a continuous diffusive process) or stable motions, which may also include long-range jumps. Through the Wright-Malécot formula, one could observe that the Brownian motions coalesce through being at the same position ('locally', no spatial correlation in the noise driving the SPDE). In contrast, the stable lineages coalesce through large-scale jumps ('globally', spatial correlations in the noise term of the SPDE). This behaviour was also observed for the genealogy consisting of stable lineages studied in [BEV13b]. In Chapter 2, we exhibit that local and global coalescence (according to a variety of kernels) are possible for any motion. For example, we demonstrate that stable ancestral lineages might

coalesce only locally. Surprisingly and more of a mathematical interest, even Brownian lineages might coalesce globally in dimension three and higher. The new regimes complicate inference of potential parameters. In more biological terms, these results indicate that, for example in the case of marine fish, it may be difficult to distinguish from genetic patterns alone whether a slower than exponential decay in relatedness originated from fish exhibiting power-law foraging patterns or power-law dispersal of eggs.

1.3.5 Project 3: varying population size

In Chapter 4, we generalise the previous results to a SLFV model with varying population size. The population size is denoted by

$$n_t(x) = \rho_t([0,1]) = \int_0^1 \rho_t(x, dk)$$

and, as for the process ρ_t, the population size is only defined up to a Lebesgue null set. The impact $u > 0$ and radius $R > 0$ of reproduction events will be fixed quantities again. To avoid the question of uniqueness and to simplify calculations, we will work on a torus $\mathbb{T}^d$ instead of the complete space $\mathbb{R}^d$. To include the varying population size, we will introduce three changes to the reproduction events:

1. Parental types will be sampled according to the composition of the population inside the ball $B(x, R)$ normalized by the average population size therein.

2. Reminiscent of Chapter 2, the newborn mass during events can differ from the removed mass. The newborn mass is modulated by a so-called growth function $r : \mathbb{T}^d \times \mathbb{R}_+ \to [-1, \infty)$, where $r(z, \overline{n})$ depends on space $z \in \mathbb{T}^d$ and the average local population size $\overline{n}$. The peculiarity of the growth function being larger than minus one will become apparent soon. The average local population size is defined as

   $$\overline{n_t}(z, R) := \frac{1}{V_R} \int_{B(z,R)} n_t(x) dx.$$

   This 'bar' notation will be frequently used to denote spatial averages in this thesis.

3. We consider a population in which the relative fraction u of individuals being replaced over a period of time of length one is independent of the local population size $\overline{n}$. This assumption could represent, for example, the death of individuals through ageing. If the population size is larger, this fraction will correspond to a larger total mass of new individuals. In turn, we would expect more parents to contribute their genetic material as the pool of offspring increases.

   We could capture this behaviour by choosing more parents in an event as the population size increases. Instead, we decided to increase the number of events as the population size increases, while the impact of each event decreases with the population size. We choose events to occur at rate $\overline{n}$ with each event replacing a proportion $\frac{u}{\overline{n}}$ of the population. Over a period of time of length one, we then roughly expect $\overline{n}$ events of impact $\frac{u}{\overline{n}}$, replacing a fraction $\overline{n}\frac{u}{\overline{n}} = u$.

In more mathematical terms, events fall at rate $\overline{n}_t + 1$ with the plus one being explained shortly. In an event at (x, t), the population size $n_t(z)$ changes to

$$\Big(1 - \frac{u}{\overline{n}_{t-}(x,R)+1}\Big)n_{t-}(z) + \frac{u}{\overline{n}_{t-}(x,R)+1}\Big(1 + r_x(\overline{n}_{t-}(x,R))\Big)\overline{n}_{t-}(x,R),$$

if $z \in B(x, R)$. Let use first come back to the lower bound of minus one of the growth function $r_z(\overline{n}_{-t})$. This prevents the new-born mass from being negative. Adding plus one to the local average population size in the denominator, prevents the term $1 - \frac{u}{\overline{n}}$ from becoming negative for small population sizes. Additionally, as the plus one appears in the rate of events, it conveys the idea that there is always a certain reproductive activity in the population, even for small population sizes.

In both the models of Chapter 2 and Chapter 4 each reproduction event leads to a change of the population size and is not extrinsically prescribed. Biologically, in the model of Section 1.2.2/Chapter 2, reproduction events, in which a proportion dies and a potentially large offspring mass is added, are seen to occur at a fixed rate. If started from the equilibrium, major changes in population size are caused by events with either a high proportion of the population prior to the event dying or an exceptionally large number of offspring. In the model introduced here, instead,

for a certain local population size, the turnover rate and size of reproduction events are almost constant. If started from the equilibrium, major changes in population size originate from changes of the growth function r over time and different values of the growth function at different event locations. The suitability of each model depends on the biological circumstances.

Under an appropriate rescaling similar to the fixed radius case in Chapter 3, the rescaled population size converges to the solution of a reaction-diffusion equation

$$\partial_t n_t = uV_R\Big(\frac{R^2}{d+2}\Delta n_t + r(n_t)n_t\Big),$$

where r denotes again the growth function and may depend on space. In parallel to Chapter 3, as the law of large numbers, we would expect the convergence of the rescaled sequence of processes to $n_t\lambda$, where by $n_t\lambda$ we denote the measure on $\mathbb{T}^d \times [0,1]$, which is uniform on types and whose spatial marginal is n_t. We will consider the fluctuations around Lebesgue measure on $\mathbb{T}^d \times [0,1]$ times a deterministic approximation m_t^N of the population size

$$Z_t^N := \sqrt{N\delta_N^{2-d}}\Big(\rho_t^N - \lambda m_t^N\Big),$$

where $m_t^N \to n_t$ as $N \to \infty$. As the deterministic approximation we take the compensating part of the semimartingale representation resulting from the martingale problem for n_t^N. The limit of the sequence of the fluctuations is given by the mild solution to the SPDE

$$dZ_t = \left[uV_R\left(\frac{R^2}{d+2}\Delta + r(n_t)\right)Z_t - \mu Z_t + \Big(\mu + uV_R n_t r'(n_t)\Big)\int_0^1 Z_t(\cdot, dk)\right]dt + dW_t,$$

for a Gaussian random field on $\mathbb{R}_+ \times \mathbb{T}^d \times [0,1]$ (see Theorem 4.3.3). The term involving $\int_0^1 Z_t(\cdot, dk)$ results from recentering by the deterministic approximation instead of the stochastic population size. We will see that the resulting Wright-Malécot formulae are invariant under terms containing $\int_0^1 Z_t(\cdot, dk)$. Whereas in Chapter 3 Wright-Malécot formulae could be guessed from the SPDEs, the corresponding calculation is more involved in this case. The Wright-Malécot formula shows that ancestral lineages experience a drift according to the population size

profile and coalescence happens at a rate inversely proportional to the population size once the ancestral lineages are at the same position (see Theorem 4.3.6). This paper links mathematical papers, obtaining a motion of ancestral lineages similar to [Eth+24], and more applied works such as [Rin+18], where barriers to gene flow are inferred.

1.4 Outline

The next three chapters will contain the three projects presented above. Chapter 2 contains the manuscript "A Λ-Fleming-Viot type model with intrinsically varying population size", which is joint work with Julian Kern. Chapter 3 and Chapter 4 contain joint works with Raphaël Forien, namely "Central limit theorems describing isolation by distance under various forms of power-law dispersal" and "Central limit theorems describing isolation by distance under varying population size". Minor adaptations to the manuscripts have been made to increase the coherence of this thesis and to cover the corrections specified by the examiners. Chapter 5 draws conclusions from these works and presents various challenges and future research directions.

Joint work with Julian Kern

2

A Λ-Fleming-Viot type model with intrinsically varying population size

Contents

2.1 Introduction **19**
2.2 Model and main results **22**
2.3 The population size process **31**
2.4 Closedness in the space of characteristics **35**
2.5 Relation to the Λ-Fleming-Viot process **38**
2.6 Lookdown construction **42**
2.7 Genealogy and duality **52**

2.1 Introduction

Genetics and Ecology Early mathematical models of evolution could often be classified as belonging either to genetics or population dynamics. The former focusses on the development of the genetic information carried by the population through biological mechanisms, most prominently mutation, selection, recombination and genetic drift. These effects already result in a high amount of stochasticity. In order to handle the models, compromises are made in other areas such as the interplay with the environment or other species. These, in turn, are the strongholds of population dynamics, which often neglect the genetic code of individuals.

It has become a major endeavour of population modelling to weaken this separation. A common characteristic of models at the intersection of ecology

and genetics is to incorporate a form of varying population size required for the implementation of ecological factors such as external catastrophes, selective sweeps, interspecies relationships or competition for a fluctuating abundance of resources. Conceptually, the population size can vary extrinsically or intrinsically. In the first case, the population size enters as a parameter into the development of the population and is driven by a randomness independent of the remaining stochasticity. This lends itself to modelling external influences such as catastrophes or the loss of habitat, see e.g. [KK03; CPS22; Cor+22]. Whereas the population size is assumed to be only slowly varying in [KK03], it is taken to undergo drastic changes (so-called *bottlenecks*) in [CPS22; Cor+22], leading to vastly different genealogies.

In contrast, models with intrinsically varying population size inextricably link the reproductive dynamics with the population size. The recent works [EK19] and [Eth+24] have emphasised the potential of the lookdown framework in analysing intrinsically varying models (see also [DK96; DK99b; DK99a]). The former provides a toolbox for population models and constructs a spatial Λ-Fleming-Viot model with varying population size, which has been a major motivation for this work. The latter analyses a locally regulated population with a two-step reproduction mechanism with a juvenile and adult stage, which allows for a wide range of population profile patterns to appear.

A varying size Λ-Fleming-Viot process We wish to further contribute to the understanding of the implications of intrinsically varying population size by considering a model based on a variation of the classical Λ-Fleming-Viot paradigm. More precisely, the model will evolve through a series of reproduction events during which a random proportion z_d of the population is killed and a random mass z_b of offspring with a type chosen uniformly from the population prior to the event is added. We will allow for a deterministic drift of the population size which arises naturally from the random death proportions and birth masses (see the discussion in Section 2.2.3). Although to our knowledge, this specific model has not yet been considered, it fits into the general framework from [DK99b]. A similar model in

the spatial context, but under much more restrictive assumptions can be found in [EK19, Section 4.2]. Our model should also be compared to the model in [Sch03], where individuals can be parental to large families, but the population size is reduced to a fixed quantity after reproduction.

The model is not a *generalisation* of the Λ-Fleming-Viot model in a strict sense: although it covers Λ-Fleming-Viot processes for a range of impact measures Λ, not all are included. Instead, it should be viewed as the natural variant for which the masses, which are removed and added in an event, do not need to be equal. This is complemented by the fact that we can approximate any classical Λ-Fleming-Viot process arbitrarily well. Despite this relatedness with classical models, the interplay between varying population size and reproduction events will necessitate additional care in the analysis. Even though duality is still available in a restricted sense, the dual coalescent is time inhomogeneous, so that our analysis of genealogical properties relies mainly on a *forward-in-time* approach.

The major tool in this analysis is the lookdown construction from [DK99b] which allows one to represent the measure-valued model as a countable particle system, where each individual has a unique label, or level, in $\mathbb{N}$. The main advantage of lookdown constructions is the fact that offspring always “look down" to their parent: from all individuals participating in a reproduction event, the one with the lowest level is parental. This incorporates the corresponding coalescent backwards in time in a consistent way and allows one to trace lines of descent forward in time, independently of their changing type.

Outline Section 2.2 contains the definition of the model, presents the main results and closes with a discussion of the model choice. In Section 2.3, we will start our analysis by collecting properties of the population size process. Section 2.4 shows the closedness of the class of processes under a natural convergence condition. In Section 2.5, we will see that any Λ-Fleming-Viot process can be recovered as a limit of our processes. Section 2.6 recalls the lookdown construction of [DK99b] and provides a well-posedness proof through the Markov Mapping Theorem. The

genealogy of the process is studied in Section 2.7. Code for simulation is provided at `github.com/mushunrek/GammaPiModel.jl`.

2.2 Model and main results

2.2.1 The $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process

Throughout this work, we consider measure-valued populations taking values in the space $\mathcal{M}_F(\mathbb{K})$ of finite measures on a compact Polish type space $\mathbb{K}$, usually $\mathbb{K} = \{a, A\}$ or $\mathbb{K} = [0, 1]$. Denote by $\mathcal{M}_1(\mathbb{K}) \subseteq \mathcal{M}_F(\mathbb{K})$ the set of probability measures on $\mathbb{K}$.

Definition 2.2.1 (Events and Characteristics)**.** In the context of deaths and births, we will always denote pairs $\boldsymbol{z}, \boldsymbol{\gamma} \in \mathbb{R}^2$ by bold letters and their entries by (z_d, z_b) and (γ_d, γ_b).

1. We denote by
$$\mathcal{Z} := \Big([0, 1] \times [0, +\infty)\Big) \setminus \{(0, 0), (1, 0)\}$$
the *event space*. Throughout this chapter, its elements will be referred to as *events* and be denoted by $\boldsymbol{z} =: (z_d, z_b)$. Here, z_d denotes the proportion of individuals to be killed, and z_b the mass of offspring born during the event $\boldsymbol{z}$.

2. We define the space $\mathfrak{C}$ of *characteristics* to contain exactly the pairs $(\boldsymbol{\gamma}, \Pi) \in \mathbb{R}^2 \times \mathcal{M}(\mathcal{Z})$ satisfying

   i) $\boldsymbol{\gamma} \geq \mathbf{0}$,

   ii) the first moment condition
   $$\int_{\mathcal{Z}} \boldsymbol{z} \ \mathrm{d}\Pi(\boldsymbol{z}) < +\infty, \tag{2.1}$$

   iii) and the balance condition
   $$\gamma_d + \int_{\mathcal{Z}} z_d \ \mathrm{d}\Pi(\boldsymbol{z}) = \gamma_b + \int_{\mathcal{Z}} z_b \ \mathrm{d}\Pi(\boldsymbol{z}). \tag{2.2}$$

In the above, $\mathcal{M}(\mathcal{Z})$ denotes the set of σ-finite measures on $\mathcal{Z}$. The inequalities are to be understood entrywise. In analogy to $\boldsymbol{z}$, we will write Π_d and Π_b for the marginals of Π.

The events $(0,0)$ and $(1,0)$ are excluded to avoid trivialities: the first does not change the population, whereas the second kills off all individuals without replacing any. The second condition (2.1) guarantees a biologically reasonable finite expected population size. It implies together with $\boldsymbol{\gamma} \geq \mathbf{0}$ that $(\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$ defines a two-dimensional subordinator with drift $\boldsymbol{\gamma}$ and jump measure Π.

Definition 2.2.2 (The $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process)**.** Let $(\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$ be a characteristic. Define the operator

$$\begin{aligned}\mathcal{L}^{\boldsymbol{\gamma},\Pi}\Phi(\sigma) = \Big(\gamma_b - \gamma_d\sigma(\mathbb{K})\Big) D\Phi\left(\sigma, \frac{\sigma}{\sigma(\mathbb{K})}\right) \\ + \int_{\mathbb{K}}\int_{\mathcal{Z}} \Big[\Phi\Big((1-z_d)\sigma + z_b\delta_\kappa\Big) - \Phi(\sigma)\Big] \, \mathrm{d}\Pi(\boldsymbol{z}) \frac{\mathrm{d}\sigma(\kappa)}{\sigma(\mathbb{K})}\end{aligned} \tag{2.3}$$

for all bounded Fréchet-differentiable $\Phi \in C_b^1\mathcal{M}_F(\mathbb{K})$ with bounded derivative. Any $\mathcal{M}_F(\mathbb{K})$-valued solution σ to the martingale problem for $\mathcal{L}^{\boldsymbol{\gamma},\Pi}$ is called a $(\boldsymbol{\gamma}, \Pi)$-*Fleming-Viot process*.

Here, $D\Phi\big(\sigma, \sigma'\big)$ denotes the derivative of Φ at σ in direction σ'. To avoid heavy notation, we designate here and afterwards both the stochastic process and a typical element of the ambient space by the same symbol. In Section 2.2.3 we give a detailed motivation of the model choice.

Any $\sigma \in \mathcal{M}_F(\mathbb{K}) \setminus \{0\}$ can be uniquely decomposed into $\sigma = N\rho$, where $N := \sigma(\mathbb{K})$ is the *population size* and the probability measure $\rho \in \mathcal{M}_1(\mathbb{K})$ is the *type distribution*. A Poisson construction of the population size process N can be found in Theorem 2.3.1 and of the complete process in Section 2.6.1. If σ is a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process, then N is a solution to the martingale problem for

$$\mathcal{L}^{\boldsymbol{\gamma},\Pi}f(N) = \Big(\gamma_b - \gamma_d N\Big) f'(N) + \int_{\mathcal{Z}} \Big[f\Big((1-z_d)N + z_b\Big) - f(N)\Big] \, \mathrm{d}\Pi(\boldsymbol{z}). \tag{2.4}$$

More precisely, σ is a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process if and only if (N, ρ) is a solution to the martingale problem for

$$\begin{aligned}&\mathcal{L}^{\boldsymbol{\gamma},\Pi} F(N,\rho)\\ &= \Big(\gamma_b - \gamma_d \sigma(\mathbb{K})\Big)\partial_N F(N,\rho)\\ &\quad + \int_{\mathbb{K}}\int_{\mathcal{Z}} \Big[F\Big((1-z_d)N + z_b, (1-\overline{\boldsymbol{z}})\,\rho + \overline{\boldsymbol{z}}\delta_\kappa\Big) - F(N,\rho)\Big]\ \mathrm{d}\Pi(\boldsymbol{z})\ \mathrm{d}\rho(\kappa),\end{aligned} \tag{2.5}$$

where we denote by

$$\overline{\boldsymbol{z}} := \overline{\boldsymbol{z}}(N, \boldsymbol{z}) := \frac{z_b}{(1-z_d)N + z_b}$$

the *effective impact* of an event $\boldsymbol{z}$ at population size N. We keep the dependence on N implicit to improve the readability. In Equation (2.4) and Equation (2.5), we abuse the notation for $\mathcal{L}^{\boldsymbol{\gamma},\Pi}$ by applying it to functions $f \in \mathcal{C}_b^1((0,\infty))$ and $F \in C_b^1((0,+\infty)\times\mathcal{M}_1(\mathbb{K}))$ respectively. Furthermore, the equivalence between the martingale problems on $\mathcal{M}_F(\mathbb{K})$ and $(0,+\infty)\times\mathcal{M}_1(\mathbb{K})$ only holds because the population does indeed not vanish, which we will prove rigorously in Section 2.3.

2.2.2 Main results

Properties of the model We show in Section 2.6 that the model can be understood with the framework of [DK99b]. In particular, the following result follows solely from studying the population size process.

Theorem 2.2.3. *The martingale problem for the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process is well-posed for any characteristic $(\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$. Furthermore, the unique solution is strongly Markovian.*

We will show existence of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process by approximation, the uniqueness will follow from the uniqueness of the lookdown construction of [DK99b], which can be conveyed to the underlying process by the Markov Mapping theorem [KN11]. The application of the Markov Mapping theorem is contained in Lemma 2.6.5. A discussion of the uniqueness argument of [DK99b] is in Section 2.6.1. The uniqueness argument comes down to showing that in the lookdown construction

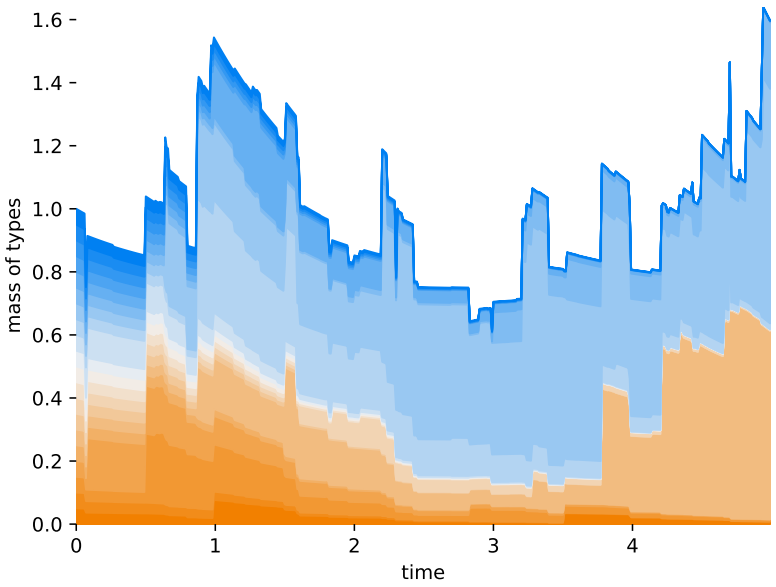


Figure 2.1: Simulation of a (γ, Π)-Fleming-Viot process: we start with 20 different types of equal mass 0.05, each represented by a different colour. Here, $\boldsymbol{\gamma} \approx (0.74, 1)$ and the marginals of Π are independent such that Π_d has density $\mathbb{1}_{[0.0001,0.3]}(x)\frac{3}{x}$, and Π_b has density $\propto \mathbb{1}_{[0.0001,0.4]}(x)\frac{1}{x}$.

the rate at which events affect a finite number of levels is finite. In [DK99b] this is shown for varying population size using a localization argument up to stopping times, these are in our case provided by Theorem 2.3.1. The concluding proof is at the end of Section 2.6.

In the context of Λ-Fleming-Viot processes, reproduction events are parametrised by an *impact* $u \in (0, 1]$ representing the proportion of the population to be replaced. Events of small impact $u \approx 0$ are usually modelled to occur at an accelerated rate of order u^{-2} as small reproduction events are expected to be much more frequent in large populations than large scale, catastrophic events. In the limit $u \downarrow 0$, the dynamics can be approximated by the Wright-Fisher diffusion which therefore can be thought of as "0-impact Fleming-Viot".

As soon as events are balanced only in mean (2.2), events $\boldsymbol{z} \approx \mathbf{0}$ can only happen at rate proportional to $|\boldsymbol{z}|^{-1}$ (see the discussion in Section 2.2.3). This is ensured by the first moment condition (2.1) which guarantees the existence of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process. Mirroring the approach from the Λ-Fleming-Viot

process, it is natural to ask whether an additional process is required to describe the "**0**-impact" behaviour. The next theorem gives a negative answer: the class of $(\boldsymbol{\gamma}, \Pi)$-models is closed under a natural notion of the convergence of characteristics.

For a Polish space E, we denote by $\mathbb{D}_{[0,+\infty)}(E)$ the space of E-valued càdlàg processes with the usual Skorokhod topology. On $\mathcal{M}_F(\mathbb{K})$, we consider the Polish topology of weak convergence induced by the topology on $\mathbb{K}$.

Definition 2.2.4. For a sequence $\left((\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})\right)_n \subset \mathfrak{C}$ of characteristics and a characteristic $(\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$, we write $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)}) \underset{n}{\longrightarrow} (\boldsymbol{\gamma}, \Pi)$ if and only if

$$\lim_n \boldsymbol{\gamma}^{(n)} \cdot \nabla g(\mathbf{0}) + \int_{\mathcal{Z}} g(\boldsymbol{z}) \ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) = \boldsymbol{\gamma} \cdot \nabla g(\mathbf{0}) + \int_{\mathcal{Z}} g(\boldsymbol{z}) \ \mathrm{d}\Pi(\boldsymbol{z}) \tag{2.6}$$

for all $g \in C_b(\mathbb{R}^2)$ differentiable in $\mathbf{0}$ such that $g(\mathbf{0}) = 0$.

This definition is equivalent to the convergence of $\gamma^{(n)}$ to γ and the vague convergence of $\Pi^{(n)}$ to Π.

Theorem 2.2.5. *The set of $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot processes is closed under limits of characteristics in $\mathfrak{C}$ in the following sense: if $\sigma_0^{(n)}$ converges weakly to some $\sigma_0 \in \mathcal{M}_F(\mathbb{K}) \setminus \{0\}$ and if $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)}) \to (\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$, then the sequence of $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})$-Fleming-Viot processes started in $\sigma_0^{(n)}$ converges weakly in $\mathbb{D}_{[0,+\infty)}\big(\mathcal{M}_F(\mathbb{K})\big)$ to the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process started in σ_0.*

Additionally, it is of interest to understand the relation between the $(\boldsymbol{\gamma}, \Pi)$- and the classical Λ-Fleming-Viot model. The next result illustrates that any Λ-Fleming-Viot process can be obtained by letting the jump measure concentrate on the diagonal $z_d = z_b$.

Theorem 2.2.6. *Let $\left((\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})\right)_n \subset \mathfrak{C}$ be a sequence of characteristics and Λ be a finite measure on $[0, 1]$. Assume that*

$$\lim_n \int_{\mathcal{Z}} g(\boldsymbol{z}) \ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) = \int_0^1 g(u, u) \ \frac{\mathrm{d}\Lambda(u)}{u^2} \tag{2.7}$$

for all $g \in C_b(\mathbb{R}^2)$ vanishing in a neighbourhood of the origin and that

$$\lim_{\epsilon \downarrow 0} \lim_n \int_{\mathcal{Z} \cap [0,\epsilon]^2} \left(\frac{z_b}{1 - z_d + z_b} \right)^2 \ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) = \Lambda(\{0\}). \tag{2.8}$$

Let $\left((N_0^{(n)}, \rho_0^{(n)})\right)_n \subset (0,+\infty) \times \mathcal{M}_1(\mathbb{K})$ *be a sequence with weak limit* $(1, \rho_0) \in (0,+\infty) \times \mathcal{M}_1(\mathbb{K})$*, and denote by* $(N^{(n)}, \rho^{(n)})$ *the* $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})$*-Fleming-Viot process started from* $(N_0^{(n)}, \rho_0^{(n)})$*. Then,* $(N^{(n)}, \rho^{(n)})$ *converges weakly to* $(1, \rho)$*, where* ρ *is a* Λ*-Fleming-Viot process started from* ρ_0*.*

For any given finite measure Λ on $[0,1]$, an approximating sequence of characteristics can be obtained by setting $\boldsymbol{\gamma}^{(n)} = 0$ and

$$\mathrm{d}\Pi^{(n)}(\mathrm{d}\boldsymbol{z}) = n^2\Lambda(\{0\})\delta_{\left(\frac{1}{n},\frac{1}{n}\right)}(\mathrm{d}\boldsymbol{z}) + \int_{1/n}^{1} \delta_{(u,u)}(\mathrm{d}\boldsymbol{z})\frac{\mathrm{d}\Lambda(u)}{u^2}\ \mathrm{d}u.$$

It is expected that one can prove corresponding statements to Theorem 2.2.5 and Theorem 2.2.6 in terms of the induced genealogies.

Genealogy and duality As in our case, the process, as an infinite-population limit, contains no notion of individuals, we will strengthen its biological interpretation by considering the model as a limit of individual-based models. In Section 2.7.1, we prove the convergence of an individual-based lookdown process to our model based on [DK99b, Theorem 3.2]. This allows us to view the coalescent originating from tracing the lowest levels in the lookdown construction as providing a notion of genealogy to the model. A construction based on marked metric measures spaces, which would extend the lookdown framework in [Guf18] and [Bla+23], goes beyond the scope of the present work.

The coalescent based on tracing the lowest levels, provides a *dual process* to the type distribution process. Indeed, the lookdown construction from Section 2.6.1 implies a pathwise moment duality conditional on the reproduction events (see [JK14, Definition 4.6, Section 4 Remark (Lookdown Construction)]). However, due to the dependence on the population size, the type distribution is not Markovian on its own, so that the duality is less exploitable. For this reason, we analyse two fundamental properties of the coalescent, *fixation* and *coming down from infinity*, by studying the genealogy through the lookdown process. We say that fixation occurs at time t if the corresponding type distribution is a Dirac measure on a single type. Through the graphical construction, this is equivalent to proving that at time

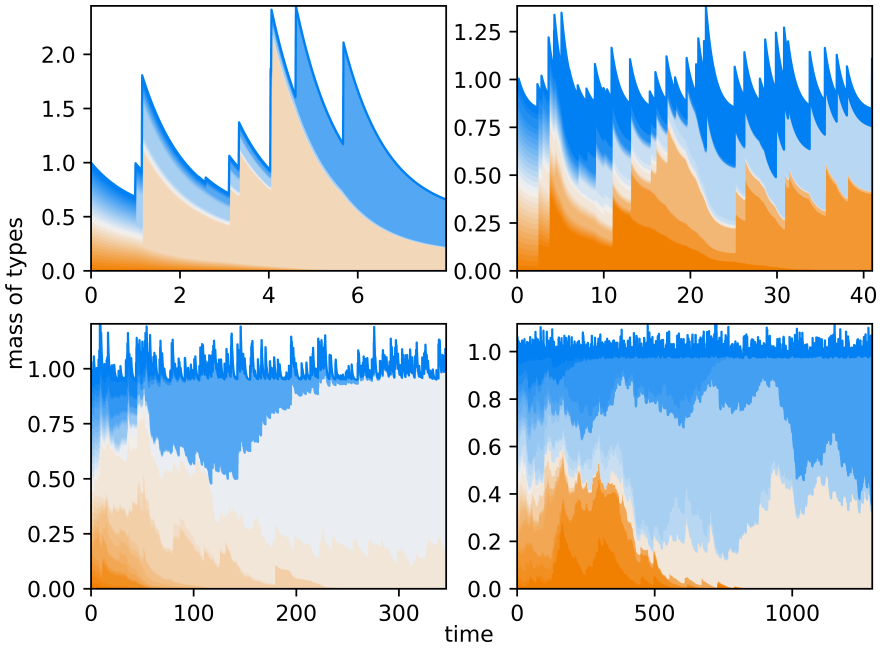


Figure 2.2: Illustration of Theorem 2.2.6 in the case of a finite number of types. The limiting process is a multi-type Wright-Fisher diffusion, the Λ-Fleming-Viot process with $\Lambda = \delta_0$. On the left-hand side of each figure, we start with 20 different types of equal proportions, each represented by a different colour. From the top left-hand figure to the bottom right-hand figure, we have used $\Pi_i := \delta_0 \otimes \mathrm{Unif}([0, k_i])$, where $k_1 = 1, k_2 = 1/3, k_3 = 1/10, k_4 = 1/20$. The drift $\boldsymbol{\gamma}_i$ is given by $(\gamma_i)_d = 1$ and

$$(\gamma_i)_b = 1 - \int_{\mathcal{Z}} z_b \, \mathrm{d}\Pi(\boldsymbol{z}) = 1 - \frac{k_i}{2}.$$

In each of the figures, the time frame shown corresponds to the appropriate rescaling of time given by Equation (2.8) as

$$\tau^i(t) = \left(\int_{[0,k_i]} \left(\frac{z_b}{1+z_b} \right)^2 \mathrm{d}z_b \right)^{-1} t.$$

t for every $n \in \mathbb{N}$, the n-th individual is a descendant of the lowest level individual. We say that fixation occurs if the above random time is a.s. finite. Weakening the above, we say that quasi-fixation occurs if for every $n \in \mathbb{N}$, there exists an a.s. finite time until the lowest n levels are descendants from the lowest level.

Theorem 2.2.7 (Quasi-fixation)**.** *Consider the graphical construction of the* $(\boldsymbol{\gamma}, \Pi)$*-Fleming-Viot process from Section 2.6. Assume that the population size process* N_t *is ergodic (cf. Remark 2.3.3) and that the marginal* Π_b *of* Π *governing the births does*

not only charge zero. Then, for any $n \geq 2$, after an a.s. finite time, the individuals associated to the lowest n levels will be descendants of the individual associated to the lowest label.

Coming down from infinity describes the property that the coalescent started from infinitely many lineages contains only finitely many lineages after a finite amount of time. Fixation cannot occur if the associated coalescent does not come down from infinity. To complement the above, we show that the coalescent *contains dust* at all times, i.e. that the coalescent started from infinitely many lineages contains infinitely many singletons at all times and does not come down from infinity. For Λ-coalescents which are not star-shaped, i.e. with $\Lambda(\{1\}) = 0$, this has been shown in the acclaimed paper [Pit99] to be equivalent to

$$\int_0^1 \frac{\Lambda(\,\mathrm{d}u)}{u} < \infty. \tag{2.9}$$

Heuristically, as $u^{-2}\Lambda(du)$ corresponds to the marginals of Π, the first moment condition (2.1) ensures that the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process satisfies Equation (2.9).

Theorem 2.2.8 (The coalescent contains dust)**.** *Consider the graphical construction of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process from Section 2.6. Then, for any time $t > 0$, the probability for the lowest lineage to not have had offspring before time t is positive.*

If Π_d does not charge 1, this guarantees that, at every time t, there are a.s. infinitely many ancestral lineages that do not coalesce in the time interval $[0, t]$.

2.2.3 Discussion

The $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot model is a natural extension of the Λ-Fleming-Viot paradigm. To make this connection apparent, let us consider the following intermediate example between the Λ- and the full $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot framework.

For small $u \approx 0$, let ν^u be a distribution on $[0, +\infty)$ with mean u. In the following, $\nu^u[f(z_b)]$ will denote the expectation of f w.r.t. $z_b \sim \nu^u$. The population

size process of the $(\mathbf{0}, \delta_u \otimes \nu^u)$-Fleming-Viot model, then, has the generator

$$\begin{aligned}\mathcal{L}^u f(N) &= \nu^u\Big[f\Big((1-u)N + z_b^u\Big) - f(N)\Big] \\ &\approx \nu^u\Big[f(N + z_b^u) - uNf'(N + z_b^u) - f(N)\Big] \\ &\approx -uNf'(N) + \nu^u\Big[f(N + z_b^u) - f(N)\Big],\end{aligned}$$

where we make use of the fact that $\mathbb{E}[f'(N + z_b^u)] \approx f'(N)$ as $u \downarrow 0$. Whereas the first term is of order u, this is not necessarily true for the second. Instead, we will make the natural assumption that it behaves in the limit like a subordinator. This is the same as asking that there are $\gamma_b \geq 0$ and a σ-finite measure Π_b on $(0, +\infty)$ satisfying $\int_0^{+\infty} z_b \ \mathrm{d}\Pi_b(z_b) < +\infty$ such that

$$\lim_{u \downarrow 0} \frac{1}{u} \nu^u[g(z_b^u)] = \gamma_b g'(0) + \int_0^{+\infty} g(z_b) \ \mathrm{d}\Pi_b(z_b) \tag{2.10}$$

for all $g : \mathbb{R} \to \mathbb{R}$ smooth enough and satisfying $g(0) = 0$. This is a particular case of Equation (2.6). Consequently, if we consider a Lévy subordinator $(L_t)_{t \geq 0}$ with characteristics (γ_b, Π_b), the measures $\nu^u := Law(L_u)$ satisfy Equation (2.10). Putting everything together, we obtain as a special case of Theorem 2.2.5 that

$$\mathcal{L}^u f(N) \approx u \cdot \left((\gamma_b - N)f'(N) + \int_0^{+\infty} f(N + z_b) - f(N) \ \mathrm{d}\Pi_b(z_b)\right). \tag{2.11}$$

Adding a non-constant death proportion does not alter the above substantially. The only part, that becomes more subtle, is to exclude the possibility of deriving a diffusion in the limit as $z_b - z_d N$ can be both positive and negative. This can be seen heuristically by the following Taylor expansion (that assumes both z_d and z_b to be of order u):

$$\begin{aligned}&\int_{\mathcal{Z}} f(N - z_d N + z_b) - f(N) \ \mathrm{d}\Pi^u(\boldsymbol{z}) \\ &\approx \int_{\mathcal{Z}} \Big[(z_b - z_d N)f'(N) + (z_b - z_d N)^2 f''(N)\Big] \ \mathrm{d}\Pi^u(\boldsymbol{z}) \\ &= \left(\int_{\mathcal{Z}} (z_b - z_d N) \ \mathrm{d}\Pi^u(\boldsymbol{z})\right) \cdot f'(N) + \left(\int_{\mathcal{Z}} (z_b - z_d N)^2 \ \mathrm{d}\Pi^u(\boldsymbol{z})\right) \cdot f''(N).\end{aligned} \tag{2.12}$$

If Π^u is a probability measure, then the first integral is of order u whereas the second is of order u^2. In particular, whenever the first integral is non-zero, there is no way to speed up the dynamics to see a diffusion. For a more complete treatment,

see also [JS03, Chapter VII, Section 3]. In the Λ-Fleming-Viot paradigm, this does not pose a problem as $N \equiv 1$ constantly, so that for $z_b \equiv u \equiv z_d$ the first order term vanishes. In the setting of a fluctuating population size, however, the first term can only vanish for at most one value of N. More precisely, under the balance condition (2.2), it vanishes exactly for $N = 1$. One should therefore see a diffusion appear only in the setting where N is forced to stay infinitesimally close to 1, i.e. in the setting of Theorem 2.2.6, leading back to the Λ-Fleming-Viot model.

The particular choice of the balance condition (2.2) is not relevant. The fraction

$$K := \frac{\gamma_b + \int_{\mathcal{Z}} z_b \ \mathrm{d}\Pi(\boldsymbol{z})}{\gamma_d + \int_{\mathcal{Z}} z_d \ \mathrm{d}\Pi(\boldsymbol{z})} \tag{2.13}$$

captures the *carrying capacity* of the system to which the population experiences a mean-reverting drift. If we consider a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process (N, ρ) not satisfying Equation (2.2), with K as in Equation (2.13), then we may define

$$\widehat{N}_t := \frac{N_t}{K}, \qquad \widehat{\gamma} := \left(\gamma_d, \frac{\gamma_b}{K}\right) \quad \text{and} \quad \int_{\mathcal{Z}} g(\boldsymbol{z}) \ \mathrm{d}\widehat{\Pi}(\boldsymbol{z}) := \int_{\mathcal{Z}} g\left(z_d, \frac{z_b}{K}\right) \mathrm{d}\Pi(\boldsymbol{z})$$

to obtain a $(\widehat{\gamma}, \widehat{\Pi})$-Fleming-Viot process $(\widehat{N}, \rho)$ that has carrying capacity 1.

Tracing the lowest levels in the lookdown construction provides the genealogy for a single fixed time. However, if one considers several points in time, the information contained in the labels about survival times, induces a bias in the genealogies. The framework developed in [Guf18] and [Bla+23] fuses genealogies in the sense of marked metric measure spaces with the lookdown space, allowing sampling at several points in time. It seems possible and enticing to carry out a similar construction for the present model in future work.

2.3 The population size process

Unsurprisingly, the population size process is crucial to the study of $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot processes. Hereafter, we will collect basic properties that we will use several times in the remainder of the chapter.

Let $(\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$ be a characteristic. We will consider the operator from Equation (2.4)

$$\mathcal{L}^{\boldsymbol{\gamma},\Pi} f(N) = (\gamma_b - \gamma_d N) f'(N) + \int_{\mathcal{Z}} \Big(f\Big((1 - z_d)N + z_b\Big) - f(N)\Big) \, \mathrm{d}\Pi(\boldsymbol{z}),$$

defined for any $f \in \mathcal{C}_c^2\Big((0, +\infty)\Big)$.

Theorem 2.3.1 (Well-Posedness)**.** *The martingale problem for $\mathcal{L}^{\boldsymbol{\gamma},\Pi}$ is well-posed. Furthermore, the unique solution for a given initial condition is a conservative strong Markov process on $(0, +\infty)$. Furthermore, if we let*

$$\tau_\epsilon := \inf\{t \geq 0 \ : \ N_t \notin (\epsilon, 1/\epsilon) \text{ or } N_{t-} \notin (\epsilon, 1/\epsilon)\}, \tag{2.14}$$

for all finite time horizons $T \geq 0$,

$$\lim_{\epsilon \downarrow 0} \mathbb{P}\Big(\tau_\epsilon \leq T\Big) = 0. \tag{2.15}$$

If $\boldsymbol{L}$ is a Lévy process of characteristic $(\boldsymbol{\gamma}, \mathbf{0}, \Pi)$, the solution can be constructed as the strong solution to

$$\mathrm{d}N_t = \Big(- N_{t-}, \ 1\Big) \cdot \mathrm{d}\boldsymbol{L}_t.$$

Equivalently, one obtains the process by taking a Poisson point process ξ on $[0, +\infty) \times \mathcal{Z}$ with mean intensity $\mathrm{d}t \otimes \Pi(\mathrm{d}\boldsymbol{z})$ and considering the unique strong solution to

$$\mathrm{d}N_t = (\gamma_b - \gamma_d N_t)\mathrm{d}t + \int_{\mathcal{Z}} (z_b - z_d N_{t-}) \, \xi(\mathrm{d}t, \mathrm{d}\boldsymbol{z}).$$

Furthermore, the process, viewed as a process on $\mathbb{R}$, is a Feller process if and only if the first marginal Π_d has no atom on 1. Here, we understand a Feller process to have a semigroup that maps the space of continuous functions vanishing at infinity into itself.

Proof. Since the extension of $\mathcal{L}^{\boldsymbol{\gamma},\Pi}$ to $C_c^2(\mathbb{R})$ maps into $C_b(\mathbb{R})$, [Kur11, Theorem 2.3] ensures that any solution to the martingale problem on the one-point-compactification of $\mathbb{R}$ is a weak solution to the SDE

$$\mathrm{d}N_t = \Big(- N_{t-}, 1\Big) \cdot \mathrm{d}\boldsymbol{L}_t$$

and vice-versa. As the coefficient is globally Lipschitz, there exists a pathwise unique, non-exploding, strong solution to this SDE. Weak uniqueness follows and we may conclude the well-posedness of the martingale problem.

Next, we prove Equation (2.15). First, assume that Π_d does not charge 1. Enumerating the jumps of the driving noise $\boldsymbol{L}$, we may bound N_t a.s. by

$$e^{-\gamma_d t}\left(\prod_{t_i\leq t}\left(1-z_d^i\right)\right)N_0 \leq N_t \leq N_0+\gamma_b t+(0,1)\cdot \boldsymbol{L}_t.$$

Since the lower bound is non-increasing and the upper bound is non-decreasing, it suffices to show that the lower bound is strictly positive and the upper bound is finite. For the upper bound, this is immediate. For the lower bound, we note that the product is nonzero if and only if $\sum_{t_i\leq t} z_d^i$ is finite, which follows from

$$\mathbb{E}\left[\sum_{t_i\leq t} z_d^i\right] = t\int_{\mathcal{Z}} z_d \ \mathrm{d}\Pi(\boldsymbol{z}) < +\infty. \tag{2.16}$$

If Π_d charges 1, events with $z_d = 1$ occur at a finite rate due to Equation (2.1). We write $I_T = \{i\in\mathbb{N} : t_i \leq T, z_d^i = 1\}$ for the finite set of such events before the horizon T. If $i\in I_T$, then N_{t_i} has the law $\nu_b = \Pi(\{1\}\times\cdot)/\Pi(\{1\}\times(0,+\infty))$ and is independent of both $(N_s)_{0\leq s<t_i}$ and the values of z_d^i, $i\notin I_T$. As the set I_T is finite,

$$N_* := \min\left(\{N_0\}\cup\left\{N_{t_i} : i\in I_T\right\}\right),$$

is almost surely positive. Hence, we have the strictly positive lower bound

$$N_t \geq e^{-\gamma_d T}\left(\prod_{\substack{t_i\leq T\\ i\notin I_T}}(1-z_d^i)\right)\cdot N_*.$$

Finally, we apply [Küh18, Corollary 1.2] which states that the Feller property for the process seen as a process on $\mathbb{R}$ is equivalent to

$$\lim_{|N|\to+\infty}\Pi\left\{\boldsymbol{z}\in\mathbb{R}^2 \ : \ z_b - Nz_d \in -B(N,r)\right\} = 0$$

for all $r>0$, where $B(N,r)$ denotes the ball of radius r around N. This is exactly the case if the first marginal of Π does not charge the point 1. □

The restriction for the process to be Feller becomes apparent, when noting that an atom of Π_d at 1 allows the process to jump from any point in $(0, +\infty)$ to the distribution ν_b defined in the proof above. In particular, functions vanishing at infinity will not be mapped onto functions that vanish at infinity. Similarly, the Feller property cannot hold for the population size process viewed as a process on $(0, +\infty)$ (in which case the test functions vanish also at zero). Indeed, from any infinitesimal $N_0 \downarrow 0$, we may jump with a fixed rate to a macroscopic population size, as the deaths are proportional to the current population size, whereas the offspring size is not.

The next result will be of practical interest when proving different tightness results.

Corollary 2.3.2 (Compact Containment)**.** *Let $(\boldsymbol{\gamma}^{(\alpha)}, \Pi^{(\alpha)})_\alpha \subset \mathfrak{C}$ be a family of characteristics satisfying*

$$\sup_\alpha \left(\gamma_b^{(\alpha)} + \int_{\mathcal{Z}} z_b \ \mathrm{d}\Pi^{(\alpha)}(\boldsymbol{z}) \right) < +\infty.$$

Furthermore, suppose that the family $(N_0^{(\alpha)})_\alpha$ of $(0, +\infty)$-valued random variables is tight. Let $(N_t^{(\alpha)})_{t \geq 0}$ be the population size process with characteristic $(\boldsymbol{\gamma}^{(\alpha)}, \Pi^{(\alpha)})$ started in $N_0^{(\alpha)}$. Then, for any finite time horizon $T \geq 0$ and any threshold $\epsilon > 0$, there exists a compact set $\Gamma_{T,\epsilon} \subseteq (0, +\infty)$ such that

$$\inf_\alpha \mathbb{P}\Big(N_t^{(\alpha)} \in \Gamma \text{ for all } t \in [0, T] \Big) \geq 1 - \epsilon.$$

Proof. Let $\boldsymbol{L}^{(\alpha)}$ denote the Lévy subordinator corresponding to $N^{(\alpha)}$. As the strong solution to

$$\mathrm{d}N_t^{(\alpha)} = \big(-N_{t-}, 1 \big) \cdot \mathrm{d}\boldsymbol{L}^{(\alpha)},$$

the process $N^{(\alpha)}$ is almost surely bounded from above by the Lévy subordinator $L_t^{(\alpha,2)} := (0, 1) \cdot \boldsymbol{L}^\alpha$ of characteristic $\big(\gamma_b^{(\alpha)}, \Pi_b^{(\alpha)}\big)$. Since subordinators are monotonically increasing, we may bound $L_t^{(\alpha,2)}$ from above by $L_T^{(\alpha,2)}$ for any $t \in [0, T]$. Now,

$$\mathbb{E}\left[L_T^{(\alpha,2)} \right] = T \left(\gamma_b^{(\alpha)} + \int_{\mathcal{Z}} z_b \ \mathrm{d}\Pi^{(\alpha)}(\boldsymbol{z}) \right) \leq C_T$$

for a constant C_T depending only on T. This yields

$$\mathbb{P}\left(\sup_{t\in[0,T]} N_t^{(\alpha)} \leq \frac{2C_T}{\epsilon}\right) \geq \mathbb{P}\Big(L_T^{(\alpha,2)} \leq \frac{2C_T}{\epsilon}\Big) \geq 1 - \mathbb{P}\left(L_T^{(\alpha,2)} \geq \frac{2C_T}{\epsilon}\right) \geq 1 - \frac{\epsilon}{2}.$$

Together with Equation (2.15), this implies the result for some compact interval of the form $\Gamma_{T,\epsilon} = \left[\delta_{T,\epsilon}, \frac{2C_T}{\epsilon}\right]$. □

$$\begin{aligned}
&\mathcal{L}^{\boldsymbol{\gamma},\Pi} f(N) \\
&= \Big(\gamma_b - \gamma_d N\Big) + \int_{\mathcal{Z}} \left[-z_d N + z_b\right] \, \mathrm{d}\Pi(\boldsymbol{z}) \\
&= -\Big(\gamma_d + \int_{\mathcal{Z}} z_d \, \mathrm{d}\Pi(\boldsymbol{z})\Big) N + \gamma_b + \int_{\mathcal{Z}} z_b \, \mathrm{d}\Pi(\boldsymbol{z})
\end{aligned}$$

Remark 2.3.3 (Ergodicity)**.** For the proof of quasi-fixation (cf. Theorem 2.2.7), we will assume that the long term behaviour of the population size is well-behaved. More precisely, it is enough for the process to be ergodic, which is expected to hold true under weak conditions on the jump measure. The interested reader may refer to [Kul09] and the references therein for relevant results. Since this is not the main focus of this chapter, it is left for future work.

2.4 Closedness in the space of characteristics

This section is concerned with the proof of Theorem 2.2.5. Recall that $(\sigma_t)_{t\geq 0}$ is a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process if and only if it is a $\mathcal{M}_F(\mathbb{K})$-valued solution to the martingale problem for $\mathcal{L}^{\boldsymbol{\gamma},\Pi}$ defined in Equation (2.3). We will make use of the uniqueness result from Section 2.6, which does not depend on this section. For the rest of this section, we consider a weakly convergent sequence $\sigma_0^{(n)} \to \sigma_0 \in \mathcal{M}_F(\mathbb{K})\backslash\{0\}$ and a sequence of characteristics $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)}) \to (\boldsymbol{\gamma}, \Pi) \in \mathfrak{C}$ converging in the sense of Definition 2.2.4. Let $\sigma^{(n)}$ denote the $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})$-Fleming-Viot process started from $\sigma_0^{(n)}$. The proof is divided into two lemmas.

Lemma 2.4.1 (Tightness)**.** *The sequence $(\sigma^{(n)})_n$ is tight.*

Proof. Corollary 2.3.2 guarantees the compact containment condition for the population size process, and thus for the process σ, as sets of measures on a compact space that are uniformly bounded by a constant are compact. By [EK86, Theorem 3.9.4], it is then sufficient to show the uniform boundedness of the generator in expectation over finite time intervals. For any fixed $\sigma \in \mathcal{M}_F(\mathbb{K})$ and $\Phi \in C_b^1\mathcal{M}_F(\mathbb{K})$, we have by Equation (2.1) and Equation (2.3)

$$\begin{aligned}\left|\mathcal{L}^{\gamma^{(n)},\Pi^{(n)}}\Phi(\sigma)\right| &\leq \left(\gamma_b^{(n)} + \gamma_d^{(n)}\sigma(\mathbb{K})\right)\|D\Phi\|_\infty \\ &\quad + \int_{\mathcal{Z}} \|D_\Phi\|_\infty\left(z_d^{(n)}\sigma(\mathbb{K}) + z_b^{(n)}\right)\, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \\ &\leq 2C\|D\Phi\|_\infty \cdot \left(1 + \sigma(\mathbb{K})\right).\end{aligned}$$

We conclude by noting that the corresponding population size process $N_t^{(n)} = \sigma_t^{(n)}(\mathbb{K})$ satisfies

$$\mathbb{E}\left[N_t^{(n)}\right] \leq \mathbb{E}\left[N_0^{(n)}\right] + Ct,$$

which is uniformly bounded on any compact time interval. □

Lemma 2.4.2. *Any limit point of the family $(\sigma^n)_{n\geq 1}$ is a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process.*

Proof. The first few lines are similar to [Eth+24, Proposition 6.6]. By [EK86, Theorem 4.8.10], it suffices to show

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\Phi(\sigma_{t+\tau}^{(n)}) - \Phi(\sigma_t^{(n)}) - \int_t^{t+\tau} \mathcal{L}^{\boldsymbol{\gamma},\Pi}\Phi(\sigma_s^{(n)})\, \mathrm{d}s\right)\prod_{i=1}^k h_i(\sigma_{t_i}^{(n)})\right] = 0 \qquad (2.17)$$

for all $k \geq 0, 0 \leq t_1 < t_2 < ... < t_k \leq t < t + \tau$ and $h_1, ..., h_k \in C_b\mathcal{M}_F(\mathbb{K})$. Equation (2.17) is satisfied in the prelimit in the sense that

$$\mathbb{E}\left[\left(\Phi(\sigma_{t+\tau}^{(n)}) - \Phi(\sigma_t^{(n)}) - \int_t^{t+\tau} \mathcal{L}^{\boldsymbol{\gamma}^{(n)},\Pi^{(n)}}\Phi(\sigma_s^{(n)})\, \mathrm{d}s\right)\prod_{i=1}^k h_i(\sigma_{t_i}^{(n)})\right] = 0$$

for every n. It is therefore sufficient to prove that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_t^{t+\tau} \left|\mathcal{L}^{\boldsymbol{\gamma}^{(n)},\Pi^{(n)}}\Phi(\sigma_s^{(n)}) - \mathcal{L}^{\boldsymbol{\gamma},\Pi}\Phi(\sigma_s^{(n)})\right|\, \mathrm{d}s \prod_{i=1}^k h_i(\sigma_{t_i}^{(n)})\right] = 0.$$

Using the tower property of conditional expectation, this statement is satisfied provided

$$\lim_{n\to\infty} \mathbb{E}\left(\int_t^{t+\tau} \left| \mathcal{L}^{\gamma^{(n)},\Pi^{(n)}}\Phi(\sigma_s^{(n)}) - \mathcal{L}^{\gamma,\Pi}\Phi(\sigma_s^{(n)}) \right| \, \mathrm{d}s \;\middle|\; \mathcal{F}_t \right) = 0.$$

Let us introduce the family of functions

$$f_\Phi^\sigma : \mathbb{R}^2 \to \mathbb{R}, \quad (z_1, z_2) \mapsto \int_{\mathbb{K}} \Phi\Big((1-z_2)\sigma + z_1\delta_\kappa\Big) - \Phi(\sigma)\frac{\mathrm{d}\sigma(\kappa)}{\sigma(\mathbb{K})},$$

which are continuously differentiable and satisfy $f_\Phi^\sigma(\mathbf{0}) = 0$. Then,

$$\nabla f_\Phi^\sigma(\mathbf{0}) = D\Phi\left(\sigma, \frac{\sigma}{\sigma(\mathbb{K})}\right)\Big(1, -\sigma(\mathbb{K})\Big)$$

and thus

$$(\gamma_b - \gamma_d\sigma(\mathbb{K}))D\Phi\Big(\sigma, \frac{\sigma}{\sigma(\mathbb{K})}\Big) = \boldsymbol{\gamma}\cdot\nabla f_\Phi(\mathbf{0}).$$

For $\sigma \in \mathcal{M}_F(\mathbb{K})$, the difference between the prelimiting and limiting generator is given by

$$\begin{aligned}
&\left|\mathcal{L}^{\boldsymbol{\gamma}^{(n)},\Pi^{(n)}}\Phi(\sigma) - \mathcal{L}^{\boldsymbol{\gamma},\Pi}\Phi(\sigma)\right| \\
&\leq \left| \Big(\gamma_b^{(n)} - \gamma_d^{(n)}\sigma(\mathbb{K})\Big)D\Phi\Big(\sigma, \frac{\sigma}{\sigma(\mathbb{K})}\Big) - \Big(\gamma_b - \gamma_d\sigma(\mathbb{K})\Big)D\Phi\Big(\sigma, \frac{\sigma}{\sigma(\mathbb{K})}\Big) \right. \\
&\qquad + \int_{\mathcal{Z}}\int_{\mathbb{K}} \left[\Phi\Big((1-z_d)\sigma + z_b\delta_\kappa\Big) - \Phi(\sigma)\right]\frac{\mathrm{d}\sigma(\kappa)}{\sigma(\mathbb{K})}\, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \\
&\qquad \left. - \int_{\mathcal{Z}}\int_{\mathbb{K}} \left[\Phi\Big((1-z_d)\sigma + z_b\delta_\kappa\Big) - \Phi(\sigma)\right]\frac{\mathrm{d}\sigma(\kappa)}{\sigma(\mathbb{K})}\, \mathrm{d}\Pi(\boldsymbol{z}) \right| \\
&\leq \left| \boldsymbol{\gamma}^{(n)}\cdot\nabla f_\Phi(\mathbf{0}) + \int_{\mathcal{Z}} f_\Phi(\boldsymbol{z})\, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) - \boldsymbol{\gamma}\cdot\nabla f_\Phi(\mathbf{0}) - \int_{\mathcal{Z}} f_\Phi(\boldsymbol{z})\, \mathrm{d}\Pi(\boldsymbol{z}) \right|.
\end{aligned}$$

Hence, by the convergence of the characteristics, we obtain the pointwise limit

$$\lim_{n\to\infty}\left|\mathcal{L}^{\boldsymbol{\gamma}^{(n)},\Pi^{(n)}}\Phi(\sigma) - \mathcal{L}^{\boldsymbol{\gamma},\Pi}\Phi(\sigma)\right| = 0.$$

Writing $N := \sigma(\mathbb{K})$ for the population size, we have

$$\begin{aligned}
\left|\mathcal{L}^{\boldsymbol{\gamma}^{(n)},\Pi^{(n)}}\Phi(\sigma) - \mathcal{L}^{\boldsymbol{\gamma},\Pi}\Phi(\sigma)\right| &\leq \|D\Phi\|_\infty\Big(\sup_n|\gamma_b^{(n)} - \gamma_b| + \sup_n|\gamma_d - \gamma_d^{(n)}|N\Big) \\
&\quad + \sup_n 2\|D\Phi\|_\infty \int_{\mathcal{Z}}(z_b + z_d)\, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}),
\end{aligned}$$

which has finite mean. The above is an integrable majorant and we may apply Lebesgue's Dominated Convergence to conclude. □

2.5 Relation to the Λ-Fleming-Viot process

This section is devoted to the proof of Theorem 2.2.6, which illustrates that $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot processes can approximate Λ-Fleming-Viot processes arbitrarily well. We prove the result in two steps. First, we prove that the population size concentrates on the constant function 1 by viewing it as the solution to an SDE with large drift. Then, we verify the convergence of the type distribution. More precisely, making use of [BES04, Lemma 5.3], it is enough to show the convergence of the type composition up to the stopping time $\tau_\delta := \inf\{t \geq 0 \ : \ N_t \leq \delta \text{ or } N_{t-} \leq \delta\}$.

Lemma 2.5.1. *Let $\left((\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})\right)_n \subset \mathfrak{C}$ be a sequence of characteristics satisfying the assumptions of Theorem 2.2.6. Then the associated sequence $(N^{(n)})_n$ of population size processes converges weakly to 1 in $\mathbb{D}_{[0,+\infty)}(\mathbb{R})$.*

Lemma 2.5.2. *Let $\left((\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})\right)_n \subset \mathfrak{C}$ be a sequence of characteristics satisfying the assumptions of Theorem 2.2.6 for a finite measure Λ on $[0,1]$, and suppose that $(N_0^{(n)}, \rho_0^{(n)}) \to (1, \rho_0)$ weakly. Denote by $\rho^{(n)}$ the associated sequence of measure-valued type processes and write ρ for the Λ-Fleming-Viot process started in ρ_0. For every $\delta > 0$, the sequences of processes $\rho^{(n)}$ stopped at τ_δ converges weakly to ρ stopped at τ_δ.*

As the total population size process converges to a deterministic function, Slutsky's Lemma ensures that the convergence of the components suffices for the joint convergence of $(N^{(n)}, \rho^{(n)})$ and Theorem 2.2.6 follows.

Remark 2.5.3. There is a subtlety in this application of Slutsky's Lemma that is worth pointing out more clearly: Slutsky's Lemma yields the convergence of the sequence in the product topology on $\mathbb{D}_{[0,+\infty)}((0,+\infty)) \times \mathbb{D}_{[0,+\infty)}([0,1])$ which is weaker than the Skorokhod topology on $\mathbb{D}_{[0,+\infty)}((0,+\infty) \times [0,1])$ ([JS03, VI.1.21, p.329]). In order to conclude that $(N^{(n)}, \rho^{(n)})$ does converge also in the latter, it suffices to note that the continuity of the limiting process strengthens the convergence automatically to uniform convergence on compact time intervals. Then, we make use of the fact that there is no difference between the two spaces from the perspective of uniform convergence on compacts.

Proof of Lemma 2.5.1. We distinguish two cases:

1. Suppose $\int_0^1 u^{-1} d\Lambda(u) < +\infty$ and $\sup_n |\boldsymbol{\gamma}^{(n)}| < +\infty$. Then, necessarily

$$\gamma_d^{(n)} - \gamma_b^{(n)} \to 0$$

and the convergence follows from the closedness of the model, see Theorem 2.2.5.

2. If $\int_0^1 u^{-1} d\Lambda(u) = \infty$ or $\sup_n |\boldsymbol{\gamma}^{(n)}| = \infty$, the proof relies on [Kat91] which considers stochastic processes that are forced onto a manifold by a large drift. In this case, the sequence $C_n := \gamma_b^{(n)} + \int_{\mathcal{Z}} z_b \ \mathrm{d}\Pi^{(n)}(\boldsymbol{z})$ diverges to infinity. By Theorem 2.3.1, we have that the population size process is a solution to

$$N_t^{(n)} = N_0^{(n)} + C_n \int_0^t (1 - N_t^{(n)}) \ \mathrm{d}t + \int_0^t \Big(- N, 1 \Big) \cdot \ \mathrm{d}\boldsymbol{L}_t^{(n)},$$

where $\boldsymbol{L}^{(n)}$ is a Lévy process with characteristic $\Big(- \int_{\mathcal{Z}} \boldsymbol{z} \ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}), \boldsymbol{0}, \Pi^{(n)} \Big)$. In this setting, the second term forces the process onto the singleton manifold $\{1\}$. The process $\boldsymbol{L}^{(n)}$ is a martingale with quadratic variation

$$[\boldsymbol{L}^{(n)}]_t = t \int_{\mathcal{Z}} \boldsymbol{z}^2 \ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}),$$

which is uniformly bounded in n, and thus uniformly integrable. In particular, [Kat91, Condition (4.2)] is satisfied for $Y = \boldsymbol{L}^{(n)}$ by choosing $\delta = \infty$ and $M = \boldsymbol{L}^{(n)}$. This implies that [Kat91, Condition (C5.1)] holds, whereas [Kat91, Condition (C5.2)] follows automatically. We apply [Kat91, Theorem 6.2] to obtain the convergence $N^{(n)} \to 1$. □

Proof of Lemma 2.5.2. The generator of the Fleming-Viot process is given by

$$A_1 \Phi(\rho) := \frac{1}{2} \int_{\mathbb{K}} \int_{\mathbb{K}} D^2 \Phi(\rho; \delta_\kappa, \delta_\chi) \ (\delta_\kappa - \rho) \, (\mathrm{d}\chi) \ \rho(\mathrm{d}\kappa),$$

where D^2 denotes the second order Fréchet derivative. The generator is defined for all functions $\Phi \in \mathcal{C}(\mathcal{M}_1(\mathbb{K}))$ for which the second derivative exists for all

$\kappa, \chi \in \mathbb{K}$ and is bounded and continuous. The generator of the Λ-Fleming-Viot is $A\Phi(\rho) = A_1^\Lambda \Phi(\rho) + A_2^\Lambda \Phi(\rho)$, where

$$A_1^\Lambda := \Lambda(\{0\}) A_1 \text{ and } A_2^\Lambda \Phi(\rho) := \int_0^1 \int_{\mathbb{K}} \left[\Phi\Big((1-u)\rho + u\delta_\kappa\Big) - \Phi(\rho) \right] \, \mathrm{d}\rho(\kappa) \frac{\mathrm{d}\Lambda(u)}{u^2}.$$

Note that the solution to the martingale problem associated to A is unique (which can be shown using a duality relationship, see [LB03, Section 5.2] and the general explanation in [Eth00, Section 1.6]). Recall that the generator of the type distribution process $\rho^{(n)}$ is given by

$$\begin{aligned} A^{(n)}\Phi(N,\rho) &= \int_{\mathcal{Z}} \int_{\mathbb{K}} \Phi\Big((1-\overline{\boldsymbol{z}})\rho + \overline{\boldsymbol{z}}\delta_\kappa\Big) - \Phi(\rho) \; \mathrm{d}\rho(\kappa) \; \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \\ &= \int_{\mathcal{Z}\cap[0,\epsilon]^2} \int_{\mathbb{K}} \Phi\Big((1-\overline{\boldsymbol{z}})\rho + \overline{\boldsymbol{z}}\delta_\kappa\Big) - \Phi(\rho) \; \mathrm{d}\rho(\kappa) \; \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \\ &\quad + \int_{\mathcal{Z}\setminus[0,\epsilon]^2} \int_{\mathbb{K}} \Phi\Big((1-\overline{\boldsymbol{z}})\rho + \overline{\boldsymbol{z}}\delta_\kappa\Big) - \Phi(\rho) \; \mathrm{d}\rho(\kappa) \; \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \\ &=: A_1^{(n,\epsilon)} + A_2^{(n,\epsilon)}. \end{aligned}$$

with $\overline{\boldsymbol{z}} := \frac{z_b}{(1-z_d)N+z_b}$ for functions $\Phi \in C_b^1(\mathcal{M}_1(\mathbb{K}))$ depending only on ρ. Adapting [BK93, Theorem 2.3] to our situation, the statement follows from

1. the sequence $(\rho^{(n)})_n$ is tight in $\mathbb{D}_{[0,T]}(\mathcal{M}_1(\mathbb{K}))$,

2. and for all $\Phi : \mathcal{M}_1(\mathbb{K}) \to \mathbb{R}$ such that Φ is three times Fréchet differentiable with bounded derivatives,

$$\sup_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} \left\| A^{(n)}\Phi(N_t^{(n)}, \cdot) \right\|_\infty \right] \leq C_\Phi \tag{2.18}$$

with $0 < C_\Phi < +\infty$, and for any $t \in [0,T]$

$$\lim_{n\to\infty} \mathbb{E}\left[\left\| A^{(n)}\Phi(N_t^{(n)}, \cdot) - A\Phi \right\|_\infty \right] = 0. \tag{2.19}$$

These conditions present a slight modification of [BK93, Theorem 2.3], where we have used that the three times Fréchet differentiable functions are dense in the domain of A and that the proof remains valid if Equation (2.18) and Equation (2.19) hold in expectation. Now, tightness of $(\rho^{(n)})_n$ follows from Equation (2.18) and the compactness of $\mathcal{M}_1(\mathbb{K})$ in combination with [EK86, Theorem 3.9.4], so that we will concentrate on proving Equation (2.18) and Equation (2.19).

Applying Taylor's expansion to the inner part, we obtain

$$\begin{aligned}&A_1^{(n,\epsilon)}\Phi(N,\rho)\\&= \int_{\mathcal{Z}\cap[0,\epsilon]^2}\int_{\mathbb{K}}\left[\overline{\boldsymbol{z}}\cdot D\Phi(\rho;\delta_\kappa-\rho)+\overline{\boldsymbol{z}}^2\cdot\frac{1}{2}D^2\Phi(\rho;\delta_\kappa-\rho,\delta_\kappa-\rho)+\mathcal{O}\left(\overline{\boldsymbol{z}}^3\right)\right]\\&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\mathrm{d}\rho(\kappa)\ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}).\end{aligned}$$

The first summand is equal to zero due to the identity

$$\int_{\mathbb{K}} D\Phi(\rho;\delta_\kappa-\rho)\ \mathrm{d}\rho(\kappa)=\int_{\mathbb{K}}\int_{\mathbb{K}}\left[D\Phi(\rho;\delta_\kappa)-D\Phi(\rho;\delta_\chi)\right]\ \mathrm{d}\rho(\kappa)\ \mathrm{d}\rho(\chi)=0.$$

For the second summand, we can use

$$\begin{aligned}&\int_{\mathbb{K}} D^2\Phi(\rho;\delta_\kappa-\rho,\delta_\kappa-\rho)\ \mathrm{d}\rho(\kappa)\\&=\int_{\mathbb{K}}\left[D^2\Phi(\rho;\delta_\kappa,\delta_\kappa-\rho)-D^2\Phi(\rho;\rho,\delta_\kappa-\rho)\right]\ \mathrm{d}\rho(\kappa)\\&=\int_{\mathbb{K}}\int_{\mathbb{K}} D^2\Phi(\rho;\delta_\kappa,\delta_\chi)\ (\delta_\kappa-\rho)(\mathrm{d}\chi)\ \mathrm{d}\rho(\kappa)\\&\qquad-\int_{\mathbb{K}}\int_{\mathbb{K}}\int_{\mathbb{K}}\left[D^2\Phi(\rho;\delta_\eta,\delta_\kappa)-D^2\Phi(\rho;\delta_\eta,\delta_\theta)\right]\ \mathrm{d}\rho(\eta)\ \mathrm{d}\rho(\theta)\ \mathrm{d}\rho(\kappa)\\&=\int_{\mathbb{K}}\int_{\mathbb{K}} D^2\Phi(\rho;\delta_\kappa,\delta_\chi)\ (\delta_\kappa-\rho)(\mathrm{d}\chi)\ \mathrm{d}\rho(\kappa)=A\Phi(\rho)\end{aligned}$$

and the third can be bounded by

$$\frac{\epsilon}{(1-\epsilon)^3}\frac{1}{N^3}\cdot\sup_n\int_{\mathcal{Z}} z_b^2\ \mathrm{d}\Pi^{(n)}(\boldsymbol{z}).$$

Altogether, we can write

$$A_1^{(n,\epsilon)}\Phi(N,\rho)=\left(\int_{\mathcal{Z}\cap[0,\epsilon]^2}\overline{\boldsymbol{z}}^2\ \mathrm{d}\Pi^{(n)}(\boldsymbol{z})\right)\cdot A_1\Phi(\rho)+\mathcal{O}\left(\frac{\epsilon}{N^3}\right),$$

which is bounded up to the stopping time τ_δ. The boundedness of $A_2^{(n,\epsilon)}$ is immediate from the fact that $\Pi^{(n)}$ is finite outside a neighbourhood of 0:

$$|A_2^{(n,\epsilon)}\Phi(N,\rho)|\le 2\|\Phi\|_\infty\sup_n\int_{\mathbb{Z}\setminus[0,\epsilon]^2}\mathrm{d}\Pi^{(n)}(\boldsymbol{z})<+\infty.$$

To prove Equation (2.19) we first bound

$$\begin{aligned}&\lim_n\mathbb{E}\left[\left\|A^{(n)}\Phi(N^{(n)},\cdot)-A\Phi\right\|_\infty\right]\\&\le\limsup_{\epsilon\downarrow 0}\lim_n\mathbb{E}\left[\left\|A_1^{(n,\epsilon)}\Phi-A_1^{\Lambda}\Phi\right\|_\infty+\left\|A_2^{(n,\epsilon)}\Phi-A_{2,2}^{(\Lambda,\epsilon)}\Phi\right\|_\infty+\left\|A_{2,1}^{(\Lambda,\epsilon)}\Phi\right\|_\infty\right],\end{aligned}\tag{2.20}$$

where we define

$$A_{2,1}^{(\Lambda,\epsilon)}\Phi(\rho) := \int_0^\epsilon \int_{\mathbb{K}} \left[\Phi\Big((1-u)\rho + u\delta_\kappa\Big) - \Phi(\rho)\right] \, \mathrm{d}\rho(\kappa)\frac{\mathrm{d}\Lambda(u)}{u^2},$$
$$A_{2,2}^{(\Lambda,\epsilon)}\Phi(\rho) := \int_\epsilon^1 \int_{\mathbb{K}} \left[\Phi\Big((1-u)\rho + u\delta_\kappa\Big) - \Phi(\rho)\right] \, \mathrm{d}\rho(\kappa)\frac{\mathrm{d}\Lambda(u)}{u^2}.$$

The first summand of the inequality (2.20) is bounded by

$$\begin{aligned}
&\left|A_1^{(n,\epsilon)}\Phi(N,\rho) - A_1^{\Lambda}\Phi(\rho)\right| \\
&\leq \left| \int_{\mathcal{Z}\cap[0,\epsilon]^2} \left[\left(\frac{z_b}{(1-z_d)N_t^{(n)} + z_b}\right)^2 - \left(\frac{z_b}{1-z_d+z_b}\right)^2 \right] \, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \right| \left\|A_1\Phi\right\|_\infty \\
&\quad + \left| \int_{\mathcal{Z}\cap[0,\epsilon]^2} \left(\frac{z_b}{1-z_d+z_b}\right)^2 \, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) - \Lambda(\{0\}) \right| \left\|A_1\Phi\right\|_\infty + \mathcal{O}\left(\frac{\epsilon}{N^3}\right)
\end{aligned}$$

Here, the second part vanishes due to the assumption on $\Lambda(\{0\})$. The first and the third parts can be treated up to τ_δ by using the convergence of the population size and by noting that

$$\begin{aligned}
&\left| \int_{\mathcal{Z}\cap[0,\epsilon]^2} \left[\left(\frac{z_b}{(1-z_d)N_t^{(n)} + z_b}\right)^2 - \left(\frac{z_b}{1-z_d+z_b}\right)^2 \right] \, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \right| \\
&\leq \left| \int_{\mathcal{Z}\cap[0,\epsilon^2]} \frac{2z_b|1-N_t^{(n)}|}{(1-\epsilon)^2 N_t^{(n)}} \cdot \frac{z_b}{1-\epsilon}\left(\frac{(1+2\epsilon)}{(1-\epsilon)N_t^{(n)}} + 1\right) \, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}) \right| \\
&\leq \frac{2|1-N_t^{(n)}|}{(1-\epsilon)^3 N_t^{(n)}}\left(\frac{(1+2\epsilon)}{(1-\epsilon)N_t^{(n)}} + 1\right) \sup_n \int_{\mathcal{Z}\cap[0,\epsilon]^2} z_b^2 \, \mathrm{d}\Pi^{(n)}(\boldsymbol{z}).
\end{aligned}$$

The convergence of the second summand of inequality (2.20) can be achieved in a similar fashion, adding and subtracting a term without the population size and using Taylor's expansion. The third summand of the inequality (2.20) can be expanded similarly to $A_1^{(n,\epsilon)}$ above and is bounded because of the boundedness of the derivatives of Φ. Taking limits, the last term vanishes as well. □

2.6 Lookdown construction

In this section, we will consider a lookdown construction of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process, which represents the process as an infinite system of particles ([DK99b]). We will abbreviate the martingale problem for the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process as FV-MP and for the lookdown process as L-MP. We have two major aims:

Aim 1 Show that the FV-MP and L-MP are well-posed, i.e. existence and uniqueness of solutions.

Aim 2 Show that the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process has a representation as a system of SDEs and a graphical representation.

A key ingredient will be the acclaimed Markov Mapping Theorem (first described in [Kur98] and further developed in e.g. [KN11; EK19]), which essentially tells us the following: suppose that the operators of the L-MP and FV-MP are consistent in an appropriate sense. Then, the existence of solutions of the FV-MP implies the existence of solutions to the L-MP. In turn, if uniqueness holds for the L-MP, uniqueness also holds for the FV-MP. With this tool at hand, we address our goals as follows:

Section 2.6.1 We introduce the lookdown construction, the system of SDEs (2.21) and its graphical representation. We show that strong solutions to the system (2.21) exist and that they solve the L-MP. Our lookdown process/L-MP is a special case of the one considered in [DK99b, Section 3.3], for which uniqueness has already been proven [DK99b, Section 4]. As any weak solution to (2.21) solves the L-MP, the uniqueness for the L-MP implies weak uniqueness for (2.21) and also that the solution to the L-MP is given by the unique strong solution to (2.21). This already shows the existence of solutions to the L-MP without invoking the Markov Mapping Theorem.

Section 2.6.2 We show existence of a solution to the FV-MP through an approximation scheme, cf. Lemma 2.6.4.

Section 2.6.3 We apply the Markov Mapping Theorem: together with the existence of solutions to the FV-MP this provides an alternative proof for the existence of solutions to the L-MP. More crucially, together with the uniqueness from Section 2.6.1, the uniqueness transfers to the FV-MP and well-posedness follows for both martingale problems.

Note that we do not follow the approach of [EK19] or [Bla+23], which consider a lookdown containing additionally the system of SDEs in order to prove well-posedness of the L-MP by an analysis of the SDEs. Instead, we use the "shortcut" that uniqueness is known for the L-MP, in order to conclude that the unique solution can be represented as the strong solution of a system of SDEs (2.21).

2.6.1 An infinite system of SDEs and its graphical representation

We will first describe the system of SDEs in order to provide a *fil d'Ariane*. Whereas the lookdown process itself comes from [DK99b, Section 3.3], we follow the more modern presentation of [Bir+05], where the case of the usual Λ-Fleming-Viot process has been treated.

The main idea will be to represent the type distribution ρ as the DeFinetti measure of an exchangeable sequence $\mathsf{k} = (\mathsf{k}(i))_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ of types from individuals sampled from the population. In order to describe the evolution of the population, we need three pieces of information: how and when do reproduction events occur? which individuals participate in a reproduction event? and who is the parent?

One advantage of the exchangeability of the sample is that we may choose the parent freely. The descriptive term *lookdown* comes from the fact that the individual, which has the lowest label i from all of those that participate in the reproduction event, is chosen as the parent. This renders the system coherent for different numbers of individuals: as the parental choice is unchanged, one can add more individuals without changing the dynamics on the already present, lowest label individuals. The two other parts are incorporated through a Poisson point process ξ on $[0,+\infty)\times\mathcal{Z}\times[0,1]^{\mathbb{N}}$ with mean intensity $\mathrm{d}t\otimes\Pi(\mathrm{d}\boldsymbol{z})\otimes\mathrm{d}u^{\otimes\mathbb{N}}$, where $\mathcal{Z}$ denotes the event space from Definition 2.2.1, $\mathrm{d}t$ is the Lebesgue measure on $[0,+\infty)$ and $\mathrm{d}u$ denotes the Lebesgue measure on $[0,1]$. The interpretation is as follows: if $(t,\boldsymbol{z},\mathsf{u})\in\xi$, a reproduction event $\boldsymbol{z}$ falls at time t. Then, every individual $i\in\mathbb{N}$ participates in the reproduction event exactly if $\mathsf{u}_i \le \overline{\boldsymbol{z}}$, where $\overline{\boldsymbol{z}}(N,\boldsymbol{z}) = \frac{z_b}{(1-z_d)N+z_b}$ denotes the effective impact at the current population size N, cf. Equation (2.5)

following. In order to ensure the exchangeability of the sample, however, we will also shift the individuals that did not participate in the reproduction event onto the next free slot. This is captured by the infinite system of SDEs

$$N_t = N_0 + \int_0^t \left(\gamma_b - \gamma_d N_s\right)\, \mathrm{d}s + \int_{[0,t]\times\mathcal{Z}\times\mathbb{U}} \left(z_b - z_d N_{s-}\right)\, \xi(\mathrm{d}s, \mathrm{d}\boldsymbol{z}, \mathrm{d}\mathsf{u}), \quad (2.21a)$$

$$\mathsf{k}_t(i) = \mathsf{k}_0(i) + \int_{[0,t]\times\mathcal{Z}\times\mathbb{U}} \left(\mathsf{k}_{s-}\left(\theta_{J(\overline{\boldsymbol{z}},\mathsf{u})}(i)\right) - \mathsf{k}_{s-}(i)\right)\, \xi(\mathrm{d}s, \mathrm{d}\boldsymbol{z}, \mathrm{d}\mathsf{u}) \qquad \forall i \in \mathbb{N}, \quad (2.21b)$$

where $J(\overline{\boldsymbol{z}}, \mathsf{u}) = \{j \in \mathbb{N} \,:\, \mathsf{u}_j \leq \overline{\boldsymbol{z}}\}$ denotes the set of individuals participating in the reproduction event and

$$\theta_J(i) := \begin{cases} i & \text{if } i < \min J \\ \min J & \text{if } i \in J \\ i - |J \cap [i]| + 1 & \text{otherwise} \end{cases}$$

describes the lookdown rule. Here and in the following, we write $[n]$ for the set of integers $\{1, \ldots, n\}$ and $\mathbb{U} := [0,1]^{\mathbb{N}}$ for notational convenience. To stress the time dependence of the effective impact, we will occasionally write $\overline{\boldsymbol{z}}_{s-} = \frac{z_b}{(1-z_d)N_{s-}+z_b}$.

The remainder of this section is dedicated to giving a graphical construction of strong solutions to Equation (2.21). This graphical construction has the major advantage of encoding all genealogical information in such a way that types of individuals can be added later by tracing the lines in the direction of the arrows.

First, let N be the unique strong solution to the SDE Equation (2.21a). In particular, N_t is a measurable function of N_0 and the restriction of ξ to $[0,t] \times \mathcal{Z}$. We build a graphical representation on the time-level picture $[0, +\infty) \times \mathbb{N}$ as follows. If $(t, \boldsymbol{z}, \mathsf{u}) \in \xi$, define the set $J(\boldsymbol{z}, \mathsf{u}) := \{j \in \mathbb{N} \,:\, \mathsf{u}_j \leq \overline{\boldsymbol{z}}\}$ as before and draw for every $i \in \mathbb{N}$ an arrow from (t, i) to $\left(t, \theta_{J(\overline{\boldsymbol{z}},\mathsf{u})}(i)\right)$ if $i \neq \theta_{J(\overline{\boldsymbol{z}},\mathsf{u})}(i)$. Note here that $\overline{\boldsymbol{z}}$ is a measurable function of N_{t-} and therefore of N_0 and the restriction of ξ to $[0,t] \times \mathcal{Z} \times \mathbb{U}$.

In order to build a solution to Equation (2.21), it is enough to mark the points $(0, i)$ with $\mathsf{k}_0(i)$ and to update the types as follows: up to the next time t at which an arrow leaves (t, i), the type $\mathsf{k}(i)$ stays constant; at the arrow it adopts the type of the individual at which the arrow points; repeat. Crucially, we will

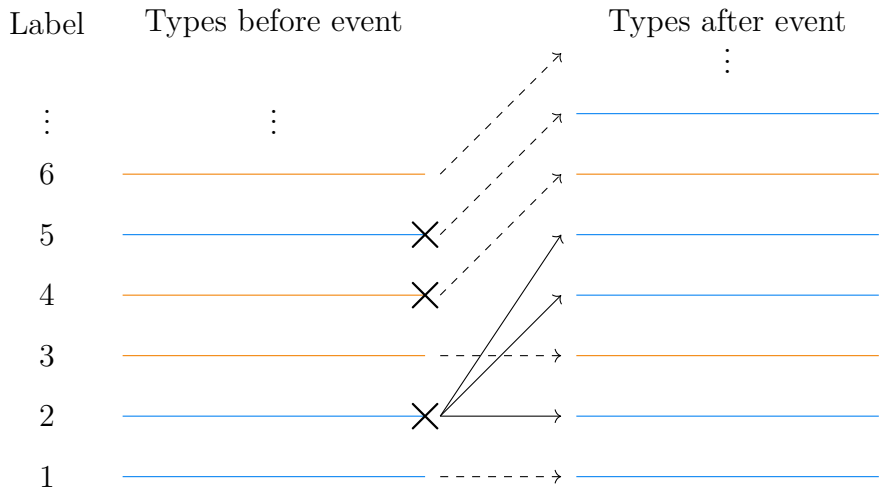


Figure 2.3: Illustration of a reproduction event: during the event, the particles with labels $2, 4, 5$ are chosen to participate. Offspring of the parental type are inserted. All other levels have to shift upwards accordingly. Types are marked by colors, drawn arrows indicate birth, dashed arrows the necessary shift.

show that every individual only changes its type a finite number of times in every finite time horizon. For illustration purposes, Figure 2.3 shows the graphical representation with only two types.

Lemma 2.6.1. *For every initial condition* $(N_0, \mathsf{k}_0) \in (0, +\infty) \times \mathbb{K}^{\mathbb{N}}$ *independent of* ξ*, there exists a strong solution to Equation* (2.21)*.*

Proof. Consider the graph constructed from ξ and define for $n \in \mathbb{N}$ the process

$$L_n(t) := \int_{\mathcal{Z}\times\mathbb{U}\times[0,t]} \sum_{i=1}^{n} \mathbb{1}_{[0,\overline{\boldsymbol{z}}_{s-}]}(\mathsf{u}_i)\, \xi(\mathrm{d}\boldsymbol{z}, \mathrm{d}\mathsf{u}, \mathrm{d}s),$$

which is either infinite or jumps exactly when there is at least one individual in $[n]$ that participates in a reproduction event. In particular, as long as L_n is finite, every individual $i \in [n]$ changes type at most a finite number of times.

Now, writing ξ_0 for the projection of ξ onto $[0, +\infty) \times \mathcal{Z}$, we have that τ_ϵ as

defined in Equation (2.14) is measurable w.r.t. N_0 and ξ_0. We can bound

$$\begin{aligned}\mathbb{E}[L_n(t\wedge\tau_\epsilon)] &= \mathbb{E}\left[\int_{[0,t\wedge\tau_\epsilon]\times\mathcal{Z}} n\overline{\boldsymbol{z}}_{s-}\ \xi_0(\mathrm{d}s,\mathrm{d}\boldsymbol{z})\right]\\ &\le n\cdot\mathbb{E}\left[\int_{[0,t]\times\mathcal{Z}}\frac{z_b}{(1-z_d)\epsilon+z_b}\ \xi_0(\mathrm{d}s,\mathrm{d}\boldsymbol{z})\right]\\ &\le tn\int_{\mathcal{Z}}\frac{z_b}{(1-z_d)\epsilon+z_b}\ \mathrm{d}\Pi(\boldsymbol{z})\\ &\le tn\int_{\mathcal{Z}}\mathbb{1}_{[1/2,1]}(z_d)+\frac{2z_b}{\epsilon}\ \mathrm{d}\Pi(\boldsymbol{z})<+\infty,\end{aligned}$$

and deduce

$$\mathbb{P}\Big(L_n(t)=+\infty\Big)\le\mathbb{P}\left(\bigcap_{\epsilon\downarrow0}\{\tau_\epsilon<t\}\right)=\lim_{\epsilon\downarrow0}\mathbb{P}(\tau_\epsilon<t)=0.$$

We conclude that every individual a.s. changes its type only a finite number of times. This in turn ensures that k is a strong solution to Equation (2.21) when considering (N,ξ) as the driving noise. Since N is also a strong solution, we conclude that (N,k) is a strong solution to Equation (2.21) started from (N_0,k_0). □

The finiteness of the jump rates implies that the solution is weakly unique, but this will follow anyway from the following consideration. First, Itô's formula implies that any weak solution to Equation (2.21) is also a solution to the lookdown martingale problem (L-MP) for $(A,\mathcal{D}(A))$ defined by

$$\begin{aligned}\mathcal{D}(A):=\Bigg\{&(N,\mathsf{k})\mapsto h(N,\mathsf{k}):=f(N)\prod_{i=1}^{n}g_i(\mathsf{k}_i)\\ &:\ f\in\mathcal{C}_b^1((0,+\infty)),g_i\in\mathcal{C}(\mathbb{K}),1\le i\le n,n\in\mathbb{N}\Bigg\}\end{aligned}\tag{2.22}$$

and

$$\begin{aligned}&Ah(N,\mathsf{k})\\ &=(\gamma_b-\gamma_dN)\frac{f'(N)}{f(N)}h(N,\mathsf{k})\\ &\quad+\int_{\mathcal{Z}}\sum_{J\subseteq[n]}\overline{\boldsymbol{z}}^{|J|}(1-\overline{\boldsymbol{z}})^{n-|J|}\cdot\Big(h\Big((1-z_d)N+z_b,\theta_J(\mathsf{k})\Big)-h(N,\mathsf{k})\Big)\ \mathrm{d}\Pi(\boldsymbol{z}),\end{aligned}\tag{2.23}$$

where $\theta_J(\mathsf{k}):=\mathsf{k}\circ\theta_J$ as in Equation (2.21). The division by $f(N)$ should be interpreted as cancelling the corresponding factor in $h(N,\mathsf{k})$.

Uniqueness for the L-MP for A has been shown in [DK99b, Section 4] using the following idea: up to stopping times, the jump rates of the levels stay finite, resulting in the uniqueness of the stopped martingale problems. If the sequence of stopping times is increasing and exhausts the complete space, [EK86, Theorem 4.6.2] yields the uniqueness of the unstopped martingale problem. In our setting, an example for an exhausting sequence is given by the τ_ϵ of Theorem 2.3.1. As the argument is based on 'localisation', it carries over to the local L-MP. We are temporarily working with local martingale problems, as we will use a localized version of the Markov Mapping Theorem ([KN11, Corollary 3.3]) in Lemma 2.6.5 (the reason will become apparent in its proof).

Combing that weak solutions to Equation (2.21) solve the L-MP for A with the uniqueness for the local L-MP for A implies that

1. weak uniqueness holds for the system (2.21) and

2. the unique solution to the local martingale problem for A is given by the strong solution to (2.21), which resulted from the graphical construction.

This gives us the following result:

Lemma 2.6.2 (See e.g. [DK99b, Section 4])**.** *The local L-MP for* $(A, \mathcal{D}(A))$ *is well-posed and the unique solution can be obtained as the strong solution to Equation* (2.21)*. In particular, every solution to the local L-MP is a solution to the L-MP.*

2.6.2 Existence of a solution for the FV-martingale problem

We first prove that paths of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process are càdlàg and the existence of solutions for the FV-MP. Then, we will apply the Markov Mapping Theorem to transfer the uniqueness of the L-MP to the uniqueness of the FV-MP. This shows the well-posedness of the FV-MP.

Lemma 2.6.3. *Every* $(\boldsymbol{\gamma}, \Pi)$*-Fleming-Viot process has a càdlàg modification.*

Proof of Lemma 2.6.3. Note that $B = (0, +\infty) \times \mathcal{M}_1(\mathbb{K})$ is separable and that $C_b^1(B)$ is separating and contains a countable set that separates points. If (N, ρ) is a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot, then N is the population size process studied in Section 2.3 and thus

$$\mathbb{P}\Big((N_t, \rho_t) \in \Gamma_{T,\epsilon} \times \mathcal{M}_1(\mathbb{K}) \text{ for all } t \in [0, T]\Big) \geq 1 - \epsilon$$

for all $T, \epsilon > 0$ and a compact $\Gamma_{T,\epsilon}$ given by Corollary 2.3.2. Hence, [EK86, Theorem 4.3.6] yields the existence of a *càdlàg* modification. □

Since any modification of a solution to a martingale problem is a solution itself, we may and will assume in the following that any solution is *càdlàg*.

Lemma 2.6.4. *For any given initial condition $(N_0, \rho_0) \in S$, there exists a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process started from (N_0, ρ_0).*

Proof. Existence is immediate as soon as the jump measure Π is finite, as a solution may be constructed as a piecewise deterministic Markov Process. For general characteristics $(\boldsymbol{\gamma}, \Pi)$, set $\Pi^{(n)}(\boldsymbol{z}) := \mathbb{1}_{\|\boldsymbol{z}\| > 1/n}\Pi(\boldsymbol{z})$ and choose $\boldsymbol{\gamma}^{(n)}$ such that $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)}) \in \mathfrak{C}$ and $\boldsymbol{\gamma}^{(n)} \to \boldsymbol{\gamma}$. Then, $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)}) \to (\boldsymbol{\gamma}, \Pi)$ in the sense of characteristics and we may apply Theorem 2.2.5 to obtain a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process as the limit of $(\boldsymbol{\gamma}^{(n)}, \Pi^{(n)})$-Fleming-Viot processes, see also Section 2.4. □

2.6.3 Application of the Markov Mapping Theorem

Next, we show that every $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process corresponds to a solution to the L-MP for A defined in Equation (2.23) through an application of the Markov Mapping Theorem ([Kur98; KN11; EK19]). Together with Lemma 2.6.4, Lemma 2.6.2 and the uniqueness of the L-MP, this shows the well-posedness of the FV-MP. In the following, we will refer to $B := (0, +\infty) \times \mathcal{M}_1(\mathbb{K})$ as the base state space and to $E := (0, +\infty) \times \mathbb{K}^{\mathbb{N}}$ as the extended state space. First, define $\gamma : E \to B$ by

$$\gamma(N, \mathsf{k}) := \lim_n \left(N, \frac{1}{n}\sum_{i=1}^{n} \delta_{\mathsf{k}(i)}\right)$$

when the limit exists and by $\gamma(N, \mathsf{k}) = (N, \delta_\kappa)$ for some fixed, but arbitrary $\kappa \in \mathbb{K}$ otherwise. Conversely, we define the Markov kernel α from B to E by

$$\alpha(N, \rho; \mathrm{d}N', \mathrm{d}\mathsf{k}) := \delta_N(\mathrm{d}N') \otimes \rho^{\otimes\mathbb{N}}(\mathrm{d}\mathsf{k})$$

that maps every type distribution ρ to a random exchangeable sample with DeFinetti measure ρ. The maps γ, α are chosen in way that

$$\alpha\Big((N, \rho), \gamma^{-1}(N, \rho)\Big) = 1$$

for $(N, \rho) \in B$. This is one of the prerequisites of the Markov mapping theorem [KN11, Corollary 3.3]. The Markov Mapping Theorem has the basic underlying aim that an extended Markov process is projected to another Markov process. This property ensures that the projection of the extended space E to the base space B is consistent with this aim.

Lemma 2.6.5. *Every solution $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process $(\hat{N}, \hat{\rho})$ started from some (N_0, ρ_0) corresponds to a solution (N, k) of the L-MP for $\Big(A, \alpha(N_0, \rho_0, \cdot\,)\Big)$ in the sense that $(N, \rho) := \gamma \circ (N, \mathsf{k})$ has the same law as $(\hat{N}, \hat{\rho})$ as elements in $\mathbb{D}_{[0,+\infty)}(B)$. Furthermore, for every $t \geq 0$, the distribution of (N_t, k_t) conditionally on $(N_s, \rho_s)_{0 \leq s \leq t}$ is given by $\alpha(N_t, \rho_t\, \cdot\,)$.*

Proof. In the following, we view α as the operator

$$\alpha h(N, \rho) := \int_E h(N', \mathsf{k})\; \alpha(N, \rho; \mathrm{d}N', \mathrm{d}\mathsf{k})$$

defined for all bounded measurable functions $h : E \to \mathbb{R}$. We obtain for the test function $h = f \prod_{i=1}^n g_i \in \mathcal{D}(A)$ that

$$\alpha h(N, \rho) = f(N) \prod_{i=1}^n \overline{g}_i(\rho),$$

where $\overline{g}_i(\rho) := \int_{\mathbb{K}} g_i(\kappa)\; \mathrm{d}\rho(\kappa)$. Then, applying α to the operator A shows

$\alpha A h(N, \rho)$

$$
\begin{aligned}
&= (\gamma_b - \gamma_d N)\frac{f'(N)}{f(N)}\alpha h(N,\rho) \\
&\quad + \sum_{J\subseteq[n]} \int_{\mathcal{Z}} f\Big((1-z_d)N + z_b\Big)\overline{\boldsymbol{z}}^{|J|}(1-\overline{\boldsymbol{z}})^{n-|J|}\left(\int_{\mathbb{K}}\prod_{i\in J} g_i(\kappa^*)\ \mathrm{d}\rho(\kappa^*)\right)\prod_{i\in[n]\setminus J}\overline{g}_i(\rho) \\
&\qquad\qquad - \alpha h(N,\rho)\ \mathrm{d}\Pi(\boldsymbol{z}) \\
&= (\gamma_b - \gamma_d N)\frac{f'(N)}{f(N)}\alpha h(N,\rho) \\
&\quad + \int_{\mathcal{Z}}\int_{\mathbb{K}}\Bigg[f\Big((1-z_d)N + z_b\Big)\prod_{i=1}^{n}\Big((1-\overline{\boldsymbol{z}})\overline{g}_i(\rho) + \overline{\boldsymbol{z}} g_i(\kappa^*)\Big) \\
&\qquad\qquad - \alpha h(N,\rho)\Bigg]\ \mathrm{d}\rho(\kappa^*)\ \mathrm{d}\Pi(\boldsymbol{z}) \\
&= (\gamma_b - \gamma_d N)\frac{f'(N)}{f(N)}\alpha h(N,\rho) \\
&\quad + \int_{\mathcal{Z}}\int_{\mathbb{K}} \alpha h\Big((1-z_d)N + z_b, (1-\overline{\boldsymbol{z}})\rho + \overline{\boldsymbol{z}}\delta_{\kappa^*}\Big) - \alpha h(N,\rho)\ \mathrm{d}\rho(\kappa^*)\ \mathrm{d}\Pi(\boldsymbol{z}),
\end{aligned}
$$

which is $\mathcal{L}^{\gamma,\Pi}$ applied to αh.

In order to conclude by invoking the Markov Mapping Theorem [KN11, Corollary 3.3], we need to check the regularity assumption [KN11, Condition 2.1] for the operator A. First, the domain $\mathcal{D}(A)$ is contained in the bounded continuous functions on E by definition, separates points, is stable under multiplication and contains the constant 1. Furthermore, $A1 = 0$ and in order to verify that A maps into the continuous functions, note that the integrand

$$
\overline{\boldsymbol{z}}^{|J|}(1-\overline{\boldsymbol{z}})^{n-|J|}\left(h\Big((1-z_d)N + z_b, \theta_J(\mathsf{k})\Big) - h(N,\mathsf{k})\right)
$$

is non-trivial only when $|J| \geq 2$. In this case, it is continuous in (N,k) for fixed $\boldsymbol{z}$, so that we may apply Lebesgue's Dominated Convergence with the integrable upper bound $z \mapsto \|h\|\overline{\boldsymbol{z}}$, where $\overline{\boldsymbol{z}}$ is computed w.r.t. a suitably chosen N. Next, we obtain the uniform bound

$$
\begin{aligned}
|Ah(N,\mathsf{k})| &\leq (\gamma_b + \gamma_d N)\|f'\|_\infty + 2^n\cdot 2\|h\|_\infty \int_{\mathcal{Z}} \overline{\boldsymbol{z}}\ \mathrm{d}\Pi(\boldsymbol{z}) \\
&\leq c_h \underbrace{\left(1 + N + \frac{1}{N}\right)}_{=:\psi(N,\mathsf{k})},
\end{aligned}
$$

where we use the same bound $\overline{\boldsymbol{z}} \leq \mathbb{1}_{[1/2,1]}(z_d) + \frac{2z_b}{N}$ as in the proof of Lemma 2.6.1.

Now, let (N, ρ) be a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process. Then, it is in particular a solution to the local $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot martingale problem. The same argument as in the proof of Lemma 2.6.1 yields that $\int_0^t \alpha\psi(N_s, \rho_s)\ \mathrm{d}s$ is almost surely finite for every $t \geq 0$. This is what necessitates local martingale problems; if we had instead a finite expectation of $\int_0^t \alpha\psi(N_s, \rho_s)\ \mathrm{d}s$, we could apply the standard Markov Mapping Theorem without the need for localisation. Then, [KN11, Corollary 3.3] implies the existence of a solution (N, k) to the local L-MP for A such that $(\hat{N}, \hat{\rho})$ has the same law on $\mathbb{D}_{[0,+\infty)}(B)$ as $(N, \rho) := \gamma \circ (N, \mathsf{k})$. Now, since the local L-MP for A is well-posed, and the unique solution is given by the strong solution to Equation (2.21), we conclude that (N, k) is a solution to the L-MP, which concludes the proof. □

Proof of Theorem 2.2.3. The uniqueness of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot martingale problem (FV-MP) is an immediate consequence of the uniqueness of the L-MP for A, which carries over to the averaged process by the Markov Mapping Theorem in the previous Lemma 2.6.5. Existence of solutions to the FV-MP had already been proven in Lemma 2.6.4. The strong Markov property, then, follows from the well-posedness by the general theory of martingale problem see e.g. [EK86, Section 4.4]. □

2.7 Genealogy and duality

As we argued in the introduction, in our case, it is more natural to refer to a coalescent of the infinite-population limit as its genealogy when it arises as the limit of the genealogies of the corresponding individual-based models. We will conclude this property for our model from the lookdown framework. Then, we will investigate fixation of types and whether the associated coalescent comes down from infinity.

2.7.1 Individual-based model

Fix a characteristic $(\boldsymbol{\gamma}, \Pi)$, an initial condition $N_0 > 0$ and $m \in \mathbb{N}$ which should be thought of as *carrying capacity* of the population. Heuristically, we will consider

the following model: we start with an initial population $N_0^m := \lfloor mN_0 \rfloor$ and impose two dynamics on the individuals:

Continuous part Each individual carries an independent Poisson clock with rate γ_d. When the clock rings, the individual dies. Independently, there is one Poisson clock with rate $m\gamma_b$. Whenever it rings, an individual is chosen at random to produce an offspring.

Discrete events Reproduction events fall as a Poisson point process on $[0, +\infty) \times \mathcal{Z}$ with intensity measure $\mathrm{d}t \otimes \Pi(\mathrm{d}\boldsymbol{z})$. At an event (t, z), $\lfloor N_t^m z_d \rfloor$ individuals die and $\lfloor mz_b \rfloor$ individuals are born and carry the type of an individual chosen uniformly at random from the population immediately before the event.

Events with $z_d < 1/N_t^m$ and $z_b < 1/m$ do not appear at all, so only a finite number of events occur on any compact time interval. In the following, we will assume the process to be stopped at the first time the population size $N_t^m = 0$ reaches zero.

To fit the above into the setting of [DK99b], we decompose the population process $N_t^m = N_0^m + B_t^m - D_t^m$ into births and deaths. The dynamics are described by the generator

$$\begin{aligned}
&\mathcal{L}^m f(n, b, d) \\
&= m\gamma_b \Big(f(n+1, b+1, d) - f(b, d) \Big) + n\gamma_d \Big(f(n-1, b, d+1) - f(b, d) \Big) \\
&\qquad + \int_{\mathcal{Z}} \Big(f(n - \lfloor nz_d \rfloor + \lfloor mz_b \rfloor, b + \lfloor mz_b \rfloor, d + \lfloor nz_d \rfloor) - f(n, b, d) \Big) \, \mathrm{d}\Pi(\boldsymbol{z}).
\end{aligned}$$

The first part is a continuous birth and death generator, whereas the second part implements reproduction events.

Similarly to the infinite population limit, we can construct this population model as a lookdown process, see [DK99b, Section 1.2]. More precisely, we start with an exchangeable sequence $\mathsf{k}_0^m \in \mathbb{K}^m$. If the population changes at time t, we denote by $k := \Delta B_t^m$ the number of births. Births work exactly as in the lookdown construction from Section 2.6.1: first, choose $k+1$ levels $i_0 < \cdots < i_k$ randomly from $1, \ldots, N_t^m$; then, insert individuals of the type $\mathsf{k}_{t-}^m(i_0)$ at levels $i_1, \ldots, i_k$ and

move all other individuals upwards. All individuals with level greater than N_t^m are considered dead, which automatically takes care of death events.

We wish to prove that this lookdown construction converges in a suitable sense to the one defined in Section 2.6.1. One verifies immediately that the the sequence $(N^m/m)_{m\in\mathbb{N}}$ converges in law to the population process N started from N_0. However, to apply [DK99b, Theorem 3.2], we need to prove a stronger convergence. More precisely, writing $[B^m]$ for the quadratic variation process of B^m, we define

$$U_t^m := \frac{[B^m]_t + B_t^m}{m^2}.$$

One then checks that $(N^m/m, U^m)$ converges weakly to (N, U), where

$$U_t = \int_{[0,t]\times\mathcal{Z}} z_b^2 \, \xi(\mathrm{d}t, \mathrm{d}\boldsymbol{z})$$

with ξ denoting the driving noise of N (see Theorem 2.3.1). Using the fact that N does not reach 0 in finite time together with the remark just after Equation (3.12) in [DK99b], we deduce the following result on genealogies.

Theorem 2.7.1 (see [DK99b, Theorem 3.2])**.** *Let $\mathsf{k}_0 \in \mathbb{K}^{\mathbb{N}}$ be an exchangeable sequence with DeFinetti measure ρ_0 and set $\mathsf{k}_0^m := (\mathsf{k}_0(i))_{1\le i\le m}$. Consider the lookdown process $(\mathsf{k}_t^m)_{t\ge 0}$ defined above and view it as a process on $\mathbb{K}^{\mathbb{N}}$ by setting $\mathsf{k}_t^m(i) := \mathsf{k}_0(i)$ if $\max_{s\le t} N_s^m \le i$ and $\mathsf{k}_t^m(i) := \mathsf{k}^m_{\tau_i^m(t)-}(i)$, where $\tau_i^m(t) := \sup\{s < t \,:\, N_s^m \ge i\}$ otherwise.*

Write $(N_t, \mathsf{k}_t)_{t\ge 0}$ for the lookdown process from Section 2.6.1 and let ρ_t be the DeFinetti measure of k_t. Then, setting

$$\rho_t^m := \frac{1}{N_t^m} \sum_{i=1}^{N_t^m} \delta_{\mathsf{k}_t^m(i)}$$

up to the extinction time and as zero afterwards, the sequence $\left(\frac{N^m}{m}, \frac{N^m}{m}\rho^m, \mathsf{k}^m\right)$ converges in law to $(N, N\rho, \mathsf{k})$. In particular, the genealogy of k approximates the genealogy of the individual-based model.

2.7.2 Fixation

We will first prove that the lowest levels of the lookdown process will eventually all be of the same type and deduce afterwards the equivalence to fixation of a certain type in the population.

Theorem 2.2.7 (Quasi-fixation)**.** *Consider the graphical construction of the* $(\boldsymbol{\gamma}, \Pi)$*-Fleming-Viot process from Section 2.6. Assume that the population size process* N_t *is ergodic (cf. Remark 2.3.3) and that the marginal* Π_b *of* Π *governing the births does not only charge zero. Then, for any* $n \geq 2$*, after an a.s. finite time, the individuals associated to the lowest* n *levels will be descendants of the individual associated to the lowest label.*

Proof. For two constants $0 < C_1 < C_2$ to be determined later and $i \geq 1$, we define $\sigma_0 := 0$ and the down- and upcrossing stopping times

$$\tau_i := \inf\{t \geq \sigma_{i-1} \ : \ N_t \leq C_1 \text{ or } N_{t-} \leq C_1\},$$
$$\sigma_i := \inf\{t \geq \tau_i \ : \ N_t \geq C_2 \text{ or } N_{t-} \geq C_2\}.$$

Due to the ergodicity, a stationary distribution μ exists, satisfying

$$\frac{1}{T}\int_0^T \mathbb{1}_{[0,C_1]}(N_t)\ \mathrm{d}t \underset{T\to+\infty}{\longrightarrow} \mu([0,C_1])\ \ a.s.$$

Choose C_1 large enough such that $\mu([0,C_1]) > 0$. Then, the above implies that τ_1 is a.s. finite. Assume for the moment that σ_1 is a.s. finite. If $C_2 \to +\infty$, then so does σ_1. For any fixed $T_0 > 0$ and $\epsilon \in (0,1)$, we may choose $C_2 = C_2(T_0, \epsilon)$ large enough so that $\sigma_1 - \tau_1 > T_0$ with probability at least $1 - \epsilon$. Enumerate the point process $\xi = \{(t_i, \boldsymbol{z}^i, u^i)\}$ from the graphical construction. Then,

$$\begin{aligned}
&\mathbb{P}\Big(\text{There is a point } (t_i, \boldsymbol{z}^i, u^i) \text{ with } z_b^i \geq \delta \text{ in the time interval } [\tau_1, \sigma_1]\Big)\\
&\geq \mathbb{P}\Big(\text{There is a point } (t_i, \boldsymbol{z}^i, u^i) \text{ with } z_b^i \geq \delta \text{ in the time interval } [\tau_1, \tau_1 + T_0]\\
&\qquad \text{and } \sigma_1 - \tau_1 > T_0\Big)\\
&= 1 - \mathbb{P}(\text{There is no such point in } [\tau_1, \tau_1 + T_0]) - \mathbb{P}(\sigma_1 - \tau_1 \leq T_0)\\
&\geq 1 - \exp\Big(-T_0 \Pi_b\Big([\delta, +\infty)\Big)\Big) - \epsilon.
\end{aligned}$$

As $\Pi_b \neq 0$, we may choose $\delta > 0$ small enough such that $\Pi_b([\delta, +\infty))$ is strictly positive. Then, choose T_0 large enough that the above remains bounded away from 0. Finally, by Theorem 2.2.3, the strong Markov property extends this argument to the interval $[\tau_2, \sigma_2]$ etc. Hence, there will be a.s. infinitely many reproduction events with effective impact

$$\overline{\boldsymbol{z}}^i = \frac{z_b^i}{(1 - z_d^i)N_{t_i-} + z_b^i} \geq \frac{\delta}{C_2 + \delta} =: C > 0.$$

Consequently, the n lowest lineages are all effected by a common reproduction event with probability at least $C^n > 0$ a.s. infinitely often and we may conclude that the n lowest lineages are effected by a common reproduction event at an a.s. finite time.

Now, if $\sigma_i = +\infty$ with positive probability, we may proceed in a more straightforward way. Indeed, this will imply that the interval $[\tau_i, \sigma_i]$ is infinite on this event, and thus contains infinitely many reproduction events satisfying $\overline{\boldsymbol{z}}^i \geq C$. Therefore, it suffices to choose $\epsilon := 1 - \mathbb{P}(\sigma_i = +\infty)$ and T_0 large enough to compensate. □

Corollary 2.7.2. *Suppose that $\mathbb{K} \subseteq [0, 1]$ is closed. For a $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process, let $F_t(\kappa) := \rho_t([0, \kappa])$ denote the cumulative distribution function of ρ_t. Under the assumptions of Theorem 2.2.7, $F_t(\kappa)$ converges for every $\kappa \in \mathbb{K}$ a.s. to a random variable $F_\infty(\kappa)$ supported on $\{0, 1\}$ such that*

$$\mathbb{P}\left(\lim_{t\to+\infty} F_t(\kappa) = 1 \;\middle|\; F_0\right) = F_0(\kappa).$$

Proof. Without loss of generality, we may consider $\mathbb{K} = \{0, 1\}$, as $F_t(\kappa)$ splits the type space into two sets. We will write $w_t := \rho_t(\{1\})$ and it suffices to show that $w_t \to w_\infty$ a.s. for a Bernoulli random variable w_∞ with conditional mean w_0, which we may choose to be deterministic. For a test function $f = f(w)$ independent of N, we obtain

$$\mathcal{L}^{\boldsymbol{\gamma},\Pi} f(N, w) = \int_{\mathcal{Z}} \left[w f\Big((1 - \overline{\boldsymbol{z}})w + \overline{\boldsymbol{z}}\Big) + (1 - w) f\Big((1 - \overline{\boldsymbol{z}})w\Big) - f(w) \right] \mathrm{d}\Pi(\boldsymbol{z}).$$

In particular, w_t is a bounded martingale and we may conclude that it converges almost surely to its limit w_∞. Additionally,

$$w_t^n = \left(\int_{\mathbb{K}} \kappa \; \mathrm{d}\rho_t(\kappa)\right)^n = \mathbb{E}\left[\prod_{i=1}^n \kappa_t(i) \;\middle|\; \rho_t\right] \quad \text{a.s.}$$

Hence, if we write τ_n for the almost surely finite time at which the n lowest lineages are all descendants from the lowest lineage, then

$$w^n_{\tau_n} = \mathbb{E}\left[\prod_{i=1}^{n} \kappa_{\tau_n}(i) \;\middle|\; \rho_{\tau_n}\right] = \mathbb{E}\left[\prod_{i=1}^{n} \kappa_0(0)\right] = \mathbb{E}[\kappa_0(0)] = w_0.$$

This implies for all $n \in \mathbb{N}$

$$\begin{aligned}\mathbb{E}[w^n_\infty] &= \lim_{t\to+\infty} \mathbb{E}[w^n_t] \\ &= \lim_{t\to+\infty} \Big(\mathbb{E}\left[w^n_t \mathbb{1}_{\tau_n \le t}\right] + \mathcal{O}(\mathbb{P}(\tau_n > t)\Big) \\ &= w_0 \lim_{t\to+\infty} \mathbb{P}(\tau_n \le t) = w_0,\end{aligned}$$

which concludes the proof. □

2.7.3 Not coming down from infinity

Coming down from infinity addresses the question whether a coalescent started from infinitely many lineages coalesces quickly enough to later have only finitely many lineages. The next result shows that our coalescent contains a non-trivial fraction of singleton lineages at any time, a situation know as containing dust. This complements the previous result as it shows that fixation cannot occur in bounded time. A related discussion can be found on [DK99b, p.196].

Theorem 2.2.8 (The coalescent contains dust)**.** *Consider the graphical construction of the $(\boldsymbol{\gamma}, \Pi)$-Fleming-Viot process from Section 2.6. Then, for any time $t > 0$, the probability for the lowest lineage to not have had offspring before time t is positive.*

If Π_d does not charge 1, this guarantees that, at every time t, there are a.s. infinitely many ancestral lineages that do not coalesce in the time interval $[0, t]$.

Proof. Let $t \ge 0$ be arbitrary. Since the rate $\Pi_d(\{1\})$ at which the entire population is replaced is finite, there is a positive probability that no such event takes place before time t. As such, we may and will assume $\Pi_b(\{1\}) = 0$.

Conditionally on the events $\xi_0 = \{(t_i, \boldsymbol{z}_i)\}$, the first lineage does not jump until time t with probability at least $\prod_{t_i \le t}(1 - \overline{\boldsymbol{z}}_i)$, where $\overline{\boldsymbol{z}}_i := \frac{z^i_b}{(1-z^i_d)N_{t_i} + z^i_b}$. Now,

$$\prod_{t_i \le t}(1 - \overline{\boldsymbol{z}}_i) > 0 \quad \text{if and only if} \quad \sum_{t_i \le t} \overline{\boldsymbol{z}}_i < +\infty.$$

Denote by $\tau_\epsilon := \inf\{t \geq 0 \ : \ N_t \leq \epsilon\}$. Then, for any $K, \epsilon > 0$, we have

$$\begin{aligned}
\mathbb{P}\left(\sum_{t_i \leq t} \overline{\boldsymbol{z}}_i < K \right) &\geq \mathbb{P}\left(\left\{ \sum_{t_i \leq t} \overline{\boldsymbol{z}}_i < K \right\} \cap \{\tau_\epsilon > t\} \right) \\
&\geq \mathbb{P}\left(\left\{ \sum_{t_i \leq t} \frac{z_b^i}{(1-z_d^i)\epsilon + z_b^i} < K \right\} \cap \{\tau_\epsilon > t\} \right) \\
&= 1 - \mathbb{P}\left(\left\{ \sum_{t_i \leq t} \frac{z_b^i}{(1-z_d^i)\epsilon + z_b^i} \geq K \right\} \cup \{\tau_\epsilon \leq t\} \right) \\
&\geq 1 - \mathbb{P}\left(\left\{ \sum_{t_i \leq t} \frac{z_b^i}{(1-z_d^i)\epsilon + z_b^i} \geq K \right\} \right) - \mathbb{P}(\tau_\epsilon \leq t) \\
&\geq 1 - K^{-1}\mathbb{E}\left[\sum_{t_i \leq t} \frac{z_b^i}{(1-z_d^i)\epsilon + z_b^i} \right] - \mathbb{P}(\tau_\epsilon \leq t) \\
&= 1 - K^{-1} t \int_{\mathcal{Z}} \frac{z_b}{(1-z_d)\epsilon + z_b} \ \mathrm{d}\Pi(\boldsymbol{z}) - \mathbb{P}(\tau_\epsilon \leq t).
\end{aligned}$$

The integral in the last line is finite as

$$\int_{\mathcal{Z}} \frac{z_b}{(1-z_d)\epsilon + z_b} \ \mathrm{d}\Pi(\boldsymbol{z}) \leq \frac{2}{\epsilon} \int_{\{z_d \leq 1/2\}} z_d \ \mathrm{d}\Pi(\boldsymbol{z}) + \int_{\{z_d \geq 1/2\}} \mathrm{d}\Pi(\boldsymbol{z}) < +\infty.$$

If we choose K large enough and ϵ small enough, the above is bounded away from zero and the statement follows. □

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	A Λ-Fleming-Viot type model with intrinsically varying population size
Publication Status	□Published □ Accepted for Publication ■Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	arxiv.org/abs/2311.06100

Student Confirmation

Student Name:	Bastian Wiederhold		
Contribution to the Paper	By the time of my joining of the project, Julian Kern had already envisioned the model and first results such as the convergence to the Wright-Fisher diffusion. I contributed in particular to the sections involving the lookdown constructions and writing the manuscript in its current form.		
Signature 		Date	11th July 2024

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Alison Etheridge		
Supervisor comments I am confident that Bastian Wiederhold made a very substantial contribution to this publication.		
Signature 	Date	11th July 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Joint work with Raphaël Forien

3

Isolation by distance under various forms of power-law dispersal

Contents

3.1 Introduction 60
3.2 Definition of the model 66
3.3 Main results 70
3.4 Proof of the main results 80
3.5 Proofs of the lemmas 92

3.1 Introduction

Probability of identity by descent In most natural settings, the population's range is too broad to be covered by a single individual during its lifetime and several generations are necessary to bridge large spatial distances. Consequently, in the presence of neutral mutations, individuals further apart in space carry, on average, more different genetic information; an effect which is known as *isolation by distance*. In *infinitely-many-alleles models*, mutations are deemed unlikely to affect the same nucleotide twice, such that mutation events create new types never seen before in the population. Under this assumption relatedness of individuals can be measured by the *probability of identity by descent* $F(x)$, that is, the probability that two individuals sampled from two locations at separation x are of the same type because no mutation occured since their most recent common ancestor.

Historically, the displacement distance between offspring and parents has been assumed to be of finite variance. Wright and Malécot ([Wri43], [Mal48]) were the first to observe that if G_t is a Gaussian transition kernel associated with Brownian motion, $F(x)$ is approximately given by the *Wright-Malécot formula* (ignoring model specific parameters)

$$F(x) \approx \int_0^\infty e^{-2\mu t} G_{2t}(x) dt, \tag{3.1}$$

where $\mu > 0$ is the rate at which neutral mutations arise in a given lineage. This formula reflects the idea of probability of identity by descent. Since $\mu > 0$ is the neutral mutation rate, $e^{-2\mu t}$ is the probability that no mutation occurred on the line of descent of either individual since the time t of their most recent common ancestor. Mutation rates estimated from biological data are often very small. If we wish to detect mutations, we therefore need to consider large time scales. Over large temporal and spatial scales, in turn, the central limit theorem tells us that jump processes of ancestral lineages are approximated by Brownian motions. Hence, the two ancestral lineages are at the same place and coalesce with a rate proportional to $G_{2t}(x)$ after time t. The universality of the central limit theorem in approximating jump processes and ancestral lineages explains why the Wright-Malécot formula has been proven for a large variety of different spatial models such as the stepping stone model or the Gaussian version of the spatial Λ-Fleming-Viot model ([CD02], [BDE02], [BKE10]).

Long-range dispersal in biology Over the past decades, the assumption of a finite-variance distance relationship between parents and offspring has been challenged by many biological examples such as butterfly populations ([FK07]), expansion behaviour of aquatic species ([Kel+14]) or the dispersal of plant seeds ([CMS00]) and the spread of pollen ([Val+17]). In plants, long-distance dispersal of seeds has even been proposed to be an instance of the Levy foraging hypothesis ([Tre+20]), which supposes that long jumps are more efficient in the search for sparse resources.

Studying the effect of exceptionally large reproduction events on genealogy has been one of the motivations for the introduction of the spatial Λ-Fleming-Viot model (SLFV) and already [BEV10] explored deviations from Kingman's coalescent. However, only recently, the paper [For22] and Smith and Weissman [SW23] have started to classify regimes of probability of identity by descent patterns under long-range reproduction. Exceptionally large events can even persist in the limit and lead to ancestral lineages described by stable processes. Additionally, coalescence of lineages may not only happen *locally* when lineages meet, but *globally*, when lineages are still separated in space. The first instance of a long-range Wright-Malécot formula was established in [For22], in which a long-range behaviour of ancestral lineages always occurred together with global coalescence. This was complemented and confirmed by the simulation-based approach of [SW23]. Exploring dimensions one and two, Smith and Weissman discovered that the same stable behaviour of ancestral lineages can appear together with both local and global coalescence. This chapter includes and considerably extends those previous classifications by generalizing the new technique of [For22].

Novel central limit theorem technique The classical formula (3.1) contains an innate mathematical difficulty: Brownian motions and thus ancestral lineages do not meet in dimension two or higher. Wright and Malécot as well as [BDE02] and [BKE10] circumvented this problem by introducing a parameter for the local scale. Only recently, the paper [For22] provided a mathematically rigorous solution to this obstacle. The key is the observation that even in higher dimensions, in all prelimiting models a non-zero probability of coalescence remains, vanishing only in the limit. By scaling the probability of identity by descent as the model converges to the limit, [For22, Theorem 2.4] has been able to recover the classical Wright-Malécot formula regardless of dimension. The arguments of [For22, Theorem 2.4] are based on a central limit theorem for the SLFV with mutation. The SLFV process was first introduced in [Eth08], while the mutation mechanism was envisioned by [VW15]. This process $\rho = (\rho_t, t \geq 0)$ evolves on $\mathbb{R}^d \times [0,1]$ and for each time and each

spatial location $x \in \mathbb{R}^d$, $\rho_t(x)$ is a probability measure encoding the composition of genetic types in the population at location x. To represent the infinite-alleles model assumption, individuals carry a real number in $[0,1]$ as a label. Reproduction is driven by a Poisson-point process and takes place through ball-shaped events of random radius r in which first a parental type is chosen uniformly from the affected ball (see Definition 3.2.1 below). Then a random proportion u of the population, i.e. of the probability measures $\rho_t(x)$ with x inside the ball, is replaced by the parental type. The parameter u is called the impact parameter. Between reproduction events, individuals mutate by assuming new uniform labels from the interval $[0,1]$ with a certain mutation rate μ. This mutation mechanism is certainly reminiscent of the uniform marker process in [DK99a, Section 6]. At the level of the process ρ, an infinite-population limit, this results in an exponential decay of $\rho_t(x)$ towards the uniform distribution on $[0,1]$ between reproduction events.

Under appropriate scaling of space, time, impact parameter and mutation rate (see Remark 3.3.3) and in a suitable topology the rescaled process $\boldsymbol{\rho}^N = (\boldsymbol{\rho}_t^N, t \geq 0)$ converges to Lebesgue measure λ on $\mathbb{R}^d \times [0,1]$. This enables us to consider the fluctuations

$$Z_t^N = \sqrt{N\eta_N}(\boldsymbol{\rho}_t^N - \lambda) \tag{3.2}$$

around the deterministic limit. Here, η_N is a sequence of positive numbers such that $N\eta_N$ converges to infinity. If η_N is appropriately chosen, Z_t^N converges in a distributional sense to the solution to a stochastic partial differential equation (SPDE). It turns out that this SPDE has a straightforward connection to the probability of identity by descent (see Section 3.3.4).

Formulae for short and long-range dispersal Crucially, this new approach yielded a generalisation of the Wright-Malécot formula to populations with long-range reproduction events, that is to say when the distribution of the radius of balls in which events take place has a heavy tail $r^{-(1+d+\alpha)}dr$ with $\alpha < 2$. Under a suitable spatial and temporal rescaling, ancestral lineages converge in the long-range

case to stable processes instead of Brownian motions (see e.g. [FP17]). In this result, the probability of identity by descent is evaluated for a pair of individuals sampled at two locations chosen at random from two probability densities ϕ, ψ. Let $P_t^N(\phi, \psi)$ denote this probability for a population evolving according to the N-th rescaled SLFV model $(\boldsymbol{\rho}_t^N, t \geq 0)$. It is shown in [For22] that

$$\lim_{t\to\infty} \lim_{N\to\infty} N\eta_N P_t^N(\phi, \psi) = \int_{(\mathbb{R}^d)^2} F(|x-y|)\phi(x)\psi(y)dxdy,$$

for some function $F : \mathbb{R}_+ \to \mathbb{R}_+$ which depends on the dispersal regime. The scaling factor $N\eta_N$ from the central limit theorem (3.2) is precisely the scaling factor of the probability of identity by descent to recover a non-trivial formula, i.e. η_N captures how quickly the probability of identity by descent converges to zero as the limit is approached. In the short-range case, where the event radius is a constant $R > 0$,

$$F(|x-y|) = \int_0^\infty e^{-2\mu t} G_{2t}(|x-y|)dt. \tag{3.3}$$

This is the classical Wright-Malécot formula, see [BDE02]. In the long-range case, where the event size distribution has a heavy tail $r^{-(1+d+\alpha)}dr$, ancestral lineages become stable processes with parameter $\alpha \in (0, 2)$. In this case, it was shown that

$$F(|x-y|) = \int_0^\infty \int_{(\mathbb{R}^d)^2} e^{-2\mu t} G_t^\alpha(x-z_1) G_t^\alpha(y-z_2) K(z_1, z_2) dz_1 dz_2 dt, \tag{3.4}$$

where

$$K(z_1, z_2) = \frac{C}{\|z_1 - z_2\|^{d+\alpha}}. \tag{3.5}$$

Reflecting the convergence of ancestral lineages, the Brownian transition kernel G_t is replaced by a stable transition kernel G_t^α. The formula involves a new term $K(z_1, z_2)$, which appears as the spatial correlations of the noise of the SPDE arising in the central limit theorem. The correlations $K(z_1, z_2)$ encode the rate of coalescence for ancestral lineages located at positions z_1 and z_2.

New phenomena We will expand the technique of [For22] to more general models. The broader regimes will unveil some new phenomena.

In [For22] there was a strong link between the behaviour of ancestral lineages and asymptotic coalescence, while our regimes show that this is no necessity. The short-range Wright-Malécot formula (3.3) can be interpreted as having a coalescence/spatial correlation term $K(x_1, x_2) = C\delta_{x_1}(x_2)$ for a constant $C > 0$. This indicates that coalescence happens only locally and ancestral lineages need to be at the same position $x_1 = x_2$ in order to coalesce asymptotically. In the long-range regime (3.4), ancestral lineages became stable processes and the coalescence/correlation term $K(x_1, x_2)$ of (3.5) indicated that coalescence happens globally over large distances. In this work, we will see mixtures of both behaviours: Brownian motions with global coalescence and stable ancestral lineages which still only coalesce locally. More generally, in most cases, local coalescence or a range of different powers α appearing in (3.5) is possible for a fixed either stable or Brownian behaviour of ancestral lineages (see Theorem 3.3.4 and the corresponding formulae in Theorem 3.3.7).

Furthermore, we will demonstrate that the coalescence/spatial correlation behaviour shows a dimension-dependent phase transition. We will identify a critical threshold for the tail behaviour below which global coalescence occurs and above which only local coalescence takes place. The critical case will exhibit local coalescence and requires a special scaling regime.

Our more general SLFV models involve two different radii r_1 and r_2 in reproduction events (Definition 3.2.1). Inside a ball of radius r_1 a parental type will be chosen. Thereafter, a proportion of the population will be replaced by the parental type inside a ball of radius r_2. Our new Wright-Malécot formulae show that the joint distribution of r_1 and r_2 determines the asymptotic behaviour of lineages while only the r_2 marginal determines the coalescence regime. Moreover, if both tails have independent heavy-tailed distributions, one of the tails dominates the parameter of the resulting stable lineages (see (3.12), (3.14) and Table 3.1).

For potential inference of long-range dispersal patterns, this new variety of formulae has far-reaching consequences, which we hope to cover in future work. The formulae illustrate that long-range reproduction leaves a clear trace in isolation by distance patterns. However, as we obtain many situations which lead to similar or even the exact same formulae, inference of more precise model parameters might be hindered.

The information has been obtained entirely forwards in time, raising the hope that the techniques can be applied to other complex spatial models, for example those incorporating fluctuating population size. Many models in population genetics assume a constant population size as this assumption facilitates the use of many tools in particular the method of duality. As we will see in the next Chapter 4, indeed, the strategy provides a route to obtaining tractable results on the behaviour of ancestral lineages of varying population size genetics models.

Outline In Section 3.2, we introduce the SLFV models considered in this work. The central limit theorem and the corresponding Wright-Malécot formulae are stated in Section 3.3. In Section 3.4, we give an overview of the proof strategy of [For22] and the adaptations needed in the present case. Section 3.5 provides proofs of all steps requiring significant changes from [For22] to cover the new regimes. Appendix A.1 proves the existence and uniqueness of the SLFV models of Section 3.2.

3.2 Definition of the model

This chapter makes use of spatial Λ-Fleming-Viot (SLFV) processes, a framework introduced in ([Eth08], [BEV10]) in which the evolution of a population is driven by a family of reproduction events. These events affect balls of varying radius and in each event, a fraction of the population inside the ball is replaced by the offspring of one or a few individuals. The SLFV process presented here differs in two ways from the most commonly studied version ([BEV10]): a different mutation mechanism, introduced in [VW15], and different areas are chosen for the origin of

parents and dispersal of offspring. We first define the state space Ξ of the process $(\rho_t, t \geq 0)$, then we clarify its evolution.

Let $\mathcal{M}\Big([0,1]\Big)$ denote the set of probability measures on the interval $[0,1]$. The state space Ξ is defined as the set of maps

$$\rho : \mathbb{R}^{\mathrm{d}} \to \mathcal{M}\Big([0,1]\Big),$$

where we identify two maps ρ_1, ρ_2, if $\{x \in \mathbb{R}^{\mathrm{d}} : \rho_1(x) \neq \rho_2(x)\}$ is a Lebesgue null set. This space is in one-to-one correspondence with the space of measures on $\mathbb{R}^d \times [0,1]$ with Lebesgue measure as a spatial marginal. At time t, the measure $\rho_t(x)$ describes the probability distribution of the type of an individual sampled from position x. To introduce a suitable topology on the space Ξ, let $(\phi_n, n \geq 1)$ be a sequence of continuous functions with compact support on $\mathbb{R}^d$, which separate points in Ξ. For $\rho_1, \rho_2 \in \Xi$ we define the metric d by

$$d(\rho_1, \rho_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \Big| \langle \rho_1, \phi_n \rangle - \langle \rho_2, \phi_n \rangle \Big|, \tag{3.6}$$

which induces the vague topology on Ξ. It is useful, for later proofs, to define p-norms as in [For22] only on the spatial argument of the test functions

$$\|\phi\|_p := \left(\int_{\mathbb{R}^{\mathrm{d}}} \sup_{k \in [0,1]} |\phi(x,k)|^p dx \right)^{1/p}. \tag{3.7}$$

We can assume without loss of generality that for multi-indices $\kappa \in \mathbb{N}^d$

$$\sup_{n \in \mathbb{N}} \Big\{ \|\partial_\kappa \phi_n\|_p : |\kappa| \leq 2, p \in \{1, \infty\} \Big\} < \infty. \tag{3.8}$$

The next definition describes the evolution of the spatial Λ-Fleming-Viot process. It extends [For22, Definition 1.1] by introducing balls of different sizes for the selection of parents and dispersal of offspring and allowing the impact of reproduction events to be random.

Definition 3.2.1. Suppose we are given a positive constant μ and a σ-finite measure ν on $(0,1] \times (0,\infty)^2$, which encodes the event impact and the event radii and satisfies

$$\int_{(0,1]\times(0,\infty)^2} u r_2^d \, \nu(du, dr_1, dr_2) < \infty. \tag{3.9}$$

Then let Π be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times (0,1] \times (0,\infty)^2$ with intensity $dt \otimes dx \otimes \nu(du, dr_1, dr_2)$. The spatial Λ-Fleming-Viot process with mutation (SLFV) is defined as a stochastic process $\rho = (\rho_t, t \geq 0)$ taking values in Ξ which evolves through reproduction and mutation.

1. Reproduction: For every point $(t, x, u, r_1, r_2) \in \Pi$, first choose a spatial position y uniformly at random from within the parent search area $B(x, r_1)$ and a parental type $k_0 \in [0,1]$ according to $\rho_{t^-}(y, dk)$. Then, replace a proportion u of the population inside the dispersal area $B(x, r_2)$, i.e. set

$$\forall z \in B(x, r_2), \;\; \rho_t(z, dk) = (1-u)\rho_{t^-}(z, dk) + u\delta_{k_0}(dk). \tag{3.10}$$

2. Mutation: If a point $x \in \mathbb{R}^{\mathrm{d}}$ is hit by events at times t_1 and t_2 with no events in between, we set for $t_1 \leq s < t_2$,

$$\rho_s(x, dk) = e^{-\mu(s-t_1)}\rho_{t_1}(x, dk) + \left(1 - e^{-\mu(s-t_1)}\right) dk. \tag{3.11}$$

In Figure 3.1, we show the effect of a reproduction event as described above on the frequency of a single allele in two cases: if the parent search radius r_1 is larger than the replacement radius r_2 and if the former is smaller than the latter.

Note that the mutation mechanism is specifically tailored to capture probabilities of identity by descent. In infinite-alleles models, if an individual acquires a mutation and a new type is created, the individual no longer contributes to the probability of identity by descent. This can be described by a uniform distribution on $[0,1]$, as specific points are not charged. Therefore, we can view neutral mutation of rate μ as a constant decay towards a uniform distribution. In [For22], more general mutation mechanisms are also considered.

Definition 3.2.1 allows us to define the infinitesimal generator of the SLFV with mutations on a suitable set of functions. For $\rho \in \Xi$, $x \in \mathbb{R}^d$, $r \in (0,\infty)$, $u \in (0,1]$ and $k_0 \in [0,1]$, define

$$\Theta_{x,r,u,k_0}(\rho)(z, dk) := \begin{cases} (1-u)\rho(z, dk) + u\delta_{k_0}(dk) & \text{if } z \in B(x,r), \\ \rho(z, dk) & \text{if } z \notin B(x,r). \end{cases}$$

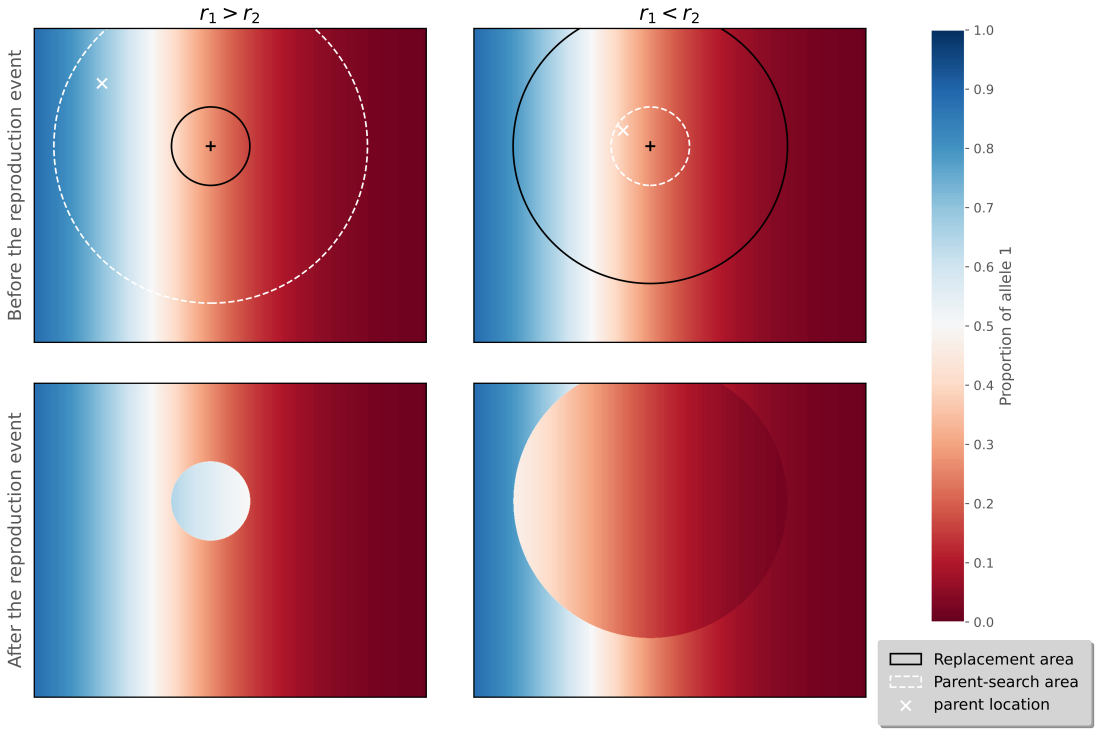


Figure 3.1: Example of the effect of a single reproduction event on the frequency of one allele in a spatial Λ-Fleming-Viot model with two alleles (0 and 1): on the left, the parent-search radius r_1 is larger than the replacement radius r_2, allowing an individual carrying allele 1 to send offspring far from its current location; on the right, the parent-search radius r_1 is smaller than the replacement radius r_2, allowing an individual to recolonise a large region.

Let $D(\mathcal{G})$ denote the set of functions $F_{f,\phi} : \Xi \to \mathbb{R}$ of the form $F_{f,\phi}(\rho) = f(\langle \rho, \phi \rangle)$, where $\phi : \mathbb{R}^d \to \mathbb{R}$ is continuous and compactly supported, and $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable. The infinitesimal generator of the SLFV of Definition 3.2.1 is the operator $\mathcal{G}$, defined on $D(\mathcal{G})$ as

$$\begin{aligned}
&\mathcal{G}F_{f,\phi}(\rho) \\
&:= \mu f'(\langle \rho, \phi \rangle)\langle \lambda - \rho, \phi \rangle \\
&\quad + \int_{\mathbb{R}^d} \int_{(0,1]\times(0,\infty)^2} \frac{1}{|B(x, r_1)|} \int_{B(x,r_1)} \int_{[0,1]} (F_{f,\phi}(\Theta_{x,r_2,u,k}(\rho)) - F_{f,\phi}(\rho)) \\
&\qquad\qquad \rho(y, dk)dy\, \nu(du, dr_1, dr_2)\, dx,
\end{aligned}$$

where $|B(x, r_1)|$ denotes the Lebesgue measure of $B(x, r_1)$ and λ denotes the constant map which associates Lebesgue measure on $[0, 1]$ to each $x \in \mathbb{R}^d$, i.e. $\lambda(x, dk) = dk$. We then say that a Ξ-valued process solves the martingale problem associated

to $(\mathcal{G}, D(\mathcal{G}))$ if, for any $F \in D(\mathcal{G})$,

$$F(\rho_t) - F(\rho_0) - \int_0^t \mathcal{G}F(\rho_s)ds$$

is an $\mathcal{F}_t$-local martingale, where $(\mathcal{F}_t, t \geq 0)$ is the natural filtration associated to the process $(\rho_t, t \geq 0)$.

Theorem 3.2.2. *Under condition* (3.9)*, there exists a unique Markov process taking values in* Ξ *which solves the martingale problem associated to* $(\mathcal{G}, D(\mathcal{G}))$.

We prove Theorem 3.2.2 in Appendix A.1, extending the arguments used in the proof of Theorem 1.2 in [EVY20]. Hereafter, the unique solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$ is called the SLFV with mutations (or simply the SLFV), with impact and radii intensity measure ν and mutation rate μ.

3.3 Main results

We generalise the results of [For22] to a spectrum of regimes exhibiting new phenomena. Its centrepiece is a central limit theorem for the fluctuations of the process, from which we deduce the asymptotic behaviour of pairs of ancestral lineages.

3.3.1 Choice of parameters

We consider two possible choices for the intensity measure ν, which are interpreted in detail in view of the main results in Subsection 3.3.5. In the *one-tail regime*, the intensity measure ν governing the reproduction events is chosen as

$$\nu_N(du, dr_1, dr_2) = \mathbb{1}_{r_2 \geq 1}\, \delta_{u_N r_2^{-\mathrm{c}}}(du)\, \delta_{r_2^{\mathrm{b}}}(dr_1)\, \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}}, \tag{3.12}$$

where $u_N \in (0,1]$, $\mathrm{a} \in (0,\infty)$ and $\mathrm{b}, \mathrm{c} \in [0,\infty)$ are parameters. Here, ancestral lineages will converge to symmetric α-stable processes with parameter

$$\alpha = \mathrm{a} \wedge \frac{\mathrm{a}}{\mathrm{b}} \wedge 2. \tag{3.13}$$

The parameter c appears in the tail distribution of the radius r_2 and the impact $u = u_N r_2^{-\mathrm{c}}$. When we consider the behaviour of lineages, both appearances

annihilate each other, so that the parameter of the stable processes (3.13) does not depend on c (see Subsection 3.3.5 for details). Introducing the parameter c allows us to modulate the impact of large events. For example, if c = d the number of offspring produced is the same for every reproduction event regardless of the radius r_2. Examples of realisations of the SLFV in the one-tail regime are displayed in Figure 3.2.

In contrast, in the *two-tails regime*, the intensity measure ν has the form

$$\nu_N(du, dr_1, dr_2) = \mathbb{1}_{r_1 \geq 1}\, \mathbb{1}_{r_2 \geq 1}\, \delta_{u_N r_1^{-\mathrm{c}_1} r_2^{-\mathrm{c}_2}}(du)\, \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1-\mathrm{c}_1} r_2^{1+\mathrm{d}+\mathrm{a}_2-\mathrm{c}_2}}, \tag{3.14}$$

where $u_N \in (0,1]$, $\mathrm{a}_1, \mathrm{a}_2 \in (0,\infty)$ and $\mathrm{c}_1, \mathrm{c}_2 \in [0,\infty)$ are parameters. In this case, limiting ancestral lineages will be symmetric α-stable processes of parameter

$$\alpha = \mathrm{a}_1 \wedge \mathrm{a}_2 \wedge 2. \tag{3.15}$$

It can then easily be checked that both these measures satisfy (3.9). Indeed,

$$\begin{aligned} &\int_{(0,1]\times(0,\infty)^2} u r_2^{\mathrm{d}}\, \nu_N(du, dr_1, dr_2) \\ &= \begin{cases} u_N \int_1^{+\infty} \frac{dr_2}{r_2^{1+\mathrm{a}}} & \text{in the one-tail regime,} \\ u_N \int_1^{+\infty} \frac{dr_1}{r_1^{1+\mathrm{a}_1}} \int_1^{+\infty} \frac{dr_2}{r_2^{1+\mathrm{a}_2}} & \text{in the two-tails regime.} \end{cases} \end{aligned}$$

In both regimes, the right hand side is finite.

Remark 3.3.1. The paper [For22] considers the one-tail regime with parameters $b = 1$ and $c = 0$, which amounts to considering only a single radius $r_1 = r_2$ and constant impact. Our results augment the corresponding results of [For22] to the above broader parameter setting and an interpretation thereof is given in Section 3.3.5.

3.3.2 Law of large numbers

Our goal is to derive a central limit theorem for the process. To derive a law of large numbers in Theorem 3.3.2, we apply the following scaling, identified in [For22], to the process. The scaling will be explained in Remark 3.3.3. We first choose a

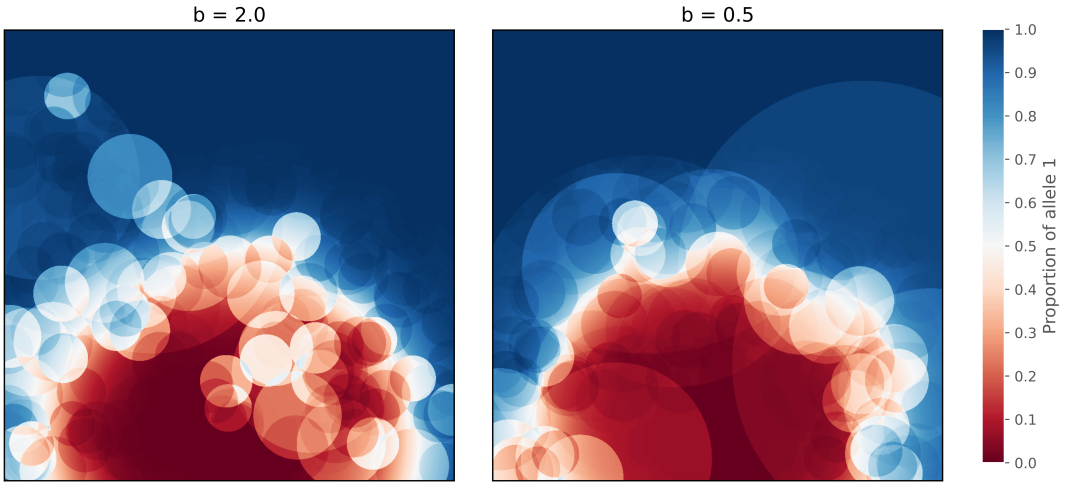


Figure 3.2: Two realisations of the spatial Λ-Fleming-Viot model in two dimensions. Each simulation is started with two alleles (1 and 0), with individuals carrying type 0 initially concentrated in a fixed ball. The plots show the proportion of individuals carrying allele 1 after some time. The parameters of the SLFV are chosen according to the one-tail regime, one where $b = 2$ (on the left) and one where $b = 0.5$ (on the right). In both simulations, the parameter a is chosen so that $\alpha = 1.3$.

sequence $(\delta_N, N \geq 1)$ of positive real numbers which converges to zero as $N \to \infty$. Let $(\rho_t^N, t \geq 0)$ be the SLFV of Definition 3.2.1, with

$$u_N = \frac{u_0}{N}, \qquad \mu_N = \delta_N^\alpha \frac{\mu}{N},$$

where α is given by (3.13) in the one-tail regime and by (3.15) in the two-tails regime. The rescaled process is then defined as

$$\boldsymbol{\rho}_t^N(x, dk) := \rho^N_{Nt/\delta_N^\alpha}\left(\frac{x}{\delta_N}, dk\right).$$

The sequence of rescaled processes $(\boldsymbol{\rho}_t^N, t \geq 0)$ converges to a deterministic limit given by Lebesgue measure λ. The convergence will be made precise in Theorem 3.3.2. This allows us to consider the fluctuations

$$Z_t^N := \sqrt{N\eta_N}(\boldsymbol{\rho}_t^N - \lambda), \tag{3.16}$$

where η_N is a sequence of real numbers capturing the magnitude of fluctuations. In the one-tail regime, we set

$$\eta_N := \begin{cases} \delta_N^{\alpha - \mathrm{a} - \mathrm{c}} & \text{if } \mathrm{a} + \mathrm{c} < \mathrm{d}, \\ \delta_N^{\alpha - \mathrm{d}} \log(1/\delta_N)^{-1} & \text{if } \mathrm{a} + \mathrm{c} = \mathrm{d}, \\ \delta_N^{\alpha - d} & \text{if } \mathrm{a} + \mathrm{c} > \mathrm{d}, \end{cases}$$

while in the two-tail regime,

$$\eta_N := \begin{cases} \delta_N^{\alpha - \mathrm{a}_2 - \mathrm{c}_2} & \text{if } \mathrm{a}_2 + \mathrm{c}_2 < \mathrm{d}, \\ \delta_N^{\alpha - \mathrm{d}} \log(1/\delta_N)^{-1} & \text{if } \mathrm{a}_2 + \mathrm{c}_2 = \mathrm{d}, \\ \delta_N^{\alpha - d} & \text{if } \mathrm{a}_2 + \mathrm{c}_2 > \mathrm{d}\,. \end{cases}$$

With those definitions at hand, we can formulate the law of large numbers extending [For22, Proposition 3.6].

Theorem 3.3.2. *Let $(\delta_N, N \geq 1)$ be a sequence of positive numbers converging to zero as $N \to \infty$ and such that $N\eta_N \to \infty$ as $N \to \infty$. Let $\boldsymbol{\rho}_t^N$ be the SLFV started from $\rho_0 = \lambda$. Then, for each fixed $T > 0$,*

$$\lim_{N\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} d(\boldsymbol{\rho}_t^N, \lambda)\right] = 0,$$

where d is the metric defined in (3.6).

Remark 3.3.3. We wish to clarify how our choice of scaling aligns with our aim to show a central limit theorem for the process. To obtain a deterministic law of large numbers, we require the impact u_N to converge to zero and we would like to consider large spatial scales. Without loss of generality, we can assume $u_N = u_0/N$ and scale space by $1/\delta_N$ for a sequence of positive reals converging to zero.

Ancestral lineages need to be among the impacted individuals in order to jump, which causes the effective speed of an ancestral lineage to be proportional to the average impact of events ([BEV10]). If we sample individuals at distance D, which in unrescaled spatial units is proportional to $1/\delta_N$, it will therefore take roughly a time $ND^\alpha \approx N/\delta_N^\alpha$ to cover this distance. We scale time by N/δ_N^α, so that the time to cover a distance D is of order $O(1)$ and lineages are capable of movement as we approach the limit. To be able to detect mutation and genetic events in the population, the expected number of mutations along the lineages needs to be of order $O(1)$ and as $1/\mu \sim N/\delta_N^\alpha$ we scale the mutations by the inverse.

3.3.3 Central limit theorem

Prior to stating the theorem, we need to introduce a few notations. Let V_r denote the volume of a ball of radius $r > 0$ in $\mathbb{R}^d$ and let $V_r(x, y)$ denote the volume of the intersection of $B(x, r)$ and $B(y, r)$ for $x, y \in \mathbb{R}^d$. Let α, β, γ, σ and ζ be given by Table 3.1. Define an operator $\mathcal{D}_\alpha$ acting on twice-continuously differentiable functions $\phi : \mathbb{R}^{\mathrm{d}} \times [0, 1] \to \mathbb{R}$ by

$$\mathcal{D}_\alpha \phi(x, k) := \begin{cases} \zeta \int_0^\infty \left(\overline{\phi}(x, k, r) - \phi(x, k)\right) \dfrac{dr}{r^{1+\alpha}} & \text{if } \alpha < 2, \\ \dfrac{\sigma^2}{2} \Delta \phi(x, k) & \text{if } \alpha = 2, \end{cases} \tag{3.17}$$

where $\overline{\phi}(x, k, r)$ is defined as the average of $y \mapsto \phi(y, k)$ over the ball $B(x, r)$ of radius r, i.e.

$$\overline{\phi}(x, k, r) = \frac{1}{V_r} \int_{B(x,r)} \phi(y, k) dy, \tag{3.18}$$

and the Laplacian operator Δ acts only on the spatial coordinates, i.e.

$$\Delta \phi(x, k) = \sum_{i=1}^{d} \frac{\partial^2 \phi}{\partial x_i^2}(x, k).$$

In the case $\alpha < 2$, $\mathcal{D}_\alpha \phi(x, k)$ can be seen as equal to

$$\frac{\zeta}{V_1(d + \alpha)} \lim_{r \to 0} \int_{\mathbb{R}^{\mathrm{d}} \setminus B(x,r)} \frac{\phi(z, k) - \phi(x, k)}{\|z - x\|^{\mathrm{d}+\alpha}} dz,$$

which in turn, if multiplied by $\zeta^{-1} V_1(\mathrm{d} + \alpha) 2^\alpha \pi^{-\mathrm{d}/2} \Gamma(\frac{\mathrm{d}+\alpha}{2}) |\Gamma(-\frac{\alpha}{2})|^{-1}$, is the fractional Laplacian applied to $x \mapsto \phi(x, k)$ (see [Kwa17]). Furthermore, define a measure K_β on $(\mathbb{R}^d)^2$ by

$$K_\beta(dx_1, dx_2) := \begin{cases} \delta_{x_2}(dx_1) dx_2 & \text{if } \beta \geq d, \\ \dfrac{1}{\|x_1 - x_2\|^\beta} dx_1 dx_2 & \text{if } \beta < d. \end{cases} \tag{3.19}$$

Our next result shows the convergence of the fluctuations (3.16) to the solution of a stochastic partial differential equation (SPDE). A majority of this work will be dedicated to proving this result. The SPDE is of a common form, and mirrors the corresponding result of [For22, Theorem 2.1, Theorem 2.3]. However, for certain

	One tail regime	Two tail regime
α	$\mathrm{a} \wedge \frac{\mathrm{a}}{\mathrm{b}} \wedge 2$	$\mathrm{a}_1 \wedge \mathrm{a}_2 \wedge 2$
β	$\mathrm{a}+\mathrm{c}$	$\mathrm{a}_2+\mathrm{c}_2$
γ	$\frac{u_0^2 V_1^2}{\mathrm{a}+\mathrm{c}-\mathrm{d}}$ if $\beta > \mathrm{d}$ $u_0^2 V_1^2$ if $\beta = \mathrm{d}$ $u_0^2 C_{d,\beta}^{(2)}$ if $\beta < \mathrm{d}$	$\frac{u_0^2 V_1^2}{(\mathrm{a}_1+\mathrm{c}_1)(\mathrm{a}_2+\mathrm{c}_2-\mathrm{d})}$ if $\beta > \mathrm{d}$ $\frac{u_0^2 V_1^2}{\mathrm{a}_1+\mathrm{c}_1}$ if $\beta = \mathrm{d}$ $\frac{u_0^2 C_{d,\beta}^{(2)}}{\mathrm{a}_1+\mathrm{c}_1}$ if $\beta < \mathrm{d}$
σ^2	$\frac{u_0 V_1}{\mathrm{d}+2}\left(\frac{1}{\mathrm{a}-2\,\mathrm{b}}+\frac{1}{\mathrm{a}-2}\right)$	$\frac{u_0 V_1}{\mathrm{d}+2}\left(\frac{1}{(\mathrm{a}_1-2)\,\mathrm{a}_2}+\frac{1}{\mathrm{a}_1(\mathrm{a}_2-2)}\right)$
ζ	$\frac{u_0 V_1}{1 \vee \mathrm{b}}$ if $\mathrm{b} \neq 1$ $\frac{C_{d,\alpha}^{(1)}}{V_1(\mathrm{d}+\alpha)}$ if $\mathrm{b} = 1$	$\frac{u_0 V_1}{\mathrm{a}_1 \vee \mathrm{a}_2}$ if $\mathrm{a}_1 \neq \mathrm{a}_2$ $u_0 V_1\left(\frac{1}{\mathrm{a}_1}+\frac{1}{\mathrm{a}_2}\right)$ if $\mathrm{a}_1 = \mathrm{a}_2$

Table 3.1: Expressions for the different parameters of the limiting fluctuations. The constants $C_{d,\alpha}^{(1)}$ and $C_{d,\beta}^{(2)}$ are given by

$$C_{d,\alpha}^{(1)} := \int_0^\infty \frac{V_r(0,e_1)}{V_r^2}\frac{dr}{r^{1+\alpha}}, \qquad C_{d,\beta}^{(2)} := \int_0^\infty V_r(0,e_1)\frac{dr}{r^{1+\mathrm{d}+\beta}},$$

where e_1 denotes an arbitrary point of the unit sphere in $\mathbb{R}^{\mathrm{d}}$.

parts of the parameter space the SPDE takes new forms specified in Table 3.1. We denote the space of tempered distributions acting on $\mathcal{S}(\mathbb{R}^d \times [0,1])$ by $\mathcal{S}'(\mathbb{R}^d \times [0,1])$. The proof of the result is given in Section 3.4.2 and differences to [For22] are detailed in Section 3.4.2.

Theorem 3.3.4. *Let Z_t^N be the fluctuations of an SLFV process with mutation as defined in* (3.16)*. Suppose that the assumptions of Theorem 3.3.2 are satisfied and that, in addition,*

$$\lim_{N\to\infty} \sqrt{\frac{\eta_N}{N}}\delta_N^{\mathrm{d}\wedge\mathrm{c}} = 0, \qquad \textit{in the one-tail regime,} \tag{3.20}$$

$$\lim_{N\to\infty} \sqrt{\frac{\eta_N}{N}}\delta_N^{\mathrm{d}\wedge\mathrm{c}_2} = 0, \qquad \textit{in the two-tail regime.} \tag{3.21}$$

Then the sequence $(Z_t^N, t \geq 0), N = 1,2,\ldots$ converges in distribution in

$D(\mathbb{R}_+, \mathcal{S}'(\mathbb{R}^{\mathrm{d}} \times [0,1]))$ *to a process* $(Z_t, t \geq 0)$ *solving the mild form of the equation*

$$\begin{cases} dZ_t = [\mathcal{D}_\alpha Z_t - \mu Z_t]dt + \sqrt{\gamma}\, dW(t), \\ Z_0 = 0. \end{cases} \tag{3.22}$$

Here, $(W(t), t \geq 0)$ *denotes a Wiener process on* $\mathbb{R}^{\mathrm{d}} \times [0,1]$*, uncorrelated in time, with covariation measure*

$$\mathcal{Q}(dx_1 dk_1 dx_2 dk_2) = K_\beta(dx_1, dx_2)\Big(dk_1 \delta_{k_1}(dk_2) - dk_1 dk_2\Big).$$

Remark 3.3.5. The results of [For22] appear in Theorem 3.3.4 as a special case of the one-tail regime when $\mathrm{b} = 1, \mathrm{c} = 0$ and $\mathrm{a} + \mathrm{c} < \mathrm{d}$.

Remark 3.3.6. The assumptions (3.20) and (3.21) on the convergence speed of the sequence $(\delta_N, N \geq 1)$ can be omitted in many situations, e.g. if either $c \geq d$ or if $\alpha = \mathrm{a}$, except if $\mathrm{a} = \mathrm{d}$ and $\mathrm{c} = 0$.

3.3.4 Wright-Malécot Formulae

From the central limit theorem, we can derive expressions for the probability of identity for two individuals sampled from a population evolving according to our SLFV model (see Section 3.4.3). The connection was established in [For22, Theorem 2.4], and the main idea can be described as follows: Let ϕ, ψ be two smooth and compactly supported probability densities. Recall from the introduction that $P_t^N(\phi, \psi)$ denotes the probability that two individuals sampled according to ϕ, ψ from the N-th rescaled SLFV model $(\boldsymbol{\rho}_t^N, t \geq 0)$ are of the same type. We can write

$$P_t^N(\phi, \psi) = \mathbb{E}\Big[\Big\langle \boldsymbol{\rho}_t^N \otimes \boldsymbol{\rho}_t^N, (\phi \otimes \psi)\mathbb{1}_\Delta \Big\rangle\Big]$$

where $\mathbb{1}_\Delta : [0,1]^2 \to \mathbb{R}$ is defined as $\mathbb{1}_\Delta(k_1, k_2) = \mathbb{1}_{k_1 = k_2}$. It was observed in [For22, p.33f] that

$$P_t^N(\phi, \psi) = (N\eta_N)^{-1}\mathbb{E}\Big[\Big\langle Z_t^N \otimes Z_t^N, (\phi \otimes \psi)\mathbb{1}_\Delta \Big\rangle\Big],$$

since $Z_t := \sqrt{N\eta_N}(\boldsymbol{\rho}_t^N - \lambda)$ and Lebesgue measure λ does not charge the diagonal $\{(k_1, k_2) : k_1 = k_2\}$. This implies

$$\lim_{N \to \infty} N\eta_N P_t^N(\phi, \psi) = \mathbb{E}\Big[\Big\langle Z_t \otimes Z_t, (\phi \otimes \psi)\mathbb{1}_\Delta \Big\rangle\Big],$$

and connects $P_t^N(\phi,\psi)$ to the fluctuations enabling us to recover Wright-Malécot formulae.

Theorem 3.3.7. *Let G_t^α be the fundamental solution associated to $\partial_t - \mathcal{D}_\alpha$. Then the Wright-Malécot formula is given by*

$$\lim_{t\to\infty}\lim_{N\to\infty} N\eta_N P_t^N(\phi,\psi) = \int_{(\mathbb{R}^{\mathrm{d}})^2} F_{d,\alpha,\beta}\Big(|x-y|\Big)\phi(x)\psi(y)dxdy, \tag{3.23}$$

where the function $F_{d,\alpha,\beta}:\mathbb{R}_+\to\mathbb{R}_+$ is defined as

$$F_{d,\alpha,\beta}\Big(|x-y|\Big) := \gamma\int_0^\infty\int_{(\mathbb{R}^{\mathrm{d}})^2} e^{-2\mu t}G_t^\alpha(x-z_1)G_t^\alpha(y-z_2)K_\beta(dz_1,dz_2)dt. \tag{3.24}$$

3.3.5 Interpretation

For a given set of parameters of the one- or two-tail regime, Table 3.1 provides the values of α and β, from which we deduce the form of $\mathcal{D}_\alpha$, $K_\beta(dx_1,dx_2)$ and the corresponding Wright-Malécot formula (3.23). The generator $\mathcal{D}_\alpha, \alpha\in(0,2]$ reflects either a stable process or Brownian behaviour of ancestral lineages. The correlations $K_\beta(dx_1,dx_2)$ are either of the form $\delta_{x_1}(x_2)$, indicating local coalescence, i.e. asymptotically ancestral lineages need to meet to coalesce, or are of the form

$$K_\beta(dx_1,dx_2) = \frac{1}{\|x_1-x_2\|^\beta}dx_1dx_2.$$

This quantity reflects the rate of long range coalescence after noting that

$$\gamma K_\beta(x_1,x_2) = u_0^2 C_{d,\beta}^{(2)}K_\beta(x_1,x_2) = u_0^2\int_{|x_1-x_2|/2}^\infty V_r(x_1,x_2)\frac{dr}{r^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}}dx_1dx_2.$$

Ancestral lineages at positions x_1 and x_2 coalesce precisely if they are affected by a reproduction event of radius bigger than $|x_1-x_2|/2$, whose centre falls into the intersection of $B(x_1,r)$ and $B(x_2,r)$ of volume $V_r(x_1,x_2)$. This happens at a rate proportional to the squared impact $u_0^2/r^{2\,\mathrm{c}}$, as both ancestral lineages need to be affected by the event. Integrating with respect to r we arrive at the above expression. Our results reveal two main new phenomena.

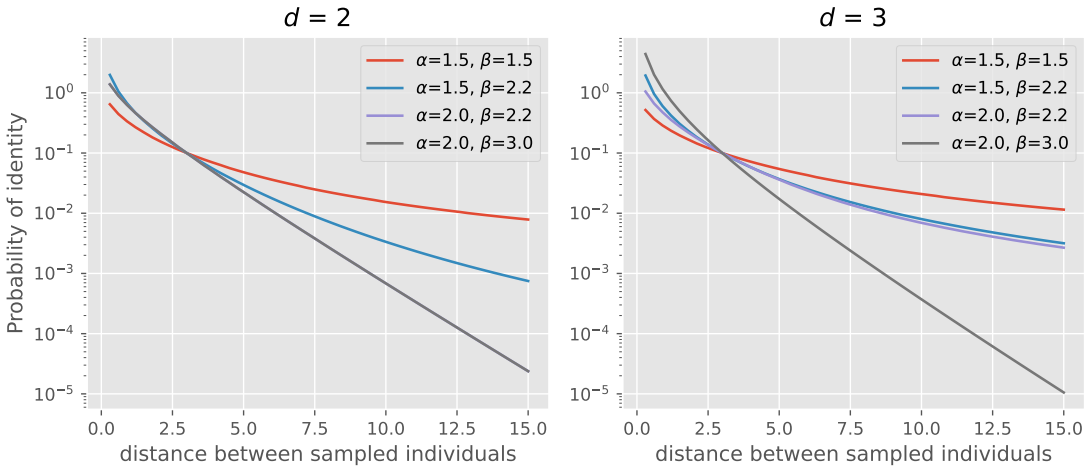


Figure 3.3: Graph of the function $r \mapsto F_{d,\alpha,\beta}(r)$ defined in (3.24) in two and three dimensions for several choices of α and β. The function is computed after setting $\zeta = 1$, $\sigma^2 = 1$, $\mu = 0.2$, and γ is chosen in order to normalise the functions at $r = 3$. On the left, $\mathrm{d} = 2$, and four curves are plotted: the red curve corresponds to long range dispersal ($\alpha = 1.5 < 2$) and long range coalescence ($\beta = 1.5 < d$), the blue curve corresponds to long range dispersal ($\alpha < 2$) with short range coalescence ($\beta = 2.2 \geq \mathrm{d}$), the purple and grey curve correspond to short range dispersal ($\alpha = 2$) and short range coalescence ($\beta \geq \mathrm{d}$). The last two curves are superposed because of the normalisation. On the right plot, four curves of the same parameters are plotted in dimension $\mathrm{d} = 3$: the red and blue curves correspond to long range dispersal ($\alpha = 1.5 < 2$) and long range coalescence ($\beta < \mathrm{d}$), the purple curve corresponds to short range dispersal ($\alpha = 2$) and long range coalescence ($\beta = 2.2 < \mathrm{d}$), and the grey curve corresponds to short range dispersal ($\alpha = 2$) and short range coalescence ($\beta = 3 \geq \mathrm{d}$). It can be noted that having either long range dispersal or long range coalescence significantly slows down the decrease of the probability of identity with geographic distance, and that this effect is more pronounced for small values of α and β.

Decorrelation of ancestral motion and coalescence behaviour Most notably, a phase transition occurs at $\beta = \mathrm{d}$. If $a + c =: \beta \geq \mathrm{d}$ or $\mathrm{a}_2 + \mathrm{c}_2 =: \beta \geq \mathrm{d}$ no spatial correlations persist and coalescence happens locally, while if $\beta < \mathrm{d}$ the spatial correlations are of non-trivial form $K_\beta(dx_1, dx_2) = \|x_1 - x_2\|^{-\beta} dx_1 dx_2$ and coalescence happens at long range. Crucially, the correlations/coalescence behaviour $K_\beta(x_1, x_2)$ and the parameter α of stable ancestral lineages are not as closely linked as suggested by the results of [For22]. Fractional Brownian lineages can now appear with only local coalescence, e.g. by choosing in the one-tail regime in dimension $\mathrm{d} = 2$,

$$\mathrm{a} = 1.5, \quad \mathrm{b} = 1, \quad \mathrm{c} = 0.7.$$

By Theorem 3.3.4, ancestral lineages are fractional ($\alpha = \mathrm{a} = 1.5$) and coalescence happens locally as $\beta = \mathrm{a} + \mathrm{c} = 2.2 > 2 = \mathrm{d}$ and therefore $K(x_1, x_2) = u_0^2 V_1^2 \delta_{x_1}(x_2)$. Even more surprisingly, in dimension $\mathrm{d} \geq 3$ we can obtain Brownian lineages, while still maintaining long range coalescence. For example, we can choose in the one-tail regime in $\mathrm{d} = 3$,

$$\mathrm{a} = 2, \;\; \mathrm{b} = 0.5, \;\; \mathrm{c} = 0.2.$$

Again, by Theorem 3.3.4, ancestral lineages are Brownian $\alpha = \frac{2}{1 \vee 0.5} \wedge 2 = 2$ and because $\beta = \mathrm{a} + \mathrm{c} = 2.2 < 3 = \mathrm{d}$ the spatial correlations/coalescence rate is

$$K(dx_1, dx_2) = u_0^2 C_{d,2}^{(2)} \frac{dx_2 dx_2}{\|x_1 - x_2\|^\beta}.$$

Both of the above cases can be found in Figure 3.3.

Asymmetric influence of r_1 and r_2 The operator $\mathcal{D}_\alpha$ is a fractional or standard Laplacian depending on the value of α. We can note that α is governed by the tails of the two marginal distributions of r_1 and r_2. More precisely, the radius with the heaviest tail dominates and determines the dispersal regime of the ancestral lineages (with a correction by a factor r_2^d in the tail of r_2 coming from the rate at which a reproduction event affects a given spatial location, see Section A.1.1 below).

On the other hand, the coalescence regime (i.e. whether K_β is concentrated on the diagonal or not) is only determined by the tail of the replacement radius r_2. Long range coalescence can only originate from reproduction events with a very large replacement radius. This is why b does not appear in the formula for β in the one-tail regime, and neither do a_1 and c_1 in the two-tail regime.

The parameter c (or c_1 and c_2 in the two-tail regime) describes how the impact of a reproduction event depends on its size. For instance, if $\mathrm{c} = 0$ in the one-tail regime, the impact is constant, and the number of offspring produced during an event is proportional to r_2^d, while if $\mathrm{c} = \mathrm{d}$, the number of offspring produced is independent of its size. The parametrisation of the two regimes in the definition of ν_N (see (3.12) and (3.14)) is such that these parameters do not affect the value

of α, but they do impact β. Hence the parameter c (or c_2) is another way of decoupling the dispersal from the coalescence regime.

3.4 Proof of the main results

At the heart of the present work lies Theorem 3.3.4, a central limit theorem for the SLFV process. First, we represent the SLFV process as a semimartingale driven by worthy martingale measures in Section 3.4.1. Worthy martingale measures were used by [Wal86] to generalise classical stochastic integrals to include space. In turn, this theory allows us to write the sequence of SLFV processes as a sequence of stochastic integrals. Second, in Section 3.4.2, we proceed by proving tightness of the sequence and uniqueness of limit points, which proves Theorem 3.3.4. The Wright-Malécot formulae will follow almost immediately from the corresponding central limit theorems (Section 3.4.3).

The overall strategy of the proofs follows [For22]. Non-trivial adaptations to cover the new regimes are necessary in several intermediate lemmas, which we introduce in the overall context in this section. We will summarize these adaptations in Section 3.4.2, before proving the lemmas in Section 3.5.

3.4.1 Semimartingale decomposition

We characterize the SLFV via a martingale problem adapting the calculations of [For22, Proposition 3.3] to include two different radii and varying impact. Let

$$E^p := L^p(\mathbb{R}^{\mathrm{d}} \times [0,1]) \cap L^\infty(\mathbb{R}^{\mathrm{d}} \times [0,1])$$

be the space of uniformly bounded L^p-functions. For $\phi \in E^p$, the averaged function $\overline{\phi}$ was defined in (3.18). Accordingly, the double averaged function is for radii $r_1, r_2 > 0$ defined as

$$\overline{\overline{\phi}}(x,k,r_1,r_2) := \frac{1}{V_{r_1} V_{r_2}} \int_{B(x,r_1)} \int_{B(y,r_2)} \phi(z,k) dz dy.$$

For $r_1, r_2 > 0$ and $\rho \in \Xi$, we define $[\rho]_{r_1,r_2} : \mathbb{R}^{\mathrm{d}} \times \mathbb{R}^{\mathrm{d}} \to \mathcal{M}([0,1])$ by

$$[\rho]_{r_1,r_2}(x_1, x_2, dk) = \begin{cases} \dfrac{1}{V_{r_2}(x_1,x_2)} \displaystyle\int_{B(x_1,r_2)\cap B(x_2,r_2)} \dfrac{1}{V_{r_1}} \int_{B(y,r_1)} \rho(z, dk)dzdy & \text{if } |x_1 - x_2| < 2r_2, \\ 0 & \text{otherwise;} \end{cases}$$

where once more $V_r(x_1, x_2) := \mathrm{Vol}(B(x_1, r) \cap B(x_2, r))$ for $r > 0$. Furthermore, we define the map

$$\begin{aligned} &\Gamma^{\nu}(\rho)(x_1, x_2, dk_1, dk_2) \\ &= \int_{(0,1]\times(0,\infty)^2} u^2 V_{r_2}(x_1, x_2) \Big[[\rho]_{r_1,r_2}(x_1, x_2, dk_1)\delta_{k_1}(dk_2) \\ &\qquad - \rho(x_1, dk_1)[\rho]_{r_1,r_2}(x_1, x_2, dk_2) - [\rho]_{r_1,r_2}(x_1, x_2, dk_1)\rho(x_2, dk_2) \\ &\qquad\qquad + \rho(x_1, dk_1)\rho(x_2, dk_2) \Big] \nu(du, dr_1, dr_2). \end{aligned} \tag{3.25}$$

The symbols $[\rho]_{r_1,r_2}$ and Γ^{ν} are used to express the predictable variation process in the semimartingale below. More precisely, $[\rho]_{r_1,r_2}$ is the distribution of the parental type in an event given that both x_1 and x_2 are in the replacement ball of the event and Γ^{ν} measures the rate at which the covariance between $\rho_t(x_1, dk)$ and $\rho_t(x_2, dk)$ increases over time due to these reproduction events.

Lemma 3.4.1. *Let $(\rho_t, t \geq 0)$ be the SLFV with mutations. For any test function $\phi \in E^1$,*

$$\begin{aligned} \mathcal{M}_t(\phi) := \left\langle \rho_t, \phi \right\rangle - \left\langle \rho_0, \phi \right\rangle - \int_0^t \Bigg\{ &\mu \left\langle \lambda - \rho_s, \phi \right\rangle \\ &+ \int_{(0,1]\times(0,\infty)^2} u V_{r_2} \left\langle \rho_s, \overline{\overline{\phi}}(\cdot, r_1, r_2) - \phi \right\rangle \nu(du, dr_1, dr_2) \Bigg\} ds \end{aligned} \tag{3.26}$$

is a square-integrable martingale adapted to the natural filtration $(\mathcal{F}_t, t \geq 0)$ with predictable quadratic variation process

$$t \mapsto \int_0^t \left\langle \Gamma^{\nu}(\rho_s), \phi \otimes \phi \right\rangle ds.$$

Proof. The proof is a straightforward adaptation of the calculations of [For22, p.18f] to reproduction events with two radii r_1 and r_2. The finite variation term is a

consequence of

$$\begin{aligned}
&\lim_{\delta t\to 0}\frac{1}{\delta t}\mathbb{E}\left[\langle\rho_{t-\delta t},\phi\rangle-\langle\rho_t,\phi\rangle|\rho_t=\rho\right]\\
&=\mu\langle\lambda-\rho,\phi\rangle\\
&\qquad+\int_{\mathbb{R}^d}\int_{(0,1]\times(0,\infty)^2}\frac{1}{V_{r_1}}\int_{B(x,r_1)\times[0,1]}\langle u\mathbb{1}_{B(x,r_2)}(\delta_{k_0}-\rho),\phi\rangle\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\rho(y,dk_0)dy\nu(du,dr_1,dr_2)dx
\end{aligned}$$

and the exchange of the order of integration. The four terms of (3.25) originate from a similar calculation applied to

$$\begin{aligned}
&\lim_{\delta t\to 0}\frac{1}{\delta t}\mathbb{E}\left[(\langle\rho_{t-\delta t},\phi\rangle-\langle\rho_t,\phi\rangle)^2|\rho_t=\rho\right]\\
&=\int_{\mathbb{R}^d}\int_{(0,1]\times(0,\infty)^2}\frac{1}{V_{r_1}}\int_{B(x,r_1)\times[0,1]}\langle u\mathbb{1}_{B(x,r_2)}(\delta_{k_0}-\rho),\phi\rangle^2\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\rho(y,dk_0)dy\nu(du,dr_1,dr_2)dx.
\end{aligned}$$

□

As a next step, we will consider the martingale problem under the rescaling introduced in Section 3.3.2. For this purpose, we define the rescaled measure ν_N^α by

$$\begin{aligned}
&\int_{(0,N]\times(0,\infty)^2}f(u,r_1,r_2)\nu_N^\alpha(du,dr_1,dr_2)\\
&:=\int_{(0,1]\times(0,\infty)^2}f(Nu,\delta_N r_1,\delta_N r_2)\delta_N^{-(\alpha+d)}\nu_N(du,dr_1,dr_2),
\end{aligned}$$

and the operator $\mathcal{L}^N$ acting on E^1 by

$$\mathcal{L}^N\phi(x,k):=\int_{(0,N]\times(0,\infty)^2}uV_{r_2}\left(\overline{\overline{\phi}}(x,k,r_1,r_2)-\phi(x,k)\right)\nu_N^\alpha(du,dr_1,dr_2).$$

The idea to determine the rescaled martingale problem is to use the equation

$$\left\langle\,\boldsymbol{\rho}_t^N,\phi\,\right\rangle=\left\langle\,\rho_{\frac{Nt}{\delta_N^\alpha}},\phi_N\,\right\rangle,$$

where $\phi_N(x,k):=\delta_N^{\mathrm{d}}\phi(\delta_N x,k)$, i.e. rewrite $\left\langle\,\boldsymbol{\rho}_t^N,\phi\,\right\rangle$ as a version of $(\rho_t,t\geq 0)$ with accelerated time integrated against test functions with the correct spatial scaling.

Lemma 3.4.2 (extends [For22, p.19f])**.** *Let $(\boldsymbol{\rho}_t^N,t\geq 0)$ be the rescaled SLFV model as described in Section 3.3.2. Then for any $\phi\in E^1$*

$$\langle\boldsymbol{\rho}_t^N,\phi\,\rangle=\left\langle\,\boldsymbol{\rho}_0,\phi\,\right\rangle+\int_0^t\left\{\mu\left\langle\,\lambda-\boldsymbol{\rho}_s^N,\phi\,\right\rangle+\left\langle\,\boldsymbol{\rho}_s^N,\mathcal{L}^N\phi\,\right\rangle\right\}ds+\boldsymbol{\mathcal{M}}_t^N(\phi),$$

where $\boldsymbol{\mathcal{M}}_t^N(\phi)$ *is a square-integrable martingale with predictable quadratic variation given by*

$$\left\langle \boldsymbol{\mathcal{M}}_t^N(\phi) \right\rangle_t = \frac{1}{N} \int_0^t \left\langle \Gamma^{\nu_N^\alpha}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle ds.$$

Hereafter, we apply the scaling of the fluctuations

$$Z_t^N := \sqrt{N\eta_N}(\boldsymbol{\rho}_t^N - \lambda) \tag{3.27}$$

to the martingale problem. The resulting martingale problem falls within the martingale measures framework developed by Walsh in [Wal86]. Martingale measures generalise the notation of martingales and stochastic integrals to involve a spatial component (Definition A.2.1). Additionally, for a suitable class called worthy (see Definition A.2.2), powerful theorems about the convergence of stochastic integrals exist (see Theorem A.3.1, Theorem A.3.3). In this framework, we can rewrite the fluctuations as a sequence of stochastic integrals, and applications of the convergence theorems will lead to Theorem 3.3.4. For functions $f : [0,T] \times \mathbb{R}^d \times [0,1] \to \mathbb{R}$ and a worthy martingale measure M, we will use the abbreviation

$$M_t(f) := \int_{[0,t]\times\mathbb{R}^d\times[0,1]} f(s,x,k) M(ds,dx,dk).$$

Rewriting the semimartingale decomposition in Lemma 3.4.2 with this new terminology and applying the scaling of the fluctuations (3.16) results in the following representation.

Proposition 3.4.3 (extends [For22, Proposition 3.1])**.** *Assume that* $\rho_0^N = \lambda$*. For any* $\phi \in E^1$*, the fluctuations* $Z_t^N := \sqrt{N\eta_N}(\lambda - \boldsymbol{\rho}_t^N)$ *can be written as*

$$\left\langle Z_t^N, \phi \right\rangle = \int_0^t \left\langle Z_s^N, \mathcal{L}^N(\phi) - \mu\phi \right\rangle ds + M_t^N(\phi),$$

where M^N *is a worthy martingale measure on* $\mathbb{R}^d \times [0,1]$ *with*

$$\left\langle M^N(\phi) \right\rangle_t = \eta_N \int_0^t \left\langle \Gamma^{\nu_N^\alpha}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle ds.$$

Proof. The calculations and argumentation of [For22, p.19f] readily carry over to the new SLFV models. A dominating measure for the martingale measures is given by

$$D_N(dt, dx_1, dk_1, dx_2, dk_2) := \eta_N |\Gamma|^{\nu_\alpha^N}(\boldsymbol{\rho}_{t^-}^N)(x_1, x_2, dk_1, dk_2)dx_1dx_2dt, \tag{3.28}$$

where we define

$$\begin{aligned}
&|\Gamma|^\nu(\rho)(x_1, x_2, dk_1, dk_2)\\
&:= \int_{(0,1]\times(0,\infty)^2} u^2 V_{r_2}(x_1, x_2)\Big[[\rho]_{r_1,r_2}(x_1, x_2, dk_1)\delta_{k_1}(dk_2)\\
&\qquad + \rho(x_1, dk_1)[\rho]_{r_1,r_2}(x_1, x_2, dk_2) + [\rho]_{r_1,r_2}(x_1, x_2, dk_1)\rho(x_2, dk_2)\\
&\qquad\qquad + \rho(x_1, dk_1)\rho(x_2, dk_2)\Big]\nu(du, dr_1, dr_2).
\end{aligned}$$

□

The operator

$$\mathcal{L}^N\phi(x,k) = \int_{(0,1]\times(0,\infty)^2} uV_{r_2}\Big(\overline{\overline{\phi}}(x,k,r_1,r_2) - \phi(x,k)\Big)\nu_N^\alpha(du, dr_1, dr_2)$$

can be interpreted as the generator of a continuous-time jump process $(X_t^N, t \geq 0)$ describing the spatial motion of an ancestral lineage in the process with behaviour as follows. At rate

$$\int_{(0,1]\times(0,\infty)^2} uV_{r_2}\nu_N^\alpha(du, dr_1, dr_2)$$

the process jumps, if located at position x, to position $x + (Y_1R_1 + Y_2R_2)$, where (R_1, R_2) is a random variable taking values in $(0,\infty)^2$ with distribution

$$\frac{\int_{(0,1]} uV_{R_2}\nu_N^\alpha(du, dR_1, dR_2)}{\int_{(0,1]\times(0,\infty)^2} uV_{r_2}\nu_N^\alpha(du, dr_1, dr_2)},$$

and Y_1, Y_2 are two independent uniform random variables on $B(0,1)$. The operator $\mathcal{L}^N - \mu$ generates the strongly-continuous semigroup P_t^N acting on E^1, defined as

$$P_t^N\phi(x,k) := e^{-\mu t}\mathbb{E}_x\left[\phi(X_t^N, k)\right]. \tag{3.29}$$

Now, the semimartingale decomposition in Proposition 3.4.3 enables us to represent $\left\langle Z_t^N, \phi\right\rangle$ as a stochastic integral with respect to a martingale measure. Indeed, by [Wal86, Theorem 5.1], we can write

$$\langle Z_t^N, \phi\rangle = \int_{[0,t]\times\mathbb{R}^d\times[0,1]} P_{t-s}^N\phi(x,k)M^N(ds, dx, dk). \tag{3.30}$$

This representation, together with a convergence theorem which can be found in [Wal86], will play a crucial role in the proof of the central limit theorem.

3.4.2 Central limit theorem

In this subsection, we prove the convergence of $(Z_N, N \geq 1)$, i.e. the central limit theorem for our SLFV models, following the overall strategy of [For22, Section 3]. Classically, tightness of the sequence will show that any subsequence admits limit points, and we shall conclude by proving uniqueness of the finite-dimensional distributions of these limit points. Afterwards, in Section 3.4.2 we highlight the main points where additional arguments are needed compared to [For22]. In turn, the proofs of these main points are detailed in Section 3.5.

Tightness of the stochastic integrals

Here, we will apply Mitoma's theorem:

Theorem 3.4.4 ([Wal86, Theorem 6.13])**.** *Let $X = ((X_t^N)_{t\in[0,1]}, N \geq 1)$ be a sequence of stochastic processes in $D([0,1], \mathcal{S}'(\mathbb{R}^{\mathrm{d}}))$. Then X is tight if and only if for each Schwartz function $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}})$, the evaluated sequence $((X_t^N(\phi))_{t\in[0,1]}, N \geq 1)$ is tight in $D([0,1], \mathbb{R})$.*

This allows us to reduce tightness of $(Z_N, N \geq 0)$ to tightness of

$$\left(\left\langle Z_t^N, \phi \right\rangle_{t\geq 0}, N \geq 1\right) \subset D(\mathbb{R}_+, \mathbb{R})$$

for $\phi \in \mathcal{S}(\mathbb{R}^d \times [0,1])$. Theorem A.3.1 will show tightness of the latter sequences, when written in the stochastic integral form of (3.30). The vague convergence $\boldsymbol{\rho}_t^N \to \lambda$ will be an additional byproduct of a bound in this theorem and we give a short proof below. To apply Theorem A.3.1 to the sequence of stochastic integrals (3.30) we need to verify three conditions. The first one requires the dominating measures of the integrators M^N to be uniformly bounded, whereas the other two conditions describe sufficient properties of the integrands $P_{t-s}^N \phi$.

Lemma 3.4.5 (extends [For22, Lemma 3.4]). *Condition 1 of Theorem A.3.1: Let D^N be the dominating measure from* (3.28). *Then in the one-tail and two-tail regimes, we can find a constant $C > 0$ such that for all $N \geq 1, 0 \leq s \leq t$ and $\phi, \psi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$,*

$$\left\langle D^N, \mathbb{1}_{[s,t]}\phi \otimes \psi \right\rangle \leq C|t-s|\Big(\|\phi\|_1\|\psi\|_1 + \|\phi\|_2\|\psi\|_2\Big). \tag{3.31}$$

Lemma 3.4.6 (extends [For22, Lemma 3.5]). *Conditions 2 - 3 of Theorem A.3.1 + convergence of the semigroups: Let P_t^N and P_t be the semigroups associated with $\mathcal{L}^N - \mu$ and $\mathcal{D}_\alpha - \mu$ according to* (3.29).

1. *If $\phi \in \mathcal{S}(\mathbb{R}^d \times [0,1])$, then $P_t^N\phi \in \mathcal{S}(\mathbb{R}^d \times [0,1])$ and $t \mapsto P_t^N\phi$ is continuous.*
2. *Then for all $t \geq 0, N \geq 1, p \geq 1$ and $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$*

$$\left\|P_t^N\phi\right\|_p \leq e^{-\mu t}\|\phi\|_p$$

and for any $\kappa \in \mathbb{N}^d$, the derivatives are bounded by

$$\left\|\partial_\kappa P_t^N\phi\right\|_p \leq e^{-\mu t}\|\partial_\kappa\phi\|_p.$$

3. *Furthermore, there exists constants $C > 0$ and $\theta > 0$ such that for all $N \geq 1$ and $t \geq 0$*

$$\left\|P_t^N\phi - P_t\phi\right\|_p \leq C\delta_N^\theta t e^{-\mu t}\max_{|\kappa|\leq 4}\|\partial_\kappa\phi\|_p.$$

The convergence of the semigroups will be required to identify limit points in the next subsection. The proof of Lemma 3.4.6 will require the boundedness and convergence of the corresponding generators.

Lemma 3.4.7 (extends [FP17, Prosition A.3]). *Let $\mathcal{D}_\alpha$ be the generator* (3.17) *and $\phi : \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$ be a twice continuously-differentiable function with $\max_{|\kappa|\leq 4}\|\partial_\kappa\phi\|_p < \infty$ for $1 \leq p \leq \infty$. Then there exists constants $\theta, C, C' > 0$ such that*

$$\left\|\mathcal{L}^N\phi\right\|_p \leq C\max_{|\kappa|\leq 2}\|\partial_\kappa\phi\|_p \tag{3.32}$$

and

$$\left\|\mathcal{L}^N\phi - \mathcal{D}_\alpha\phi\right\|_p \leq C'\delta_N^\theta\max_{|\kappa|\leq 4}\|\partial_\kappa\phi\|_p. \tag{3.33}$$

To cover our new regimes, the proofs of Lemma 3.4.5 and Lemma 3.4.7 require significant changes and will therefore be detailed in Section 3.5. The remainder of this subsection aligns with the arguments of [For22, p.22f].

Proof of Lemma 3.4.6. Recall the definition of $P_t^N\phi$ in (3.29). As the random walk is invariant by translation

$$\mathbb{E}_x\left[\phi(X_t^N,k)\right]=\mathbb{E}_0\left[\phi(x+X_t^N,k)\right],$$

which by Fubini's theorem implies

$$\int_{\mathbb{R}^d}\mathbb{E}_x\left[\phi(X_t^N,k)\right]dx=\int_{\mathbb{R}^d}\phi(x,k)dx.$$

This allows us to write, by Jensen's inequality,

$$\begin{aligned}\|P_t^N\phi\|_p &\leq \left(\int_{\mathbb{R}^d}\sup_{k\in[0,1]}\left|e^{-\mu t}\mathbb{E}_x\left[\phi(X^N,k)\right]\right|^p dx\right)^{1/p}\\ &\leq e^{-\mu t}\left(\int_{\mathbb{R}^d}\mathbb{E}_x\left[\sup_{k\in[0,1]}\left|\phi(X_t^N,k)\right|^p\right]dx\right)^{1/p}\\ &\leq e^{-\mu t}\|\phi\|_p.\end{aligned}$$

The bounds on the derivatives and that the $P_t^N\phi$ are Schwartz functions follows from similar calculations. To prove property 3., let us consider the difference

$$g_t^N=P_t^N\phi-P_t\phi$$

whose time derivative is

$$\partial_t g_t^N=(u\mathcal{L}^{N,\alpha}-\mu)g_t^N+u(\mathcal{L}^N-\mathcal{D}_\alpha)P_t\phi.$$

As the difference g_0^N is zero, this yields the representation

$$g_t^N=u\int_0^t P_{t-s}^N(\mathcal{L}^N-\mathcal{D}_\alpha)P_s\phi ds.$$

Taking the L^p-norm on both sides, we get

$$\|g_t^N\|_p\leq u\int_0^t e^{-\mu(t-s)}\left\|(\mathcal{L}^N-\mathcal{D}_\alpha)P_s\phi\right\|_p ds,$$

and we can apply (3.33) in Lemma 3.4.7. □

We are now in a position to apply Theorem A.3.1 to obtain the tightness of the stochastic integrals and the convergence to Lebesgue measure.

Corollary 3.4.8. *Under the assumptions of Theorem A.3.1, the sequence*

$$\Big(\Big\langle Z_t^N, \phi \Big\rangle_{t\geq 0}, N \geq 1 \Big) \subset D(\mathbb{R}_+, \mathbb{R})$$

is tight for any $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$.

Proof. With minor adaptations (see [For22, Proof of Proposition 3.6]), Lemmas 3.4.5 and 3.4.6 prove conditions 1 - 3 of Theorem A.3.1, which yields the tightness of the stochastic integrals (3.30). □

Proof of Theorem 3.3.2. In addition to the tightness, Theorem A.3.1 shows that there exists a constant $C^* > 0$ such that for all $T > 0, N \geq 1$ and $\phi \in \mathcal{S}(\mathbb{R}^d \times [0,1])$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| \Big\langle Z_t^N, \phi \Big\rangle \right|^2 \right]^{1/2} \leq C^* \max_{p\in\{1,2\}} \max_{|\kappa|\leq 2} \|\partial_\kappa \phi\|_p. \tag{3.34}$$

This allows us to bound

$$\mathbb{E}\left[\sup_{t\in[0,T]} d(\boldsymbol{\rho}_t^N, \lambda) \right] \leq \frac{C^*}{\sqrt{N\eta_N}} \sum_{n=1}^{\infty} \frac{1}{2^n} \max_{q\in\{1,2\}} \max_{|\kappa|\leq 2} \|\partial_\kappa \phi_n\|_p,$$

where d is the vague metric on Ξ defined in (3.6) and ϕ_n the corresponding sequence of test functions. Using equation (3.8), the right-hand side is bounded by $\frac{C}{\sqrt{N\eta_N}}$ for some $C > 0$, proving the result. □

Uniqueness of finite-dimensional distributions

We need to prove the convergence of the integrands and integrators of (3.30) to show convergence of the finite-dimensional distributions. In the last subsection, we have seen the L^p-convergence of the integrand $P_t^N \to P_t$. The convergence in law of the martingale measures M^N will follow from Theorem A.3.2 applied to $M^N(\phi)$ for test functions $\phi \in \mathcal{S}(\mathbb{R}^d \times [0,1])$ with the limit given by

$$M(\phi) = \sqrt{\gamma} \int_{\mathbb{R}^d \times [0,1]} \phi(x,k) W(dxdk)$$

for the Wiener process W of Theorem 3.3.4. With both convergence of integrands and integrators at hand, by an adaptation Theorem A.3.3 of [Wal86, Theorem 7.12] this suffices for the convergence of the finite dimensional distributions

$$\Big(\Big\langle Z^N_{t_i}, \phi_i \Big\rangle \Big)_{i=1,\dots,n} \xrightarrow[N\to\infty]{d} \left(\int_{[0,t_i]\times\mathbb{R}^{\mathrm{d}}\times[0,1]} P_{t_i-s}\phi_i(x,k) M(ds,dx,dk) \right)_{i=1,\dots,n}$$

for $t_1, ..., t_n \geq 0$ and functions $\phi_1, ..., \phi_n \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$. In turn, by [Wal86, Theorem 6.15] we will conclude that the sequence $((Z^N_t)_{t\geq 0}, N \geq 1)$ converges in law in $D(\mathbb{R}_+, \mathcal{S}'(\mathbb{R}^{\mathrm{d}} \times [0,1]))$ to the solution $(Z_t, t \geq 0)$ of (3.22) defined by

$$\Big\langle Z_t, \phi \Big\rangle = \int_{[0,t]\times\mathbb{R}^{\mathrm{d}}\times[0,1]} P_{t-s}\phi(x,k) M(ds,dx,dk). \tag{3.35}$$

We separate the prerequisites of the martingale convergence theorem into two lemmas:

Lemma 3.4.9 (extends [For22, Lemma 3.10 Part 1])**.** *For any Schwartz function $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$ the jumps of the martingale measures*

$$\sup_{t\geq 0} \Big| M^N_t(\phi) - M^N_{t^-}(\phi) \Big| \xrightarrow[N\to\infty]{} 0$$

vanish almost surely.

Lemma 3.4.10 (extends [For22, Lemma 3.10 Part 2])**.** *Let $\mathcal{Q}$ be defined as*

$$\mathcal{Q}(dx_1 dk_1 dx_2 dk_2) = \gamma K_\beta(dx_1, dx_2)\Big(dk_1 \delta_{k_1}(dk_2) - dk_1 dk_2\Big), \tag{3.36}$$

where γ and $K_\beta(x_1, x_2)$ are defined as in Table 3.1 and (3.19)*. Then for any fixed $t \geq 0$ and for any Schwartz function $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$ the covariation process converges in probability*

$$\Big\langle M^N(\phi) \Big\rangle_t \xrightarrow[N\to\infty]{} t \Big\langle \mathcal{Q}, \phi \otimes \phi \Big\rangle.$$

We postpone the proofs of Lemma 3.4.9 and Lemma 3.4.10 to Section 3.5.3, as both require adaptations to hold for the generalised regimes. Following [For22], the limiting martingale measure of $(M^N, N \geq 1)$ is then characterised by:

Lemma 3.4.11 (extends [For22, Lemma 3.9])**.** *Suppose M is a continuous martingale measure on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times [0,1]$ with covariation*

$$dt\delta_t(ds)\gamma K_\beta(dx_1, dx_2)\Big(dk_1\delta_{k_1}(dk_2) - dk_1 dk_2\Big).$$

Then the sequence $(M^N, N \geq 1)$ converges in distribution to M in $D(\mathbb{R}_+, \mathcal{S}'(\mathbb{R}^{\mathrm{d}} \times [0,1]))$.

Proof. Let $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$ be a Schwartz function and consider the sequence $((M_t^N(\phi)_{t\geq 0}, N \geq 1)$. Lemma 3.4.9 and Lemma 3.4.10 amount to the two conditions of Theorem A.3.2, yielding the convergence of $(M_t^N(\phi), t \geq 0)$ to $(M_t(\phi), t \geq 0)$ in distribution in $D(\mathbb{R}_+, \mathbb{R})$.
Now let $\phi_1, ..., \phi_k \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$ be Schwartz functions. Since both $\left\langle M^N(\phi_i) \right\rangle_t$ and $\left\langle M^N(\phi_j) \right\rangle_t$ converge, the covariation $\left\langle M^N(\phi_i), M^N(\phi_j) \right\rangle_t$ converges by polarization. Applying Theorem A.3.2 again, we get convergence of finite-dimensional distributions $(M_t^N(\phi_1), ..., M_t^N(\phi_k), t \geq 0)$. Consequently, $(M^N, N \geq 1)$ satisfies the prerequisites of [Wal86, Theorem 6.15] and the convergence of $(M^N, N \geq 1)$ to M follows. □

Summary of significant changes required to the proofs of [For22]

In order to unify the arguments, we will generalise the various lemmas of the proof of [For22] to the one- and two-tail regimes. While the general form of the lemmas has been preserved, the proofs require non-trivial adaptations, which we provide in Section 3.5. In summary:

1. Lemma 3.4.5 deals with the bound on the dominating measures. Here, we are able to use the same bounds in new combinations to cover the new regimes.

2. Prior calculations cannot be transfered to Lemma 3.4.7, because in the one- and two-tail regime one of the averages of $\mathcal{L}^N$ vanishes asymptotically.

3. For general sequences $(\delta_N, N \geq 1)$ converging to zero, the proof of Lemma 3.4.9 does not carry over. It is necessary to impose the new conditions (3.20) and (3.21) on the convergence speed of δ_N relative to N in the one- respectively two-tail regime.

4. Lemma 3.4.10 analyses the quadratic variation of the martingale measures. If $\beta < \mathrm{d}$, convergence becomes much harder to justify, which led us to develop the auxiliary Corollary 3.5.1. The case $\beta > \mathrm{d}$ has been adapted to two different radii and the case $\beta = \mathrm{d}$ involves new calculations to account for the novel logarithmic scaling $\eta_N = \delta_N^{\alpha-\mathrm{d}} \log^{-1}(1/\delta_N)$.

3.4.3 Derivation of the Wright-Malécot formulae

Let us explain how one derives the Wright-Malécot formula from the central limit theorem, analogously to [For22, Section 4]. We are interested in $P_t^N(\phi, \psi)$, the probability that two individuals sampled according to ϕ, ψ from the N-th rescaled SLFV model $(\boldsymbol{\rho}_t^N, t \geq 0)$ are of the same type. The probability densities ϕ, ψ are assumed to be smooth and compactly supported. Recall from Section 3.3.4 that

$$P_t^N(\phi, \psi) = \mathbb{E}\Big[\Big\langle \boldsymbol{\rho}_t^N \otimes \boldsymbol{\rho}_t^N, (\phi \otimes \psi)\mathbb{1}_\Delta \Big\rangle \Big]$$

where $\mathbb{1}_\Delta : [0,1]^2 \to \mathbb{R}$ is defined as $\mathbb{1}_\Delta(k_1, k_2) = \mathbb{1}_{k_1 = k_2}$. Since $Z_t^N := \sqrt{N\eta_N}(\boldsymbol{\rho}_t^N - \lambda)$ and Lebesgue measure λ does not charge the diagonal $\{(k_1, k_2) : k_1 = k_2\}$, this is equal to

$$P_t^N(\phi, \psi) = (N\eta_N)^{-1}\mathbb{E}\Big[\Big\langle Z_t^N \otimes Z_t^N, (\phi \otimes \psi)\mathbb{1}_\Delta \Big\rangle \Big].$$

As a first step, we need to justify that the exchange of the limit $N \to \infty$ with the expectation is possible.

Lemma 3.4.12. *For $t \geq 0$, two compactly supported and smooth densities ϕ, ψ and Z_t the solution to* (3.22) *we have*

$$\lim_{N\to\infty} N\eta_N P_t^N(\phi, \psi) = \mathbb{E}\Big[\Big\langle Z_t \otimes Z_t, (\phi \otimes \psi)\mathbb{1}_\Delta \Big\rangle \Big].$$

Proof. From Proposition 3.4.3 and Equation (3.30) we know

$$N\eta_N P_t^N(\phi, \psi) = \mathbb{E}\left[\eta_N \int_0^t \Big\langle \Gamma^{\nu_N^\alpha}(\boldsymbol{\rho}_s^N), \Big(P_{t-s}^N\phi \otimes P_{t-s}^N\psi\Big)\mathbb{1}_\Delta \Big\rangle ds \right].$$

Combining Lemma 3.4.5 and Lemma 3.4.6 yields

$$\begin{aligned} \Big| \Big\langle \eta_N \Gamma^{\nu_N^\alpha}(\boldsymbol{\rho}_s^N), \Big(P_{t-s}^N\phi \otimes P_{t-s}^N\psi\Big)\mathbb{1}_\Delta \Big\rangle - \Big\langle \eta_N \Gamma^{\nu_N^\alpha}(\boldsymbol{\rho}_s^N), \Big(P_{t-s}\phi \otimes P_{t-s}\psi\Big)\mathbb{1}_\Delta \Big\rangle \Big| \\ \leq C|t-s|^3 \delta_N^\theta e^{-2\mu(t-s)}. \end{aligned}$$

As in Lemma 3.4.10 one can then deduce

$$\eta_N \int_0^t \Big\langle \Gamma^{\nu_N^\alpha}(\boldsymbol{\rho}_s^N), \big(P_{t-s}\phi \otimes P_{t-s}\psi\big)\mathbb{1}_\Delta \Big\rangle\, ds \xrightarrow[N\to\infty]{} \int_0^t \Big\langle \mathcal{Q}, \big(P_s\phi \otimes P_s\psi\big)\mathbb{1}_\Delta \Big\rangle\, ds.$$

Again by Lemma 3.4.5 and Lemma 3.4.6, the dominated convergence theorems allows to exchange limit and expectation. □

Proof of Theorem 3.3.7. From the previous proof we know

$$\mathbb{E}\Big[\Big\langle Z_t \otimes Z_t, (\phi\otimes\psi)\mathbb{1}_\Delta \Big\rangle\Big] = \int_0^t \Big\langle \mathcal{Q}, \big(P_s\phi \otimes P_s\psi\big)\mathbb{1}_\Delta \Big\rangle\, ds.$$

The integral on the right-hand side stays finite by Lemma 3.4.5 and Lemma 3.4.6 as $t \to \infty$, which shows

$$\lim_{t\to\infty} \mathbb{E}\Big[\Big\langle Z_t \otimes Z_t, (\phi\otimes\psi)\mathbb{1}_\Delta \Big\rangle\Big] = \int_0^\infty \Big\langle \mathcal{Q}, \big(P_s\phi \otimes P_s\psi\big)\mathbb{1}_\Delta \Big\rangle\, ds.$$

If we let G_t^α be the fundamental solution associated to $\partial_t - \mathcal{D}_\alpha$, we can rewrite the above equation with the representation

$$P_t\phi(x,k) = e^{-\mu t}\int_{\mathbb{R}^d} G_{ut}^\alpha(x-y)\phi(y,k)dy.$$

The expressions of Theorem 3.3.7 follow as $\mathbb{1}_\Delta$ integrated with respect to the types part of $\mathcal{Q}$ found in (3.36) yields

$$\int_{[0,1]^2} \mathbb{1}_\Delta(k_1,k_2)(\delta_{k_1}(dk_2)dk_1 - dk_1dk_2) = 1.$$ □

3.5 Proofs of the lemmas

This section collects all proofs requiring significant changes to generalise [For22] to the new regimes.

3.5.1 Bound on dominating measures

Proof of Lemma 3.4.5. The first estimates are the same as in [For22, p.20f], then we use new combinations of the bounds for the various cases. Recall

$$\Big\langle D^N, \mathbb{1}_{[s,t]}\phi\otimes\psi \Big\rangle = \int_s^t \Big\langle \eta_N|\Gamma|^{\nu_N^\alpha}(\boldsymbol{\rho}_v^N), \phi\otimes\psi \Big\rangle\, dv.$$

Considering the four summands of $|\Gamma|^{\nu_N^\alpha}$ separately, we can bound $\left\langle D^N, \mathbb{1}_{[s,t]}\phi\otimes\psi \right\rangle$ by

$$\begin{aligned} 4|t-s| \int_{(0,1]\times(0,\infty)^2} \eta_N(Nu)^2 \int_{(\mathbb{R}^{\mathrm{d}})^2} V_{\delta_N r_2}(x_1,x_2) \\ \times \sup_{k\in[0,1]} |\phi(x_1,k)| \sup_{k\in[0,1]} |\psi(x_2,k)| dx_1 dx_2 \delta_N^{-(\mathrm{d}+\alpha)} \nu_N(du,dr_1,dr_2). \end{aligned} \tag{3.37}$$

We defined L^p-norms in (3.7) to involve the supremum over $k\in[0,1]$ and thus applying $V_{\delta_N r_2}(x_1,x_2) \le V_{\delta_N r_2}$ the inner integral is bounded by $V_{\delta_N r_2}\|\phi\|_1\|\psi\|_1$. Alternatively, the Cauchy-Schwarz inequality results in

$$\int_{(\mathbb{R}^{\mathrm{d}})^2} V_{\delta_N r_2}(x_1,x_2) \sup_{k\in[0,1]} |\phi(x_1,k)| \sup_{k\in[0,1]} |\psi(x_2,k)| dx_1 dx_2 \le V_{\delta_N r_2}^2 \|\phi\|_2\|\psi\|_2, \tag{3.38}$$

since

$$\int_{\mathbb{R}^{\mathrm{d}}} V_{\delta_N r_2}(x_1,x_2) dx_2 = V_{\delta_N r_2}^2.$$

We only demonstrate how the two estimates can be used to bound (3.37) for the one-tail regime. The two-tail regime follows in a similar fashion, as the double integrals can be written as a product of two integrals.

One-tail regime If $\beta = \mathrm{a}+\mathrm{c} < d$, we split the integral in (3.37) and apply both bounds to obtain

$$\begin{aligned} 4|t-s| \Bigg\{ & \int_1^{1/\delta_N} \delta_N^{\alpha-\mathrm{a}-\mathrm{c}} \frac{u_0^2}{r_2^{2\,\mathrm{c}}} V_{\delta_N r_2}^2 \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} \|\phi\|_2\|\psi\|_2 \\ & + \int_{1/\delta_N}^\infty \delta_N^{\alpha-\mathrm{a}-\mathrm{c}} \frac{u_0^2}{r_2^{2\,\mathrm{c}}} V_{\delta_N r_2} \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} \|\phi\|_1\|\psi\|_1 \Bigg\}. \end{aligned}$$

The first integral can be bounded by a finite quantity independent of N

$$\begin{aligned} \int_1^{1/\delta_N} \delta_N^{\alpha-\mathrm{a}-\mathrm{c}} \frac{u_0^2}{r_2^{2\,\mathrm{c}}} V_{\delta_N r_2}^2 \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} &= u_0^2 \int_1^{1/\delta_N} V_{\delta_N r_2}^2 \delta_N \frac{dr_2}{(\delta_N r_2)^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\ &= u_0^2 V_1^2 \int_{\delta_N}^1 \frac{dr_2}{r_2^{1+\beta-\mathrm{d}}} \le \frac{u_0^2 V_1^2}{\mathrm{d}-\beta}, \end{aligned}$$

since $\beta < \mathrm{d}$. The second integral is equal to

$$\begin{aligned} \int_{1/\delta_N}^\infty \delta_N^{\alpha-\mathrm{a}-\mathrm{c}} \frac{u_0^2}{r_2^{2\,\mathrm{c}}} V_{\delta_N r_2} \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} &= u_0^2 \int_{1/\delta_N}^\infty V_{\delta_N r_2} \delta_N \frac{dr_2}{(\delta_N r_2)^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\ &= u_0^2 V_1 \int_1^\infty \frac{dr_2}{r_2^{1+\mathrm{a}+\mathrm{c}}} = \frac{u_0^2 V_1}{\beta}. \end{aligned}$$

Similarly, if $\beta = \mathrm{d}$, the first integral is bounded by

$$\begin{aligned}
&\int_1^{1/\delta_N} \frac{\delta_N^{\alpha-\mathrm{d}}}{\log(1/\delta_N)} \frac{1}{r_2^{2\,\mathrm{c}}} V^2_{\delta_N r_2} \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} \|\phi\|_2 \|\psi\|_2 \\
&= \log(1/\delta_N)^{-1} V_1^2 \int_1^{1/\delta_N} \frac{dr_2}{r_2} \|\phi\|_2 \|\psi\|_2 \\
&= V_1^2 \|\phi\|_2 \|\psi\|_2,
\end{aligned}$$

and the second integral can be bounded by

$$\begin{aligned}
&\int_{1/\delta_N}^{\infty} \frac{\delta_N^{\alpha-\mathrm{d}}}{\log(1/\delta_N)} \frac{1}{r_2^{2\,\mathrm{c}}} V_{\delta_N r_2} \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} \|\phi\|_1 \|\psi\|_1 \\
&= \frac{\delta_N^{-\mathrm{d}}}{\log(1/\delta_N)} V_1 \int_{1/\delta_N}^{\infty} \frac{dr_2}{r_2^{1+\mathrm{d}}} \|\phi\|_1 \|\psi\|_1 \\
&= \frac{V_1 \|\phi\|_1 \|\psi\|_1}{\mathrm{d} \log(1/\delta_N)}.
\end{aligned}$$

If $\beta > \mathrm{d}$, we can apply the Cauchy-Schwarz bound directly to (3.37)

$$\begin{aligned}
\int_1^{\infty} \delta_N^{\alpha-\mathrm{d}} \frac{1}{r_2^{2\,\mathrm{c}}} V^2_{\delta_N r_2} \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} \|\phi\|_2 \|\psi\|_2 &= V_1^2 \int_1^{\infty} \frac{dr_2}{r_2^{1+\mathrm{a}+\mathrm{c}-\mathrm{d}}} \|\phi\|_2 \|\psi\|_2 \\
&= \frac{V_1^2 \|\phi\|_2 \|\psi\|_2}{\beta - \mathrm{d}}.
\end{aligned}$$

This concludes the proof of the lemma. □

In the proof of Lemma 3.4.10, we will require a version of Lemma 3.4.5 for functions $\Psi \in \mathcal{S}((\mathbb{R}^d)^2 \times [0,1]^2)$.

Corollary 3.5.1. *Let D^N be the dominating measure from* (3.28)*. Then in the one-tail and two-tail regimes we can find constants $C > 0, q > 1$ such that for all $N \geq 1, 0 \leq s \leq t$ and $\Psi \in \mathcal{S}((\mathbb{R}^d)^2 \times [0,1]^2)$*

$$\left\langle D^N, \mathbb{1}_{[s,t]} \Psi \right\rangle \leq C|t-s| \left(\|\Psi\|_1 + \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \left| \Psi(x_1, x_2) \right|^q dx_2 \right)^{1/q} dx_1 \right), \tag{3.39}$$

where we abbreviate $\sup_{k_1,k_2 \in [0,1]^2} \Psi(x_1, x_2, k_1, k_2)$ *by* $\Psi(x_1, x_2)$*.*

Proof. To prove (3.39), we need to bound the inner integral of (3.37) for functions $\Psi \in \mathcal{S}((\mathbb{R}^d)^2 \times [0,1]^2)$. Using $V_{\delta_N r_2}(x_1, x_2) \leq V_{\delta_N r_2}$, we obtain the bound $V_{\delta_N r_2} \|\Psi\|_1$.

Alternatively, we can use Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$ and set $\xi = 1/p \in [0,1)$

$$\begin{aligned}
&\int_{(\mathbb{R}^{\mathrm{d}})^2} V_{\delta_N r_2}(x_1,x_2)\Psi(x_1,x_2)dx_1dx_2\\
&\quad\le \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} V_{\delta_N r_2}(x_1,x_2)^{p\xi}dx_2\right)^{1/p}\left(\int_{\mathbb{R}^d}\underbrace{V_{\delta_N r_2}^{q(1-\xi)}(x_1,x_2)}_{\le V_{\delta_N r_2}}\left|\Psi(x_1,x_2)\right|^q dx_2\right)^{1/q}dx_1\\
&\quad\le V_{\delta_N r_2}^{\frac{2}{p}+\frac{1}{q}}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\left|\Psi(x_1,x_2)\right|^q dx_2\right)^{1/q}dx_1.
\end{aligned}$$

We can choose $\frac{2}{p}+\frac{1}{q}$ arbitrarily close to 2, which allows us to apply the same arguments as in Lemma 3.4.5. □

3.5.2 Convergence of generators

In this section, we show the convergence of the generator $\mathcal{L}^N$ from (3.27) given by

$$\mathcal{L}^N\phi(x) = N\delta_N^{-\alpha}\int_{(0,N]\times(0,\infty)^2} uV_{r_2}\left(\overline{\overline{\phi}}(x,\delta_N r_1,\delta_N r_2)-\phi(x)\right)\nu_N(du,dr_1,dr_2)$$

to the fractional Laplacians $\mathcal{D}_\alpha$ defined in (3.17). Since the operators $\mathcal{L}^N$ only act on the spatial variable, we omit the dependence on the second variable k in this whole section. The main idea of the proof is that under our scaling one of the averaging radii is much smaller than the other, and in the limit only the largest averaging radius matters. The only case where the two radii are precisely the same, that is $\mathrm{b}=1$ in the one-tail regime, was already covered by [FP17, Appendix A]. The proof will require the following lemmas.

Lemma 3.5.2. *Assume that $\phi:\mathbb{R}^{\mathrm{d}}\to\mathbb{R}$ has continuous derivatives up to the second order and $\max_{|\kappa|\le 2}\|\partial_\kappa\phi\|_q$ is finite for $1\le q\le\infty$. Then for all $1\le p\le\infty$*

$$\left\|\overline{\phi}(r)-\phi\right\|_p \le \min\left\{2\|\phi\|_p, \frac{\mathrm{d}}{2}r^2\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p\right\}. \tag{3.40}$$

Proof. The first estimate of (3.40) stems from $\|\overline{\phi}(r)\|_p\le\|\phi\|_p$. The second estimate of (3.40) is derived by Taylor's theorem in [FP17, Proposition A.1]. □

Lemma 3.5.3. *Assume that $\phi:\mathbb{R}^{\mathrm{d}}\to\mathbb{R}$ has continuous derivatives up to the fourth order and $\max_{p\in[1,\infty]}\max_{|\kappa|=4}\|\partial_\kappa\phi\|_p$ is finite. Then for all $1\le p\le\infty$ and all $r,s\in[0,\infty)$*

$$\left\|\overline{\overline{\phi}}(x,r,s)-\phi(x)-\frac{r^2+s^2}{2(d+2)}\Delta\phi(x)\right\|_p \le \frac{d^3}{3!}\max_{|\kappa|=4}\|\partial_\kappa\phi\|_p(r^4+s^4). \tag{3.41}$$

Proof. This is a generalization of [FP17, Proposition A.1] to two radii. □

Proof of Lemma 3.4.7. In the following proof, we focus on the more difficult second bound (3.33). The first bound (3.32) can be shown using the same estimates.

One-tail regime In the one-tail regime, substituting ν_N by its expression in (3.12), we can write

$$\begin{aligned}\mathcal{L}^N\phi(x) &= \delta_N^{-\alpha}\int_1^\infty \frac{u_0}{r^c}V_r\left(\overline{\overline{\phi}}(x,\delta_N r^b,\delta_N r)-\phi(x)\right)\frac{dr}{r^{1+d+a-c}}\\ &= u_0V_1\delta_N^{-\alpha}\int_1^\infty\left(\overline{\overline{\phi}}(x,\delta_N r^b,\delta_N r)-\phi(x)\right)\frac{dr}{r^{1+a}}.\end{aligned} \tag{3.42}$$

The Brownian case Suppose that $a>2$ and $\frac{a}{b}>2$, then $\alpha=2$. Let $\sigma>0$ be defined as

$$\begin{aligned}\sigma^2 &= \frac{u_0V_1}{d+2}\int_1^\infty (r^{2\,\mathrm{b}}+r^2)\frac{dr}{r^{1+\mathrm{a}}}\\ &= \frac{u_0V_1}{d+2}\left(\frac{1}{\mathrm{a}-2\,\mathrm{b}}+\frac{1}{\mathrm{a}-2}\right).\end{aligned}$$

Let also $(A_N, N\geq 1)$ denote a sequence taking values in $(1,\infty]$ which will be chosen later. We can then bound, using the definition of σ^2,

$$\begin{aligned}&\left\|\mathcal{L}^N\phi-\frac{\sigma^2}{2}\Delta\phi\right\|_p\\ &\leq u_0V_1\int_1^{A_N}\left\|\overline{\overline{\phi}}(\cdot,\delta_N r_2^{\mathrm{b}},\delta_N r)-\phi-\frac{\delta_N^2 r^{2\,\mathrm{b}}+\delta_N^2 r^2}{2(\mathrm{d}+2)}\Delta\phi\right\|_p\delta_N^{-2}\frac{dr}{r^{1+\mathrm{a}}}\\ &+u_0V_1\int_{A_N}^\infty\left\|\overline{\overline{\phi}}(\cdot,\delta_N r^b,\delta_N r)-\phi\right\|_p\delta_N^{-2}\frac{dr}{r^{1+\mathrm{a}}}+\frac{u_0V_1}{2(d+2)}\int_{A_N}^\infty(r^{2b}+r^2)\frac{dr}{r^{1+\mathrm{a}}}\|\Delta\phi\|_p.\end{aligned}$$

Using (3.41) in the first integral and (3.40) in the second we obtain

$$\begin{aligned}&\left\|\mathcal{L}^N\phi-\frac{\sigma^2}{2}\Delta\phi\right\|_p\\ &\leq C\left(\delta_N^2\int_1^{A_N}(r^{4b}+r^4)\frac{dr}{r^{1+\mathrm{a}}}+2\int_{A_N}^\infty(r^{2b}+r^2)\frac{dr}{r^{1+\mathrm{a}}}\right)\max_{0\leq|\kappa|\leq 4}\|\partial_\kappa\phi\|_p,\end{aligned} \tag{3.43}$$

for some constant $C>0$. First note that, since $a>2b$ and $a>2$,

$$\int_{A_N}^\infty(r^{2b}+r^2)\frac{dr}{r^{1+\mathrm{a}}}=\frac{1}{(a-2b)A_N^{a-2b}}+\frac{1}{(a-2)A_N^{a-2}}.$$

On the other hand

$$\int_1^{A_N} r^4 \frac{dr}{r^{1+\mathrm{a}}} = \begin{cases} \frac{A_N^{4-a}-1}{4-a} & a \neq 4, \\ \log(A_N) & a = 4, \end{cases}, \qquad \int_1^{A_N} r^{4b} \frac{dr}{r^{1+\mathrm{a}}} = \begin{cases} \frac{A_N^{4b-a}-1}{4b-a} & a \neq 4b, \\ \log(A_N) & a = 4b. \end{cases}$$

We can thus choose $A_N = \infty$ if $a > 4 \vee 4b$, and $A_N = \delta_N^{-\xi}$ with $0 < \xi < \frac{2}{4-a}$ if $a \leq 4$ and $0 < \xi < \frac{2}{4b-a}$ if $a \leq 4b$. In each case, the right hand side of (3.43) is bounded by $C\delta_N^\theta \max_{0\leq|\kappa|\leq 4} \|\partial_\kappa \phi\|_p$ for some constant $C > 0$.

The stable case When $b < 1$ and $\mathrm{a} < 2$, $\alpha = \mathrm{a}$ and, changing variables in the integral over r in (3.42), we obtain

$$\mathcal{L}^N \phi(x) = u_0 V_1 \int_{\delta_N}^\infty \left(\overline{\overline{\phi}}(x, \delta_N^{1-b} r^b, r) - \phi(x)\right) \frac{dr}{r^{1+a}}. \tag{3.44}$$

On the other hand, if $b > 1$ and $\frac{a}{b} < 2$, $\alpha = \frac{a}{b}$ and we instead write

$$\mathcal{L}^N \phi(x) = \frac{u_0 V_1}{b} \int_{\delta_N}^\infty \left(\overline{\overline{\phi}}(x, r, \delta_N^{1-1/b} r^{1/b}) - \phi(x)\right) \frac{dr}{r^{1+\frac{a}{b}}}.$$

Note that the two expressions above are of the same form, with $1/b$ in the second one playing the role of b in the first one, so that the case $b > 1$ can be treated in exactly the same way as the case $b > 1$. The case $b = 1$ has already been covered in [FP17].

We start by detailing the case $b < 1$. We see that, in (3.44), $\delta_N^{1-b} r^b \to 0$ as $N \to \infty$. Let us set

$$\mathcal{D}^N \phi(x) := u_0 V_1 \int_{\delta_N}^\infty \left(\overline{\phi}(x, r) - \phi(x)\right) \frac{dr}{r^{1+\mathrm{a}}}.$$

As a result,

$$\|\mathcal{L}^N \phi - \mathcal{D}^N \phi\|_p \leq u_0 V_1 \int_{\delta_N}^\infty \|\overline{\overline{\phi}}(\cdot, \delta_N^{1-b} r^b, r) - \overline{\phi}(\cdot, r)\|_p \frac{dr}{r^{1+\mathrm{a}}}.$$

If $b = 0$, we can use (3.40) to obtain

$$\begin{aligned} \|\mathcal{L}^N \phi - \mathcal{D}^N \phi\|_p &\leq C\delta_N^2 \int_{\delta_N}^\infty \frac{dr}{r^{1+\mathrm{a}}} \max_{|\kappa|=2} \|\partial_\kappa \phi\|_p \\ &\leq C\delta_N^{2-\mathrm{a}} \max_{|\kappa|=2} \|\partial_\kappa \phi\|_p. \end{aligned}$$

If $b > 0$, using (3.40) for $r \leq \delta_N^{-\frac{1-b}{b}}$, we obtain, for $a \neq 2b$,

$$\begin{aligned}\|\mathcal{L}^N\phi - \mathcal{D}^N\phi\|_p &\leq C\left(\delta_N^{2(1-b)}\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p\int_{\delta_N}^{\delta_N^{-\frac{1-b}{b}}} r^{2b-a-1}dr + 2\|\phi\|_p\int_{\delta_N^{-\frac{1-b}{b}}}^{\infty}\frac{dr}{r^{1+\mathrm{a}}}\right)\\ &\leq C\left(\delta_N^{2(1-b)}\frac{\delta_N^{-(2b-a)\frac{1-b}{b}} - \delta_N^{2b-a}}{2b-a} + \frac{2}{a}\delta_N^{\frac{a(1-b)}{b}}\right)\max_{0\leq|\kappa|\leq 2}\|\partial_\kappa\phi\|_p\\ &\leq C\left(\delta_N^{\frac{a(1-b)}{b}} + \delta_N^{2-a}\right)\max_{0\leq|\kappa|\leq 2}\|\partial_\kappa\phi\|_p.\end{aligned}$$

If $a = 2b$, we instead obtain

$$\|\mathcal{L}^N\phi - \mathcal{D}^N\phi\|_p \leq C\left(\delta_N^{2(1-b)}|\log(\delta_N)| + \delta_N^{\frac{a(1-b)}{b}}\right)\max_{0\leq|\kappa|\leq 2}\|\partial_\kappa\phi\|_p.$$

Since $\mathrm{a} < 2$ and $b < 1$, the right hand side is of the form $C\delta_N^\theta \max_{0\leq|\kappa|\leq 2}\|\partial_\kappa\phi\|_p$ for some constant $C > 0$.

It thus remains to bound $\|\mathcal{D}^N\phi - \mathcal{D}_\alpha\phi\|_p$. Using (3.40),

$$\begin{aligned}\|\mathcal{D}^N\phi - \mathcal{D}_\alpha\phi\|_p &\leq u_0V_1\int_0^{\delta_N}\|\overline{\phi}(\cdot,r) - \phi\|_p\frac{dr}{r^{1+\mathrm{a}}}\\ &\leq C\int_0^{\delta_N} r^{2-\mathrm{a}-1}dr\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p\\ &\leq C\delta_N^{2-\mathrm{a}}\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p,\end{aligned}$$

for some constant $C > 0$. This concludes the proof of (3.33).

Two-tails regime In the two-tails regime, similar calculations as in (3.42) lead to

$$\mathcal{L}^N\phi(x) = u_0V_1\delta_N^{-\alpha}\int_1^\infty\int_1^\infty\left(\overline{\overline{\phi}}(x,\delta_Nr_1,\delta_Nr_2) - \phi(x)\right)\frac{dr_1dr_2}{r_1^{1+a_1}r_2^{1+a_2}}. \tag{3.45}$$

The Brownian case Suppose that $\mathrm{a}_1 > 2$ and $\mathrm{a}_2 > 2$. Then $\alpha = 2$ and we set

$$\begin{aligned}\sigma^2 &= \frac{u_0V_1}{d+2}\int_1^\infty\int_1^\infty(r_1^2+r_2^2)\frac{dr_1dr_2}{r_1^{1+\mathrm{a}_1}r_2^{1+\mathrm{a}_2}}\\ &= \frac{u_0V_1}{d+2}\left(\frac{1}{(\mathrm{a}_1-2)\,\mathrm{a}_2} + \frac{1}{\mathrm{a}_1(\mathrm{a}_2-2)}\right).\end{aligned}$$

Then for $A_N \in (1, \infty]$,

$$\begin{aligned}
&\left\| \mathcal{L}^N \phi - \frac{\sigma^2}{2} \Delta \phi \right\|_p \\
&\leq u_0 V_1 \int_1^{A_N} \int_1^{A_N} \left\| \overline{\overline{\phi}}(\cdot, \delta_N r_1, \delta_N r_2) - \phi - \frac{\delta_N^2 r_1^2 + \delta_N^2 r_2^2}{2(d+2)} \Delta \phi \right\|_p \delta_N^{-2} \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} \\
&\quad + u_0 V_1 \int_1^{A_N} \int_{A_N}^{\infty} \left(\left\| \overline{\overline{\phi}}(\cdot, \delta_N r_1, \delta_N r_2) - \phi \right\|_p \delta_N^{-2} + \frac{r_1^2 + r_2^2}{2(d+2)} \|\Delta \phi\|_p \right) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} \\
&\quad + u_0 V_1 \int_{A_N}^{\infty} \int_1^{\infty} \left(\left\| \overline{\overline{\phi}}(\cdot, \delta_N r_1, \delta_N r_2) - \phi \right\|_p \delta_N^{-2} + \frac{r_1^2 + r_2^2}{2(d+2)} \|\Delta \phi\|_p \right) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}}.
\end{aligned}$$

Using (3.41) and (3.40), we obtain

$$\begin{aligned}
&\left\| \mathcal{L}^N \phi - \frac{\sigma^2}{2} \Delta \phi \right\|_p \leq C \int_1^{A_N} \int_1^{A_N} \delta_N^2 (r_1^4 + r_2^4) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} \max_{|\kappa|=4} \|\partial_\kappa \phi\|_p \\
&+ C \left(\int_1^{A_N} \int_{A_N}^{\infty} (r_1^2 + r_2^2) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} + \int_{A_N}^{\infty} \int_1^{\infty} (r_1^2 + r_2^2) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} \right) \max_{|\kappa|=2} \|\partial_\kappa \phi\|_p.
\end{aligned}$$

The bracketed term is equal to

$$\frac{1 - A_N^{-\mathrm{a}_2}}{\mathrm{a}_2(\mathrm{a}_1 - 2) A_N^{\mathrm{a}_1 - 2}} + \frac{1 - A_N^{-(\mathrm{a}_2 - 2)}}{\mathrm{a}_1(\mathrm{a}_2 - 2) A_N^{\mathrm{a}_1}} + \frac{1}{(\mathrm{a}_1 - 2)\, \mathrm{a}_2\, A_N^{\mathrm{a}_2}} + \frac{1}{\mathrm{a}_1(\mathrm{a}_2 - 2) A_N^{\mathrm{a}_2 - 2}} \leq C A_N^{-\theta},$$

for some constants $C > 0$ and $\theta > 0$. On the other hand, if $\mathrm{a}_1 \neq 4$,

$$\int_1^{A_N} \int_1^{A_N} \delta_N^2 r_1^4 \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} = \delta_N^2 \frac{1 - A_N^{-\mathrm{a}_2}}{\mathrm{a}_2} \times \frac{A_N^{4-\mathrm{a}_1} - 1}{4 - \mathrm{a}_1},$$

while if $\mathrm{a}_1 = 4$,

$$\int_1^{A_N} \int_1^{A_N} \delta_N^2 r_1^4 \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} = \delta_N^2 \frac{1 - A_N^{-\mathrm{a}_2}}{\mathrm{a}_2} \log(A_N).$$

The same integral with r_2^4 instead of r_1^4 is computed in the same way. As a result, if both $\mathrm{a}_1 > 4$ and $\mathrm{a}_2 > 4$, we can take $A_N = \infty$, otherwise we chose $A_N = \delta_N^{-\xi}$ where ξ is such that

$$0 < \xi < \frac{2}{(4 - \mathrm{a}_1)^+} \wedge \frac{2}{(4 - \mathrm{a}_2)^+}.$$

This yields

$$\left\| \mathcal{L}^N \phi - \frac{\sigma^2}{2} \Delta \phi \right\|_p \leq C \delta_N^\theta \max_{0 \leq |\kappa| \leq 4} \|\partial_\kappa \phi\|_p,$$

for some $C > 0$ and $\theta > 0$.

The stable case Let us now treat the case where $\mathrm{a}_1 \wedge \mathrm{a}_2 < 2$. From (3.45), it is clear that a_1 and a_2 play symmetric roles, so let us assume that $\mathrm{a}_1 < \mathrm{a}_2$. In this case, $\alpha = \mathrm{a}_1$ and we change variables in the integral over r_1 to obtain

$$\mathcal{L}^N\phi(x) = u_0 V_1 \int_{\delta_N}^{\infty}\int_1^{\infty} \left(\overline{\overline{\phi}}(x, r_1, \delta_N r_2) - \phi(x)\right) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}}.$$

Then, setting

$$\mathcal{D}^N\phi(x) = \frac{u_0 V_1}{\mathrm{a}_2} \int_{\delta_N}^{\infty} \left(\overline{\phi}(x, r) - \phi(x)\right) \frac{dr}{r^{1+\mathrm{a}_1}},$$

and using (3.40), we write

$$\begin{aligned}
\|\mathcal{L}^N\phi - \mathcal{D}^N\phi\|_p &\le u_0 V_1 \int_{\delta_N}^{\infty}\int_1^{\infty} \|\overline{\overline{\phi}}(\cdot, r_1, \delta_N r_2) - \overline{\phi}(\cdot, r_1)\|_p \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} \\
&\le C \int_{\delta_N}^{\infty}\int_1^{\infty} (1 \wedge (\delta_N r_2)^2) \frac{dr_1 dr_2}{r_1^{1+\mathrm{a}_1} r_2^{1+\mathrm{a}_2}} \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p \\
&\le C\delta_N^{-\mathrm{a}_1} \left(\delta_N^2 \int_1^{1/\delta_N} r^{1-\mathrm{a}_2} dr + \int_{1/\delta_N}^{\infty} \frac{dr}{r^{1+\mathrm{a}_2}}\right) \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p.
\end{aligned}$$

If $\mathrm{a}_2 \neq 2$, this yields

$$\begin{aligned}
\|\mathcal{L}^N\phi - \mathcal{D}^N\phi\|_p &\le C\delta_N^{-\mathrm{a}_1} \left(\delta_N^2 \frac{\delta_N^{-(2-\mathrm{a}_2)} - 1}{2 - \mathrm{a}_2} + \delta_N^{\mathrm{a}_2}\right) \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p \\
&\le C\delta_N^{-\mathrm{a}_1} \left(\frac{\delta_N^{\mathrm{a}_2} - \delta_N^2}{2 - \mathrm{a}_2} + \delta_N^{\mathrm{a}_2}\right) \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p.
\end{aligned}$$

The right hand side above is of the form $C\delta_N^\theta \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p$ for $C > 0$ and $\theta > 0$ since $\mathrm{a}_1 < \mathrm{a}_2$ and $\mathrm{a}_1 < 2$. If $\mathrm{a}_2 = 2$, the above yields instead

$$\|\mathcal{L}^N\phi - \mathcal{D}^N\phi\|_p \le C\delta_N^{-\mathrm{a}_1} \left(\delta_N^2 |\log(\delta_N)| + \delta_N^{\mathrm{a}_2}\right) \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p,$$

which can be bounded by $C\delta_N^\theta \max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi\|_p$ for some $C > 0$ and $\theta > 0$ for the same reasons. □

3.5.3 Convergence of the martingales measures

Martingale Jumps

We now prove that the martingale measures become continuous in the limit.

Proof of Lemma 3.4.9. We start as in [For22, p.26f] and then adapt bounds to implement our regimes. Note

$$\begin{aligned} M_t^N(\phi) - M_{t^-}^N(\phi) &= \left\langle Z_t^N - Z_{t-}^N, \phi \right\rangle - \int_{t-}^t \left\langle Z_s^N, \mathcal{L}^N(\phi) - \mu\phi \right\rangle ds \\ &= \sqrt{N\eta_N} \left\langle \boldsymbol{\rho}_t^N - \boldsymbol{\rho}_{t-}^N, \phi \right\rangle \\ &= \sqrt{N\eta_N} \Big(\left\langle \rho^N_{Nt/\delta_N^\alpha}, \phi_N \right\rangle - \left\langle \rho^N_{(Nt/\delta_N^\alpha)^-}, \phi_N \right\rangle \Big). \end{aligned}$$

If $(t, x_0, u, r_1, r_2) \in \Pi$,

$$\begin{aligned} &\left| \left\langle \rho_t^N, \phi_N \right\rangle - \left\langle \rho_{t^-}^N, \phi_N \right\rangle \right| \\ &\leq u \sup_{k_0 \in [0,1]} \left| \int_{\mathbb{R}^{\mathrm{d}} \times [0,1]} \phi_N(x,k)\, \mathbb{1}_{|x-x_0|<r_2} \left(\delta_{k_0}(dk) - \rho_{t^-}^N(x, dk)\right) dx \right| \\ &\leq 2u \int_{\mathbb{R}^{\mathrm{d}}} \delta_N^{\mathrm{d}} \sup_{k \in [0,1]} |\phi(\delta_N x, k)| \mathbb{1}_{|x-x_0|<r_2} dx. \end{aligned}$$

Using $\mathbb{1}_{|x-x_0|<r_2} \leq 1$ on the one hand and $|\phi(\delta_N x, k)| \leq \|\phi\|_\infty$ on the other we get

$$\int_{\mathbb{R}^{\mathrm{d}}} \delta_N^{\mathrm{d}} \sup_{k \in [0,1]} |\phi(\delta_N x, k)| \mathbb{1}_{|x-x_0|\leq r_2} dx \leq \min\Big\{ \|\phi\|_1, V_{r_2} \delta_N^d \|\phi\|_\infty \Big\}.$$

As a result, in the one-tail regime, there exists a constant $C > 0$, depending only on ϕ and d, such that

$$\begin{aligned} \sup_{t \geq 0} \left| M_t^N(\phi) - M_{t^-}^N(\phi) \right| &\leq C \sqrt{\frac{\eta_N}{N}} \sup_{r_2 > 1} \left(\frac{1 \wedge (\delta_N r_2)^{\mathrm{d}}}{r_2^{\mathrm{c}}} \right) \\ &\leq C \sqrt{\frac{\eta_N}{N}} (\delta_N)^{\mathrm{d} \wedge \mathrm{c}}. \end{aligned}$$

Here, we used the fact that, if $(t, x_0, u, r_1, r_2) \in \Pi$, then $u = u_0/(r_2^{\mathrm{c}} N)$. In the two-tails regime, there exists a constant $C > 0$ such that

$$\begin{aligned} \sup_{t \geq 0} \left| M_t^N(\phi) - M_{t^-}^N(\phi) \right| &\leq C \sqrt{\frac{\eta_N}{N}} \sup_{r_1 > 1, r_2 > 1} \left(\frac{1 \wedge (\delta_N r_2)^{d}}{r_1^{\mathrm{c}_1} r_2^{\mathrm{c}_2}} \right) \\ &\leq C \sqrt{\frac{\eta_N}{N}} (\delta_N)^{d \wedge \mathrm{c}_2}. \end{aligned}$$

In both regimes, the right-hand side converges to zero as $N \to \infty$ by (3.20) and (3.21), which concludes the proof of the lemma. □

Spatial Correlations

In this section, we prove Lemma 3.4.10 to characterize the correlations of the limiting noises in (3.22). We first give a motivating calculation. We have already seen that

$$\left\langle M^N(\phi) \right\rangle_t = \eta_N \int_0^t \left\langle \Gamma^{\nu_\alpha^N}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle ds,$$

where Γ^ν was defined in (3.25), and claim

$$\eta_N \int_0^t \left\langle \Gamma^{\nu_\alpha^N}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle ds \xrightarrow[N\to\infty]{} t \left\langle \mathcal{Q}, \phi \otimes \phi \right\rangle,$$

where $\mathcal{Q}$ is defined as

$$\mathcal{Q}(dx_1 dk_1 dx_2 dk_2) = \gamma K_\beta(dx_1, dx_2)(dk_1 \delta_{k_1}(dk_2) - dk_1 dk_2). \tag{3.46}$$

The measure $K_\beta(dx_1, dx_2)$ encodes the spatial correlations and coalescence rates in each regime and was defined in (3.19). The form of (3.46) follows from $\Gamma^{\nu_\alpha^N}$ by letting $\boldsymbol{\rho}_t^N \to \lambda$. Recall that

$$\begin{aligned}
&\eta_N \Gamma^{\nu_\alpha^N}(\boldsymbol{\rho}_s^N)(x_1, x_2, dk_1, dk_2) \\
&\quad = \eta_N \int_{(0,1]\times[0,\infty)^2} (Nu)^2 \int_{\mathbb{R}^{\mathrm{d}}} \mathbb{1}_{B(x_1,\delta_N r_2)\cap B(x_2,\delta_N r_2)}(x) \frac{1}{V_{\delta_N r_1}} \int_{B(x,\delta_N r_1)} \\
&\qquad \Big[\boldsymbol{\rho}_s^N(y, dk_1)\delta_{k_1}(dk_2) - \boldsymbol{\rho}_s^N(y, dk_1)\boldsymbol{\rho}_s^N(x_2, dk_2) \\
&\qquad - \boldsymbol{\rho}_s^N(x_1, dk_1)\boldsymbol{\rho}_s^N(y, dk_2) + \boldsymbol{\rho}_s^N(x_1, dk_1)\boldsymbol{\rho}_s^N(x_2, dk_2) \Big] \\
&\qquad\qquad dy dx \frac{1}{\delta_N^{\mathrm{d}+\alpha}} \nu_N(du, dr_1, dr_2).
\end{aligned} \tag{3.47}$$

If we replace ρ_s^N by λ, the right-hand side of (3.47) should converge to

$$\begin{aligned}
&\eta_N \int_{(0,1]\times[0,\infty)^2} (Nu)^2 \int_{\mathbb{R}^{\mathrm{d}}} \mathbb{1}_{B(x_1,\delta_N r_2)\cap B(x_2,\delta_N r_2)}(x) \frac{1}{V_{\delta_N r_1}} \int_{B(x,\delta_N r_1)} \\
&\qquad \Big[dk_1 \delta_{k_1}(dk_2) - dk_1 dk_2 \Big] dy dx \frac{1}{\delta_N^{\mathrm{d}+\alpha}} \nu_N(du, dr_1, dr_2) \\
&= \underbrace{\eta_N \int_{(0,1]\times(0,\infty)^2} (Nu)^2 V_{\delta_N r_2}(x_1, x_2) \frac{1}{\delta_N^{\mathrm{d}+\alpha}} \nu_N(du, dr_1, dr_2)}_{} \Big[dk_1 \delta_{k_1}(dk_2) - dk_1 dk_2 \Big].
\end{aligned}$$

The term in brackets contains all spatial information and will equal $\gamma K_\beta(dx_1, dx_2)$. For example, in the one-tail regime, if $\beta = \mathrm{a} + \mathrm{c} < \mathrm{d}$, we obtain

$$\begin{aligned}
&\eta_N \int_{(0,1]\times(0,\infty)^2} (Nu)^2 V_{\delta_N r_2}(x_1, x_2) \delta_N^{-(\mathrm{d}+\alpha)} \nu_N(du, dr_1, dr_2) \\
&\qquad = \delta_N^{\alpha - \mathrm{a} - \mathrm{c}} u_0^2 \int_1^\infty \frac{1}{r_2^{2\,\mathrm{c}}} V_{\delta_N r_2}(x_1, x_2) \delta_N^{-(d+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}-\mathrm{c}}} \\
&\qquad = u_0^2 \int_1^\infty V_{\delta_N r_2}(x_1, x_2) \delta_N \frac{dr_2}{(\delta_N r_2)^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\
&\qquad = u_0^2 \int_{\delta_N}^\infty V_{r_2}(x_1, x_2) \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}}.
\end{aligned}$$

Taking the limit $\delta_N \to 0$ and abbreviating $\frac{x_1 - x_2}{\|x_1 - x_2\|} = e_1$, this converges to

$$\begin{aligned}
&u_0^2 \int_0^\infty V_{r_2}(x_1, x_2) \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\
&= u_0^2 \int_{\mathbb{R}^d} \int_0^\infty \mathbb{1}_{\{|x_1 - x_2 - y| < r_2\}} \mathbb{1}_{\{|0-y|<r_2\}} \frac{dr_2}{r_2^{1+\mathrm{d}+\beta}} dy \\
&= u_0^2 \int_{\mathbb{R}^d} \int_0^\infty \mathbb{1}_{\{|e_1 - y| < r_2/\|x_2 - x_1\|\}} \mathbb{1}_{\{|y|<r_2/\|x_2-x_1\|\}} \frac{dr_2}{r_2^{1+\mathrm{d}+\beta}} \|x_1 - x_2\|^{\mathrm{d}} dy \\
&= u_0^2 \int_{\mathbb{R}^d} \int_0^\infty \mathbb{1}_{\{|e_1 - y| < r_2\}} \mathbb{1}_{\{|y|<r_2\}} \frac{1}{\|x_1 - x_2\|^\beta} \frac{dr_2}{r_2^{1+\mathrm{d}+\beta}} dy \\
&= u_0^2 C_{d,\beta}^{(2)} K_\beta(x_1, x_2).
\end{aligned}$$

If $\beta > \mathrm{d}$, we instead arrive at

$$\begin{aligned}
&\eta_N \int_{(0,1]\times(0,\infty)^2} u^2 V_{r_2}(x_1, x_2) \nu_N^\alpha(dr_1, dr_2) \\
&\qquad = u_0^2 \eta_N \int_{(0,1]\times(0,\infty)^2} (Nu)^2 V_{\delta_N r_2}(x_1, x_2) \delta_N^{-(\mathrm{d}+\alpha)} \nu_N(du, dr_1, dr_2) \\
&\qquad = u_0^2 \delta_N^{\alpha - \mathrm{d}} \int_1^\infty V_{\delta_N r_2}(x_1, x_2) \delta_N^{-(\mathrm{d}+\alpha)} \frac{dr_2}{r_2^{1+\mathrm{d}+\alpha+\mathrm{c}}} \\
&\qquad = u_0^2 V_1^2 \int_1^\infty \frac{V_{\delta_N r_2}(x_1, x_2)}{V_{\delta_N r_2}^2} \frac{dr_2}{r_2^{1+\alpha+\mathrm{c}-\mathrm{d}}},
\end{aligned}$$

and as this quantity is integrated against test functions

$$\begin{aligned}
&u_0^2 V_1^2 \int_1^\infty \int_{(\mathbb{R}^{\mathrm{d}})^2} \phi(x_1, k_1) \phi(x_2, k_2) \frac{V_{\delta_N r_2}(x_1, x_2)}{V_{\delta_N r_2}^2} dx_1 dx_2 \frac{dr_2}{r_2^{1+\alpha+\mathrm{c}-\mathrm{d}}} \\
&\qquad \xrightarrow[N\to\infty]{} u_0^2 V_1^2 \int_1^\infty \frac{dr_2}{r_2^{1+\alpha+\mathrm{c}-\mathrm{d}}} \int_{\mathbb{R}^{\mathrm{d}}} \phi(x, k_1) \phi(x, k_2) dx.
\end{aligned}$$

Noting $\int_1^\infty r_2^{-(1+\alpha+\mathrm{c}-\mathrm{d})} dr_2 = (\alpha + \mathrm{c} - \mathrm{d})^{-1}$ the above expression matches the corresponding result in Table 3.1. The spatial correlations vanish and the noise

becomes white in space. A similar calculation holds in the case $\beta = \mathrm{d}$. The next proof will demonstrate that all these calculations are rigorously justified.

Proof of Lemma 3.4.10. In this proof, we will focus on the one-tail regime as the two-tail regime can be treated with similar arguments. First, we will consider the case $\beta = \mathrm{a} + \mathrm{c} < \mathrm{d}$ extending [For22, Subsection 3.5, Proof of Lemma 3.10 (3.32), stable case]. Second, we will adapt [For22, Subsection 3.5, Proof of Lemma 3.10 (3.32), fixed case] to analyse the case $\beta > \mathrm{d}$, where the spatial correlations vanish. Afterwards, we will treat the case $\beta = \mathrm{d}$. The covariation process can be written as

$$\left\langle M^N(\phi) \right\rangle_t \tag{3.48}$$

$$= \eta_N \int_0^t \int_{(0,1]\times[0,\infty)^2} (Nu)^2 \left\langle \Gamma^{(\delta_N r_1, \delta_N r_2)}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle \delta_N^{-(\alpha+d)} \nu_N(du, dr_1, dr_2) ds, \tag{3.49}$$

where $\Gamma^{(\delta_N r_1, \delta_N r_2)}$ is short for evaluation of Γ^ν at the Dirac measure at $(1, \delta_N r_1, \delta_N r_2)$.

Case $\boldsymbol{\beta < \mathrm{d}}$ Substituting integrals, (3.48) translates to:

$$\left\langle M^N(\phi) \right\rangle_t = \int_0^t u_0^2 \int_{\delta_N}^\infty \left\langle \Gamma^{(\delta_N^{1-\mathrm{b}} r_2^{\mathrm{b}}, r_2)}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds. \tag{3.50}$$

In (3.50) we can change the limit of integration of the inner integral from $[\delta_N, \infty)$ to $[0, \infty)$, as the difference converges to zero:

$$\int_0^t u_0^2 \int_0^{\delta_N} \left\langle \Gamma^{(\delta_N^{1-\mathrm{b}} r_2^{\mathrm{b}}, r_2)}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds$$
$$\leq 4u_0 t \|\phi\|_2^2 \int_0^{\delta_N} V_{r_2}^2 \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \xrightarrow[N\to\infty]{} 0,$$

where we used again (3.38). Defining

$$\mathfrak{Q}_{r_2}(x_1, x_2, dk_1, dk_2) := V_{r_2}(x_1, x_2)(dk_1 \delta_{k_1}(dk_2) - dk_1 dk_2)$$

it remains to show

$$\int_0^t u_0^2 \int_0^\infty \left\langle \Gamma^{(\delta_N^{1-\mathrm{b}} r_2^{\mathrm{b}}, r_2)}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds$$
$$\xrightarrow[N\to\infty]{} \int_0^t u_0^2 \int_0^\infty \left\langle \mathfrak{Q}_{r_2}, \phi \otimes \phi \right\rangle \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds.$$

We proceed by dominated convergence. Calculations in Lemma 3.4.5 provide us with the necessary bound in (r_1, r_2). To see pointwise convergence of the integrand, note we can write

$$\left\langle \Gamma^{(r_1,r_2)}(\boldsymbol{\rho}), \phi \otimes \phi \right\rangle = V_{r_2}^2 \left\langle \boldsymbol{\rho}, \Psi^{(1)}_{(r_1,r_2)} \right\rangle + \left\langle \boldsymbol{\rho} \otimes \boldsymbol{\rho}, \Psi^{(2)}_{(r_1,r_2)} \right\rangle$$

for two functions $\Psi^{(1)}_{(r_1,r_2)}(x_1, k_1), \Psi^{(2)}_{(r_1,r_2)}(x_1, x_2, k_1, k_2)$. The first function is defined as

$$\Psi^{(1)}_{(r_1,r_2)}(x_1, k_1) := \frac{1}{V_{r_1}} \int_{B(x_1,r_1)} \overline{\phi}(x, k_1, r_2)^2 dx$$

and the second function as

$$\begin{aligned}
&\Psi^{(2)}_{(r_1,r_2)}(x_1, x_2, k_1, k_2) \\
&:= V_{r_2}(x_1, x_2)\phi(x_1, k_1)\phi(x_2, k_2) - \frac{V_{r_2}}{V_{r_1}}\phi(x_1, k_1) \int_{B(x_1,r_1)\cap B(x_2,r_2)} \overline{\phi}(y, k_2, r_2) dy \\
&\quad - \frac{V_{r_2}}{V_{r_1}}\phi(x_2, k_2) \int_{B(x_1,r_1)\cap B(x_2,r_2)} \overline{\phi}(y, k_1, r_2) dy.
\end{aligned}$$

Since $\Psi^{(1)} \in \mathcal{S}(\mathbb{R} \times [0, 1])$, we know from (3.34) that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| \left\langle Z_t^N, \Psi^{(1)} \right\rangle \right|^2 \right]^{1/2} \leq C \max_{q\in\{1,2\}} \max_{|\kappa|\leq 2} \|\partial_\kappa \Psi^{(1)}\|_p. \tag{3.51}$$

The derivatives on the right-hand side are

$$\frac{\partial}{\partial x_1} V_{r_2}^2 \Psi^{(1)}_{(r_1,r_2)}(x_1, k_1) = \frac{2V_{r_2}^2}{V_{r_1}} \int_{B(x_1,r_1)} \overline{\phi}(y, k_1, r_2)\overline{\phi_x}(y, k_1, r_2) dy$$

and

$$\frac{\partial^2}{\partial x_1^2} V_{r_2}^2 \Psi^{(1)}_{(r_1,r_2)}(x_1, k_1) = \frac{2V_{r_2}^2}{V_{r_1}} \int_{B(x_1,r_1)} \overline{\phi_x}(y, k_1, r_2)^2 + \overline{\phi}(y, k_1, r_2)\overline{\phi_{xx}}(y, k_1, r_2) dy.$$

The Schwartz space $\mathcal{S}(\mathbb{R}^d \times [0, 1])$ is closed under multiplication and averages. This, together with Lemma 3.5.2, allows us to bound (3.51) for $\Psi^{(1)}$ independently of r_1 and show pointwise convergence for the first summand. To see the pointwise convergence of the second summand note

$$\mathbb{E}\Big[\left\langle Z_t^N \otimes Z_t^N, \Psi^{(2)} \right\rangle \Big] \leq \mathbb{E}\Big[\left\langle D^N, \mathbb{1}_{[0,t]} \Psi^{(2)} \right\rangle \Big].$$

Corollary 3.5.1 provides us with a uniform bound on the right-hand side. This implies the pointwise convergence of $\left\langle (\boldsymbol{\rho}_t^N - \lambda) \otimes (\boldsymbol{\rho}_t^N - \lambda), \Psi^{(2)} \right\rangle$. The identity

$$\boldsymbol{\rho}_t^N \otimes \boldsymbol{\rho}_t^N - \lambda \otimes \lambda = \lambda \otimes (\boldsymbol{\rho}_t^N - \lambda) + (\boldsymbol{\rho}_t^N - \lambda) \otimes \lambda + (\boldsymbol{\rho}_t^N - \lambda) \otimes (\boldsymbol{\rho}_t^N - \lambda)$$

allows us to make the transition to

$$\left\langle \boldsymbol{\rho}_t^N \otimes \boldsymbol{\rho}_t^N - \lambda \otimes \lambda, \Psi^{(2)} \right\rangle = (N\eta_N)^{-1/2} \left\langle Z_t^N, \tilde{\Psi} \right\rangle + (N\eta_N)^{-1} \left\langle Z_t^N \otimes Z_t^N, \Psi^{(2)} \right\rangle. \tag{3.52}$$

Here, $\tilde{\Psi}(x,k) = \left\langle \lambda, \Psi^{(2)}(\cdot, x, k) \right\rangle + \left\langle \lambda, \Psi^{(2)}(x, \cdot, k) \right\rangle$ and it remains to prove the pointwise convergence of the middle term. We can represent $\left\langle \lambda, \Psi^{(2)}(\cdot, x, k) \right\rangle$ through averaged functions

$$\begin{aligned}
&\left\langle \lambda, \Psi^{(2)}(\cdot, x_2, k_2) \right\rangle \\
&= \int_{\mathbb{R}^d} V_{r_2}(x_1, x_2) \int_0^1 \phi(x_1, k_1) dk_1 dx_1 \phi(x_2, k_2) \\
&\quad - \frac{V_{r_2}}{V_{r_1}} \int_{\mathbb{R}^d} \int_0^1 \phi(x_1, k_1) dk_1 \int_{B(x_1, r_1)} \mathbb{1}_{B(x_2, r_2)} \overline{\phi}(y, k_2, r_2) dy dx_1 \\
&\quad - \frac{V_{r_2}}{V_{r_1}} \phi(x_2, k_2) \int_{\mathbb{R}^d} \int_{B(x_1, r_1)} \mathbb{1}_{B(x_2, r_2)}(y) \int_0^1 \overline{\phi}(y, k_1, r_2) dk_1 dy dx_1 \\
&= \overline{\overline{\phi}}(x_2, r_2) \phi(x_2, k_2) V_{r_2}^2 - V_{r_2}^2 \overline{\overline{\phi}(\cdot, r_1) \overline{\phi}(\cdot, k_2, r_2)}(x_2, r_2) - V_{r_2}^2 \overline{\overline{\phi}}(x_2, r_2) \phi(x_2, k_2).
\end{aligned}$$

Again, this time for $\left\langle \lambda, \Psi^{(2)}(\cdot, x, k) \right\rangle$, we can bound the right-hand side of the inequality (3.51) independently of r_1 and the missing term of (3.52) converges.

Case $\boldsymbol{\beta > d}$ We use a separate strategy, which extends [For22] to different radii r_1, r_2 for the impacted area and range from where to choose parents. Our strategy is to approximate the four summands of $\left\langle \Gamma^{(r_1, r_2)}(\rho), \phi \otimes \phi \right\rangle$ by more tractable terms. The first summand equals

$$V_{r_2}^2 \left\langle \overline{\rho}_{r_1}, \overline{\phi}(\cdot, r_2)^2 \right\rangle,$$

where we define for $r > 0$

$$\overline{\rho}_r(x, dk) := \frac{1}{V_r} \int_{B(x,r)} \rho(z, dk) dz.$$

The second term is of the form

$$\int_{\mathbb{R}^d} \int_{[0,1]} \int_{B(x,r_2)} \phi(x_1, k_1) dx_1 \overline{\rho}_{r_1}(x, dk_1) \int_{B(x,r_2)} \int_{[0,1]} \phi(x_2, k_2) \rho(x_2, dk_2) dx_2 dx. \tag{3.53}$$

If we replace $\phi(x_1, k_1)$ by $\phi(x, k_1)$ and $\phi(x_2, k_2)$ by $\phi(x, k_2)$ in (3.53), the distance between the new expression

$$V_{r_2}^2 \left\langle \overline{\rho}_{r_1} \cdot \overline{\rho}_{r_2}, \phi \otimes \phi \right\rangle = V_{r_2}^2 \int_{\mathbb{R}^{\mathrm{d}} \times [0,1]^2} \phi(x, k_1)\phi(x, k_2)\overline{\rho}_{r_1}(x, dk_1)\overline{\rho}_{r_2}(x, dk_2)dx$$

and (3.53) is smaller than

$$\min\left\{2r_2 V_{r_2}^2 \|\phi\|_1 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty, 4V_{r_2}^2 \|\phi\|_1 \|\phi\|_\infty\right\}.$$

This can be seen by applying

$$\left| \int_{B(x,r)\times[0,1]} (\phi(x_1, k_1) - \phi(x, k_1))\rho(x_1, dk_1)dx_1 \right| \leq \min\left\{rV_r \max_{|\kappa|=1} \|\partial_\kappa \phi\|_1, 2V_r \|\phi\|_\infty\right\}.$$

The last two terms of $\left\langle \Gamma^{(r_1,r_2)}, \phi \otimes \phi \right\rangle$ can be approximated in the same way to obtain

$$\begin{aligned}\Big| \left\langle \Gamma^{(r_1,r_2)}(\rho), \phi \otimes \phi \right\rangle - V_{r_2}^2 \Big(\left\langle \overline{\rho}_{r_1}, \overline{\phi}(\cdot, r_2)^2 \right\rangle - \left\langle \overline{\rho}_{r_1} \cdot \overline{\rho}_{r_2}, \phi \otimes \phi \right\rangle \Big) \Big| \\ \leq \min\left\{8r_2 V_{r_2}^2 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty \|\phi\|_1, 16V_{r_2}^2 \|\phi\|_\infty \|\phi\|_1\right\}.\end{aligned}$$

In total, if $\mathrm{a} + \mathrm{c} > \mathrm{d}$ this leads to

$$\begin{aligned}&\Big| \langle M^N(\phi) \rangle_t - u_0^2 \delta_N^{-2d} \int_0^t \int_1^\infty V_{\delta_N r_2}^2 \Big(\left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}}, \overline{\phi}(\cdot, \delta_N r_2)^2 \right\rangle \\ &\qquad - \left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}} \cdot \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2}, \phi \otimes \phi \right\rangle \Big) \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds \Big| \\ &\leq u_0^2 t \delta_N^{-2\,\mathrm{d}} 8 \|\phi\|_1 \int_1^\infty \min\left\{\delta_N r_2 V_{\delta_N r_2}^2 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty, 2V_{\delta_N r_2}^2 \|\phi\|_\infty\right\} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\ &\leq u_0^2 t \delta_N^{\mathrm{a}+\mathrm{c}-\mathrm{d}} 8 \|\phi\|_1 \int_{\delta_N}^\infty \min\left\{r_2 V_{r_2}^2 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty, 2V_{r_2}^2 \|\phi\|_\infty\right\} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\ &\leq u_0^2 t \delta_N^{\mathrm{a}+\mathrm{c}-\mathrm{d}} 8 \|\phi\|_1 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty \int_{\delta_N}^1 r_2 V_{r_2}^2 \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\ &\qquad + u_0^2 t \delta_N^{\mathrm{a}+\mathrm{c}-\mathrm{d}} 16 \|\phi\|_1 \|\phi\|_\infty \int_1^\infty V_{r_2}^2 \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}}.\end{aligned}$$

This right-hand side converges to zero as $N \to \infty$. Again from Theorem 3.3.2 we know $\boldsymbol{\rho}_t^N \to \lambda$ in probability, which implies $\overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}} \to \lambda$ and $\overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2} \to \lambda$ in the same sense. This shows

$$\begin{aligned}\left\langle M_t^N(\phi) \right\rangle_t &\xrightarrow[N\to\infty]{} u_0^2 V_1^2 \int_0^t \int_1^\infty \left(\left\langle \lambda, \phi^2 \right\rangle - \left\langle \lambda \cdot \lambda, \phi \otimes \phi \right\rangle \right) \frac{dr_1}{r_1^{1+\mathrm{a}+\mathrm{c}-\mathrm{d}}} ds \\ &= \left\langle \mathcal{Q}, \phi \otimes \phi \right\rangle.\end{aligned}$$

Case $\boldsymbol{\beta = \mathrm{d}}$ Similar to the argument following (3.37), we can show the bound

$$\left| \left\langle \Gamma^{(r_1,r_2)}(\rho), \phi \otimes \phi \right\rangle \right| \leq 4V_{r_2}\|\phi\|_1^2.$$

The following distance vanishes as $N \to \infty$:

$$\begin{aligned}
&\Bigg| \langle M^N(\phi) \rangle_t - u_0^2 \frac{\delta_N^{-2d}}{\log(1/\delta_N)} \int_0^t \int_1^{1/\delta_N} V_{\delta_N r_2}^2 \Big(\left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}}, \overline{\phi}(\cdot, \delta_N r_2)^2 \right\rangle \\
&\qquad - \left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}} \cdot \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2}, \phi \otimes \phi \right\rangle \Big) \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds \Bigg| \\
&\leq u_0^2 t \frac{\delta_N^{-2\,\mathrm{d}}}{\log(1/\delta_N)} 8\|\phi\|_1 \int_1^{1/\delta_N} \delta_N r_2 V_{\delta_N r_2}^2 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\
&\qquad + u_0^2 t \frac{\delta_N^{-2\,\mathrm{d}}}{\log(1/\delta_N)} \int_{1/\delta_N}^\infty \left| \left\langle \Gamma^{(\delta_N r_1, \delta_N r_2)}(\boldsymbol{\rho}_s^N), \phi \otimes \phi \right\rangle \right| \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\
&\leq u_0^2 t \frac{\delta_N^{-2\,\mathrm{d}}}{\log(1/\delta_N)} 8\|\phi\|_1 \int_1^{1/\delta_N} \delta_N r_2 V_{\delta_N r_2}^2 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \\
&\qquad + u_0^2 t \frac{\delta_N^{-2\,\mathrm{d}}}{\log(1/\delta_N)} 4\|\phi\|_1^2 \int_{1/\delta_N}^\infty V_{\delta_N r_2} \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} \to 0.
\end{aligned}$$

Substituting $u := \frac{\log(r_2)}{\log(1/\delta_N)}$ we obtain

$$\begin{aligned}
&u_0^2 \frac{\delta_N^{-2d}}{\log(1/\delta_N)} \int_0^t \int_1^{1/\delta_N} V_{\delta_N r_2}^2 \Big(\left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}}, \overline{\phi}(\cdot, \delta_N r_2)^2 \right\rangle \\
&\qquad - \left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}} \cdot \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2}, \phi \otimes \phi \right\rangle \Big) \frac{dr_2}{r_2^{1+\mathrm{d}+\mathrm{a}+\mathrm{c}}} ds \\
&= u_0^2 V_1^2 \int_0^t \int_1^{1/\delta_N} \Big(\left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}}, \overline{\phi}(\cdot, \delta_N r_2)^2 \right\rangle \\
&\qquad - \left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2^{\mathrm{b}}} \cdot \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N r_2}, \phi \otimes \phi \right\rangle \Big) \frac{dr_2}{\log(1/\delta_N) r_2} ds \\
&= u_0^2 V_1^2 \int_0^t \int_0^1 \Big(\left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N^{1-\mathrm{b}\,u}}, \overline{\phi}(\cdot, \delta_N^{1-u})^2 \right\rangle \\
&\qquad - \left\langle \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N^{1-\mathrm{b}\,u}} \cdot \overline{(\boldsymbol{\rho}_s^N)}_{\delta_N^{1-u}}, \phi \otimes \phi \right\rangle \Big) du ds.
\end{aligned}$$

Following a similar argument based on dominated convergence as in the case $\beta > \mathrm{d}$, we can conclude the proof. □

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

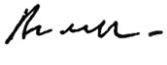
Title of Paper	Central limit theorems describing isolation by distance under various forms of power-law dispersal
Publication Status	□Published □ Accepted for Publication ■Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	arxiv.org/abs/2211.16286

Student Confirmation

Student Name:	Bastian Wiederhold		
Contribution to the Paper	I wrote the majority of the document, whereas Raphaël Forien provided crucial ideas regarding the precise scaling regimes as well as more complex calculations and possible simplifications.		
Signature 		Date	11th July 2024

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Alison Etheridge		
Supervisor comments I am confident that Bastian Wiederhold made a very substantial contribution to this publication.		
Signature 	Date	11th July 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Joint work with Raphaël Forien

4

Isolation by distance under varying population size

Contents

4.1 Introduction 110
4.2 Definition of the model 114
4.3 Main results 122
4.4 Convergence of the population size process 133
4.5 Semimartingale decomposition of the fluctuations . . . 139
4.6 Proof of the central limit theorem 151
4.7 Proof of the Wright-Malécot formula 162

4.1 Introduction

Isolation by distance The genetic patterns of a scattered population are spatially correlated if individuals move across space and successfully reproduce at positions different to their origin. In many species, individual movements during their lifetime will not exhaust the complete range of the species due to environmental, behavioural or energetic constraints. New genetic information created by mutation and recombination is therefore diffusing over space at a finite rate. The genetic similarity of individuals separated in space will decrease as a function of their separation. This phenomenon is called isolation by distance and the patterns of genetic isolation will reflect the spatial gene flow in the population.

Wright-Malécot formula To quantify patterns of genetic isolation we need to find an appropriate scale of genetic similarity. The simplest measure for the genetic relatedness of two individuals is to ask whether they are identical or not. While in reality the genetic information of an individual from a non-clonal species rarely matches any other individual, considering only a part of the genetic information such as a single locus renders the measure applicable. We could then measure isolation by distance by the probability of identity.

Wright and Malécot ([Wri43], [Mal48]) observed that the decay of the probability of identity with spatial separation is roughly exponential. This approximation holds for many models and is today named after them. The universality of the approximation reflects the fact that ancestral lineages, the sequences of spatial positions of the ancestors, are as jump processes well-approximated by Brownian motions over sufficiently large scales. To see the same type in two sampled individuals, tracing backwards in time ancestral lineages need to have coalesce before either acquiring mutations. In turn, the Brownian motions approximating their motion are described by Gaussian transition kernels leading to the observed exponential decay in relatedness. Later works noticed that some of their assumptions were inconsistent ([Fel75]) and reproduced the results more convincingly in space ([KW64], [BDE02]).

Natural forces An exact exponential decay in real populations is unlikely, due to influences such as the local structure, rare events or inhomogenous dispersal dynamics. Ideally, we would like to be able to describe isolation by distance under the multifarious forces acting on natural populations to provide a tool to infer information about the population from its genetic patterns. Simulation-based approaches have incorporated different influences on isolation by distance with recent examples found in ([Rin+18], [Sga+22]).

A first step towards an analytical description of new influences was undertaken with the introduction of the class of spatial Λ-Fleming-Viot processes ([BEV10], [BEV13a]). In this framework, it was shown how long-range reproduction events

might alter the genealogy. Several ancestral lineages of sampled individuals can merge at the same time, as the contained ancestors were offspring to the same parent of a long-range event. This is accompanied by a slower than exponential decay of the probability of identity. Corresponding long-range Wright-Malécot formulae were discovered in [SW23], [For22] and Chapter 3. Here, ancestral lineages behave according to stable processes, which can be thought of as a generalization of Brownian motion with larger jumps. Moreover, even if ancestral lineages are still separated in space, there exists a positive rate of coalescence due to large spatial events.

Aim and motivation of this work In this paper, we derive a Wright-Malécot formula to quantify isolation by distance under the influence of varying population size. To this end, we propose a spatial Λ-Fleming-Viot type model with a reproduction mechanism which might change the local population density. In reproduction events, a proportion of the population in a region dies and is replaced by a not necessarily equal mass governed by a so-called growth function that depends on the local population size. If the carrying capacity is larger, more individuals are expected to have reproductive success. We model this by letting the frequency of events increase with higher population size while reducing the relative fraction replaced by offspring in the population. Our motivation is twofold:

1. The resulting formula is more realistic than previous results, as it explains for example the increased genetic isolation if the population is separated through barriers to gene flow such as by unfavourable terrain (see Figure 4.2). This is an important step towards bridging the gap between the aforementioned analytical work and more applied papers incorporating precisely these effects such as [Rin+18]. Additionally, although in this paper the bias in the gene flow results from differences in population size, we expect a similar formula to hold if there are general biases in the displacement of individuals. For example, octopus dispersal patterns are more likely to follow the circular drift of the Antarctic circumpolar current (see [Lau+23]).

2. Mathematically, the varying population size makes it more difficult to find duality relationships with link the forwards in time process with a corresponding backwards in time structure. Ancestral lineages might be non-Markovian and correlated in a complicated way, as the paths of ancestral lineages might contain information about the population profile. In such a setting, only few rigorous mathematical results have been obtained so far ([BCD16], [Eth+24], [BDS24]). Notably, in [Eth+24] the motion of ancestral lineages is extracted from a complex spatial model using lookdown constructions. We will arrive at a comparable motion of ancestral lineages from a completely different angle using a strategy based on a central limit theorem ([For22]).

   In previous papers ([For22], Chapter 3), the considered process was of population size one allowing one to derive the probability of identity and associated behaviour of ancestral lineages in a rather straight-forward manner from a corresponding central limit theorem. In the present paper, we will consider a central limit theorem originating from a process which is of varying population size. This renders the derivation of the formula for the probability of identity more complex, but demonstrates the feasibility of the technique in a much wider context.

Forwards in time central limit theorem The main tool of this work will be a connection between the probability of identity and central limit theorems. Individuals will be equipped with neutral genetic types, tracing related individuals in the population. Mutation then corresponds to assuming uniform new types from the real interval $[0, 1]$ with a neutral mutation rate μ. As two uniform samples on the real interval $[0, 1]$ are never equal, in this way, two individuals are only of the same type if neither has experienced mutation since their most recent common ancestor.

The process describing the limiting fluctuations of a spatial model is given as the solution of an SPDE driven by a Gaussian random field. A spatial differential operator such as the Laplacian disperses the 'spikes' of the Gaussian random field across space. The genetic fluctuations of the SPDE in two distinct spatial positions

can only be correlated if individuals from those positions had a common ancestor back in the past. This explains the emergence of the coalescence behaviour of the ancestral lineages of two samples in the central limit theorem. The differential operator carries the fluctuations across space and is therefore linked to the behaviour of ancestral lineages. A coalescence will cause a significant deviation from the average behaviour, and hence the noise and its spatial correlations reflect the coalescence behaviour.

This work illustrates that the correspondence between central limit theorems and the probability of identity and the behaviour ancestral lineages is a general phenomenon, while providing more realistic analytical expressions to approximate isolation by distance.

Outline The next section introduces and motivates our model. In Section 4.3 we state and interpret our main results. In Section 4.4 we deduce the convergence of the population size to the solution of a reaction-diffusion equation. Section 4.5 starts analysing the evolution of genetic types under the model by stating and scaling the corresponding martingale problems. The proof of the central limit theorem is contained in Section 4.6, and in Section 4.7 we deduce the corresponding Wright-Malécot formula. In Appendix B.1 we will provide details on Figure 4.2, which illustrates the accuracy of the predicted probability of identity by comparing to simulations. Further appendices deal with technical details.

4.2 Definition of the model

Class of models Our analysis will use a spatial Λ-Fleming-Viot (SLFV) model, a class of measure-valued stochastic processes first developed in [Eth08] and [BEV10]. Up to null-sets, at every space-time position the process assigns a distribution on a type space, representing the composition of the population at this point. The population is driven by a Poisson point process that determines the time and location of events, the region affected and the 'impact' u of the event. Inside the area affected by an event, first a parental type is sampled from the population, and then a proportion u of the population is replaced by this type.

Varying population size SLFV models were first defined rigorously in [EK19]. There a general offspring mechanism was considered based on the assumption that the 'mass' of offspring is random and has an expectation given by the affected area multiplied by the event impact. This idea is also present in our more specific choice of reproduction events. Additionally, events fall at a rate proportional to the local population size and the replaced proportion of the population decreases if the population size is larger. Another modification of our model compared to the classical SLFV model is a mutation mechanism, which was first studied in [VW15]. Although the population is already an infinite-population limit, it can be thought of as each individual mutating with a certain rate μ to a new type represented by a uniform choice from the real interval $[0, 1]$. To guarantee existence of the model, we will work on the d-dimensional torus $\mathbb{T}^d$.

Choice of events Throughout this work, we will denote the population size by n (defined almost everywhere) and the local average of the population size over a ball $B(x, R)$ with radius $R > 0$ as

$$\overline{n}(x, R) := \frac{1}{V_R} \int_{B(x,R)} n_t(z) dz.$$

Heuristically, during a time period of length one, we wish to replace a certain proportion u of the population reflecting the turnover through death and reproduction. However, if the local population size $\overline{n}$ is larger, we expect more individuals have reproductive success and to be parental to some part of the proportion u of new individuals. There are two different ways of implementing this in the SLFV:

1. Either we have events replacing a proportion $u \in (0, 1)$, but allow a number of parents proportional to the local population size $\overline{n}$,

2. or we could have the rate of events proportional to the local population size $\overline{n}$, each with a single parent and replacing only a proportion $u/\overline{n}$ of the population.

We opted for the second possibility. Events will cover a ball $B(x, R)$ of radius $R > 0$ around the centre x of the event. To be exact, for each spatial position $y \in B(x, R)$ the population size changes according to

$$n(y) \mapsto \Big(1-\frac{u}{\overline{n}(x,R)+1}\Big)n(y)+\frac{u}{\overline{n}(x,R)+1}\Big(1+r_x(\overline{n}(x,R))\Big)\overline{n} \text{ at rate } \overline{n}+1, \quad (4.1)$$

where $r : \mathbb{T}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ is a function capturing the regulation of the population size. The first argument of $r_x(\overline{n})$ allows the population growth to depend on space, whereas the second is the local population size. Typically, the growth function will be positive if local population size is smaller than the carrying capacity of the location and negative if the local population size is larger.

One might note that in (4.1) we accelerate by $\overline{n} + 1$ instead of $\overline{n}$ and replace a proportion of $u/(\overline{n} + 1)$ instead of $u/\overline{n}$. Adding one to the rate of reproduction events captures that there is always reproductive activity even if the local population density is low. The same replacement in the denominator also prevents the population size from becoming negative, as otherwise for small $\overline{n}$ the term $(1 - u/\overline{n})$ could be negative.

To provide a precise definition, we will now first clarify the state space Ξ of the process and then describe its evolution (similar to [For22, Section 1]).

State space We will choose the growth function in such a way that the population size does not pass a threshold $n_{\max} > 0$. The type space will be the real interval $[0, 1]$. At any location, the population can be thought of assuming values in $\mathcal{M}_{n_{\max}}([0, 1])$, the set of all measures with mass smaller or equal to $n_{\max}$ on $[0, 1]$. The full state space $\Xi = (\Xi', d)$ will be

$$\Xi' = \Big\{\rho : \mathbb{T}^d \to \mathcal{M}_{n_{\max}}([0,1]) \text{ is a measurable function}\Big\}\Big|_{\sim}$$

under the equivalence relationship

$$\rho_1 \sim \rho_2 \Leftrightarrow \{z \in \mathbb{T}^d : \rho_1(z) \neq \rho_2(z)\} \text{ is a Lebesgue null set.}$$

We equip Ξ' with the topology of weak convergence. To this end, let $(\phi_n)_{n\geq 1}$ be a sequence of test functions such that

$$\exists k : \langle \rho_1, \phi_k \rangle \neq \langle \rho_2, \phi_k \rangle \Leftrightarrow \rho_1 \neq \rho_2 \text{ in } \Xi'.$$

To simplify some of the later calculations, in the same way as [For22], we introduce p-norms as

$$\|\phi\|_p := \left(\int_{\mathbb{T}^d} \sup_{k\in[0,1]} |\phi(z,k)|^p dz \right)^{1/p}. \tag{4.2}$$

and additionally assume a uniform bound of the partial derivatives for indices $\kappa \in \mathbb{N}^d$

$$\sup_{n\in\mathbb{N}} \left\{ \|\partial_\kappa \phi_n\|_p : |\kappa| \leq 2, p \in \{1, \infty\} \right\} < \infty. \tag{4.3}$$

The weak metric d inducing the topology is given by

$$d(\rho_1, \rho_2) := \sum_{n=1}^{\infty} \frac{1}{2^n} \Big| \langle \rho_1, \phi_n \rangle - \langle \rho_2, \phi_n \rangle \Big|. \tag{4.4}$$

In the same way as for the local population size $\overline{n}$, we will write

$$\overline{\phi}(x,k,R) := \frac{1}{V_R} \int_{B(x,R)} \phi(y,k) dy$$

for the average of a function $\phi : \mathbb{T}^d \times [0,1] \to \mathbb{R}$ over a ball of radius R with volume V_R. Suppose a stochastic process evolves in Ξ. For a space-time position (z,t), the value $\rho_t(z)$ can be interpreted as assigning masses to genetic types of the interval $[0,1]$ (up to a Lebesgue null-set). The total mass

$$n_t(z) = \int_0^1 \rho_t(z, dk) := \mathcal{H}\rho_t(z)$$

encodes the population size. We will use the symbol $\mathcal{H}$ to denote the mass of a measure $M \in \mathcal{M}([0,1])$ or a function $\phi : [0,1] \to \mathbb{R}$ if interpreted as a density, i.e.

$$\mathcal{H}M := \int_0^1 M(dk), \qquad \mathcal{H}\phi := \int_0^1 \phi(k) dk. \tag{4.5}$$

Restrictions on the growth function We wish to make three additional assumption on our model, which we can guarantee by assuming certain properties of the growth function. Recall from (4.1) that we add a mass of size

$$\frac{u}{\overline{n}(x,R)+1}\Big(1+r_x(\overline{n}(x,R))\Big)\overline{n}$$

of the sampled parental type in reproduction events.

1. It is natural to assume that the new-born mass in events is positive. Strictly speaking, the above term $1+r_x(\overline{n})$ requires the growth function to be bounded from below by minus one. As we will scale the growth function by a term converging to zero (4.7), a lower bound will suffice to ensure this property for the scaled processes.

2. We wish to ensure upper and lower bounds on the population size $0 < n_{\min} \leq n_{\max}$. If the growth function $r_x(\overline{n})$ is positive, the population grows locally. To ensure a lower bound it is therefore enough to assume that as the local population size $\overline{n}$ becomes small, the growth function stays larger than a positive value for all locations. An upper bound can be guaranteed by assuming that $(1+r_x(\overline{n}))\overline{n}$ stays bounded.

   These assumptions significantly reduce the technicalities required for the proofs in this paper. For example, if the population size had no upper bound, one would need to establish a control on large values in order to ensure convergence of the sequence of rescaled population size processes.

3. In addition, we will make a regularity assumption on the growth function. This allows us to follow the arguments of [FP17] for the convergence of the population size and ensures that the limit, which will be a solution to a reaction-diffusion equation, exists.

Each of these ideas can be formalized as follows.

Assumption 4.2.1. We assume the following properties of the growth function $r : \mathbb{T}^d \times \mathbb{R}_+ \to \mathbb{R}_+$:

1. (Positive birth) There exists a lower bound $C \in \mathbb{R}$ on the growth function such that $\forall x \in \mathbb{T}^d$ and $\forall n > 0$

$$r_x(n) \geq C.$$

If we consider the model without rescaling, we assume $C = -1$.

2. (Upper and lower bound on the population size) There exists $C' > 0$ such that $\forall x \in \mathbb{T}^d$ we have

$$\lim_{n \to 0+} r_x(n) > C'$$

and there exists $n_{\max} > 0$ such that all population sizes $n > 0$ and all locations $x \in \mathbb{T}^d$

$$(1 + r_x(n))n < n_{\max}.$$

3. (Regularity of the growth function) Assume that for all population sizes $n > 0$ the function $r_\cdot(n) \in C(\mathbb{T}^d)$ and that

$$\sup_{x \in \mathbb{T}^d} \sup_{n \geq 0} \sup_{0 \leq |\kappa| \leq 4} \left| \frac{\partial^\kappa}{\partial n^\kappa} r_x(n) \right| < \infty.$$

Definition 4.2.2. Let μ, R be positive constants, $u \in (0,1)$, $r : \mathbb{T}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ a growth function satisfying Assumption 4.2.1 and Π be a Poisson point process on $\mathbb{R}_+ \times \mathbb{T}^d \times [0,1]$ with intensity $(n_{\max} + 1)(dt \otimes dx \otimes dw)$.

We define the modified spatial Λ-Fleming-Viot process (mSLFV) as the stochastic process $\rho = (\rho_t, t \geq 0)$ with population size $n_t(z) = \mathcal{H}\rho_t(z)$, which evolves in Ξ according to:

1. Reproduction: For any point $(t, x, w) \in \Pi$ we perform the following steps. In case $w \geq (\overline{n}_t(x) + 1)/(n_{\max} + 1)$, we ignore the event. Otherwise, we choose a parental type k_0 according to the normalized mass distribution in the ball $B(x, R)$, i.e. with respect to

$$\frac{1}{\int_{B(x,R)} n_{t-}(z)dz} \int_{B(x,R)} \rho_{t-}(y, dk)dy.$$

Then we remove a proportion

$$u' := \frac{u}{\overline{n}(x,R)+1}$$

of the population inside the affected area $B(x,R)$ and introduce a mass

$$u'\Big(1+r(\overline{n}(x,R))\Big)\overline{n}(x,R)$$

of offspring of type k_0, i.e. for every $z \in B(x,R)$ we set

$$\rho_t(z,dk) = (1-u')\rho_{t-}(z,dk) + u'\Big(1+r_x(\overline{n}(x,R))\Big)\overline{n}(z,R)\delta_{k_0}(dk).$$

2. Mutation: If t_1 and t_2 are two consecutive times at which a spatial location $z \in \mathbb{T}^d$ is affected by events, we set for times s with $t_1 \leq s < t_2$,

$$\rho_s(z,dk) = e^{-\mu(s-t_1)}\rho_{t_1}(z,dk) + n_{t_1}(z)\left(1-e^{-\mu(s-t_1)}\right)dk.$$

An interpretation of the mutation mechanism was given in Section 4.1 If there is no ambiguity, we will abbreviate $\overline{n}(z,R)$ by $\overline{n}$. As the core $D^{core}(\mathcal{G})$ of the domain $D(\mathcal{G})$ of the generator $\mathcal{G}$ of the mSLFV process, we use the set of functions

$$D^{core}(\mathcal{G}) := \Big\{F_{f,\phi} : \Xi \to \mathbb{R}, \rho \mapsto f(\langle\rho,\phi\rangle) \;\; : \;\; \phi \in C_c(\mathbb{T}^d), f \in C^1(\mathbb{R})\Big\}.$$

The generator of the process acting on functions in $D^{core}(\mathcal{G})$ is

$$\begin{aligned}
&\mathcal{G}F_{f,\phi}(\rho)\\
&= \mu\Big\langle\rho, \int_{[0,1]}(\phi(\cdot,k')-\phi(\cdot,k))dk'\Big\rangle f'(\langle\rho,\phi\rangle)\\
&\quad + \int_{\mathbb{T}^d}\frac{1}{\int_{B(x,R)\times[0,1]}\rho(y,dk)dy}\int_{B(x,R)\times[0,1]}\\
&\qquad (\overline{n}+1)\Big(f\Big(\Big\langle\rho + \frac{u}{\overline{n}+1}\mathbb{1}_{B(x,R)}(\delta_{k_0}[1+r(\overline{n})]\overline{n}-\rho), \phi\Big\rangle\Big) - f(\langle\rho,\phi\rangle)\Big)\\
&\qquad\qquad \rho(y,dk_0)dydx.
\end{aligned} \tag{4.6}$$

We are working on the torus $\mathbb{T}^d$ and our assumptions ensure a finite total rate of events.

Proposition 4.2.3. *Under Assumption 4.2.1, the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$ is well-posed, i.e. there exists a unique solution.*

Remark 4.2.4. The model could be generalised; this remark presents some of the possibilities.

As we will see in the next section, ancestral lineages follow the gradient of the population size. This vector field is irrotational. There are situations such as the aforementioned octopus population subject to the circular drift around Antarctica [Lau+23], where one can imagine rotational biases of the ancestral lineages. In general, an additional vector field might be used to describe biases in the dispersal patterns. One could implement this in the model, for example, by incorporating a bias in the sampling location of the parental type.

As already mentioned, weakening the existence of $n_{\max}$ would require more work to control how rarely large population sizes are attained to prove e.g. tightness of the sequence of processes under scaling. Removing the assumed lower bound $n_{\min}$ on the population size, would complicate any bound used on the inverse of the population.

Currently, as we fix the impact u and the radius $R > 0$, given the population profile n and event centre $z \in \mathbb{T}^d$ the mass added in an event is deterministic. We could incorporate a random factor ξ with $\mathbb{E}(\xi) = 1$ in the mass added in reproduction.

Remark 4.2.5. Working on a torus $\mathbb{T}^d$ ensures uniqueness of the process, as events then occur at a finite rate. Existence of the process can be guaranteed even on $\mathbb{R}^d$ using a sequence of approximating processes. Uniqueness of a varying population size SLFV on $\mathbb{R}^d$ has yet to be shown. To our knowledge uniqueness for SLFV processes has been obtained so far using either duality with a simpler process or lookdown constructions, where the process is constructed as the De-Finetti measure of an exchangeable sequence. The varying population size renders the dual process intractable in our case. A potential lookdown construction of our model could be based on [EK19, Section 4.2]. Unfortunately, the varying population size leads to a particle representation in which individuals change level in every event (so at an infinite rate), which prevents a uniqueness proof.

4.3 Main results

The centrepiece of this work is a central limit theorem for the process. We will see that it allows us to calculate the probability of identity by descent and reflects the behaviour of ancestral lineages. These results generalise the findings of the fixed radius case of [For22] to varying population size.

4.3.1 Law of large numbers for the population size

We will apply the following scaling to the process. As $N \to \infty$, we will let the impact $u \to 0$. This corresponds to letting the strength of the genetic drift become smaller and smaller, which in turn emulates a larger population size (even though the model is already an infinite-population limit). As 'individuals' participate in events at a rate proportional to the impact, we will accelerate time by a scaling inverse to the scaling of the impact. In addition, we rescale space and time in the usual way to obtain a Brownian motion from a jump process. The mutation rate is scaled inversely to time, so that mutation remains of an order visible in the limit. Equally, the growth function needs to decrease to behave reasonably as we accelerate time.

Under this scaling the population size will converge to the solution n_t of a reaction-diffusion equation, while the typed process ρ_t^N converges to Lebesgue measure λ on $\mathbb{T}^d \times [0,1]$ multiplied by the limiting population size n_t. Lebesgue measure appears as the mutation mechanism dominates in the limit. Equivalently, we could view λ as the element of Ξ associating to almost every position a uniform distribution on $[0,1]$.

More formally, let δ_N be an arbitrary sequence of positive numbers converging to zero. We define the rescaled process $\rho_t^N(z, dk)$ in two steps.

1. First, we let $\mathfrak{p}^N = (\mathfrak{p}_t^N, t \geq 0)$ be the mSLFV process with rescaled parameters and growth function

$$u_N := \frac{u_0}{N}, \qquad \mu_N := \delta_N^2 \frac{\mu}{N}, \qquad r_z^N := \delta_N^2 r_{\delta_N z} \tag{4.7}$$

for $z \in \delta_N^{-1}\mathbb{T}^d$.

2. Then, for $z \in \mathbb{T}^d$, we apply the spatial and temporal scaling

$$\rho_t^N(z, dk) := \mathfrak{p}^N_{Nt/\delta_N^2}\Big(\frac{z}{\delta_N}, dk\Big). \tag{4.8}$$

As we see in Lemma 4.3.1 the evolution of the rescaled population size can be characterised by a martingale problem. The parental choice and subsequent dispersal of offspring results in the double average operator

$$\begin{aligned}\mathcal{L}^N\psi(z) &:= \frac{uV_R\delta_N^{-2}}{V^2_{\delta_N R}}\int_{B(z,\delta_N R)}\int_{B(x,\delta_N R)}\Big[\psi(y)-\psi(z)\Big]dydx \\ &= uV_R\delta_N^{-2}\Big(\overline{\overline{\psi}}(z,\delta_N R)-\psi(z)\Big),\end{aligned} \tag{4.9}$$

whereas the growth of the population size will be described by

$$\begin{aligned}\mathcal{R}_n^N\psi(z) &:= uV_R\frac{1}{V^2_{\delta_N R}}\int_{B(z,\delta_N R)} r_x(\overline{n}(x,\delta_N R))\int_{B(x,\delta_N R)}\psi(y)dydx \\ &= uV_R\overline{r(\overline{n})\overline{\psi}}(z,\delta_N R).\end{aligned} \tag{4.10}$$

Lemma 4.3.1. *Let $\phi \in L^1(\mathbb{T}^d)$. Then,*

$$Y_t^N(\phi) := \Big\langle\, n_t^N, \phi\,\Big\rangle - \Big\langle\, n_0^N, \phi\,\Big\rangle - uV_R\int_0^t \Big\langle\, n_s^N, (\mathcal{L}^N + \mathcal{R}^N_{n_s^N})\phi\,\Big\rangle\, ds$$

is a square-integrable martingale with predictable quadratic variation

$$\begin{aligned}&\Big\langle\, Y^N(\phi)\,\Big\rangle_t \\ &= \int_0^t \frac{u^2}{\delta_N^{d+2}}\int_{\mathbb{T}^d}\frac{1}{\overline{(\overline{n_s^N}(x,\delta_N R)+1)N}} \\ &\quad\left(\int_{\mathbb{T}^d}\mathbb{1}_{B(x,\delta_N R)}(z)\Big([1+\delta_N^2 r_x(\overline{n_s^N}(x,\delta_N))]\overline{n_s^N}(x,\delta_N R)-n_s^N(z)\Big)\phi(z)dz\right)^2 dxds.\end{aligned} \tag{4.11}$$

In the limit as $N\to\infty$, the first operator $\mathcal{L}^N$ will result in a Laplacian, and the growth operator $\mathcal{R}_n^N$ in a reaction term so that the population size converges to the solution of a reaction-diffusion equation.

Proposition 4.3.2. *Fix $T>0$, and let n_t^N denote the population size of the rescaled process ρ_t^N. Then,*

$$\lim_{N\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]} d\Big(n_t^N, n_t\Big)\right] = 0, \tag{4.12}$$

where n_t solves the reaction-diffusion equation

$$\frac{dn_t}{dt} = uV_R\Big(\frac{R^2}{d+2}\Delta n_t + r(n_t)n_t\Big), \tag{4.13}$$

and d is the metric defined in (4.4).

The proof of this theorem is contained in Section 4.4. It follows from [FP17, Section 3.2-4.3], but we included it as the preliminary result Lemma 4.4.3 plays a crucial role in the later proof of the central limit theorem.

4.3.2 Central limit theorem

We will consider the fluctuations of ρ_t^N around Lebesgue measure on $\mathbb{T}^d \times [0,1]$ multiplied by a deterministic approximation m_t^N to the population size n_t^N

$$Z_t^N := \sqrt{N\delta_N^{2-d}}\Big(\rho_t^N - \lambda m_t^N\Big). \tag{4.14}$$

Underlying this scaling is the idea that the process ρ_t^N converges to λn_t - although we will not explicitly formulate this as a result. Naturally, we assume that the sequence δ_N is chosen in such a way that $N\delta_N^{2-d} \to \infty$ as $N \to \infty$. In fact, for the proof of Lemma 4.6.4, we will assume even

$$N^{\frac{1}{2}}\delta_N^{1+\frac{d}{2}} \xrightarrow{N\to\infty} \infty. \tag{4.15}$$

The deterministic approximation to the population size that we take is defined by

$$\Big\langle m_t^N, \phi \Big\rangle := \Big\langle m_0^N, \phi \Big\rangle + \int_0^t \Big\langle m_s^N, (\mathcal{L}^N + \mathcal{R}_{m_s^N}^N)\phi \Big\rangle \, ds. \tag{4.16}$$

Comparing to Lemma 4.3.1, this formula differs from a representation of n_t^N by a missing martingale term which would contain the randomness. The scaling (4.14) is the same as in specific settings of [For22] and Chapter 3. The term $N\delta_N^{2-d}$, the square of $\sqrt{N\delta_N^{2-d}}$, will compensate a corresponding term in the predictable quadratic variation term (4.11) of Lemma 4.3.1, which results form a change of variables of (4.11).

We denote by $\mathcal{D}(\mathbb{T}^d \times [0,1])$ functions $\phi : \mathbb{T}^d \times [0,1] \to \mathbb{R}$ which are smooth in the spatial component. Let $\mathcal{D}'(\mathbb{T}^d \times [0,1])$ be the dual space of generalised functions and $\mathbb{D}(\mathbb{R}_+, \mathcal{D}'(\mathbb{T}^d \times [0,1]))$ the space of càdlàg paths in the dual space. Recall from (4.5) that $\mathcal{H}Z(\cdot) := \int_0^1 Z_t(\cdot, dk)$.

Theorem 4.3.3. *The sequence of fluctuations* $(Z_t^N, t \geq 0), N = 1, 2, \ldots$ *converges in distribution in* $D(\mathbb{R}_+, \mathcal{D}'(\mathbb{T}^d \times [0,1]))$ *to the mild solution* $(Z_t, t \geq 0)$ *of the stochastic partial differential equation*

$$\begin{cases} dZ_t = \left[uV_R \left(\frac{R^2}{d+2} \Delta + r(n_t) \right) Z_t - \mu Z_t + \Big(\mu + uV_R n_t r'(n_t) \Big) \mathcal{H} Z_t \right] dt + dW_t, \\ Z_0 = 0. \end{cases} \tag{4.17}$$

The Gaussian random field W *is uncorrelated in time and its covariation measure acts on functions* $\phi \in \mathcal{D}(\mathbb{T}^d \times [0,1])$ *through*

$$\Big\langle \mathcal{Q}_t, \phi \otimes \phi \Big\rangle := u^2 V_R^2 \int_{\mathbb{T}^d} \frac{n_t(z)^2}{n_t(z)+1} \int_{[0,1]} \left(\phi(z,k) - \int_{[0,1]} \phi(z,k') dk' \right)^2 dk dz. \tag{4.18}$$

We will prove the theorem in Section 4.6. The arguments synthesise and generalise ideas of [For22] and [FP17].

Remark 4.3.4. The idea behind the classical central limit theorem is to recentre by the mean of the random variables involved. In fact, there is more freedom regarding the recentring, as long the sequence used to recentre converges to the 'limiting mean'. The same freedom is present in the spatial central limit theorems considered here. In Equation (4.14), recentring by λn_t^N would equally have been possible and would result in the above SPDE without the term involving $\mathcal{H} Z_t$. We chose to recentre by the deterministic approximation λm_t^N as the calculations will illustrate that the Wright-Malécot formula is the same with or without the terms involving $\mathcal{H} Z_t$, so invariant under the recentring.

4.3.3 Wright-Malécot formula

Next, we explain how the fluctuations Z_t^N can be related to the probability of identity. As the processes assume values in Ξ and are only unique up to a spatial nullset, the probability of identity $P_t^N(\psi_1, \psi_2)$ will depend on two probability densities ψ_1, ψ_2 describing the sampling location. The probability of sampling the same type from two probability distributions d_1, d_2 is given by evaluating the distributions d_1, d_2

with respect to the indicator of the diagonal $\mathbb{1}_{\diagdown}(k_1, k_2) := \mathbb{1}_{k_1=k_2}$. Accounting for the varying population size, we hence define the probability of identity as

$$\begin{aligned} &P_t^N(\psi_1, \psi_2) \\ &= \mathbb{E}\left[\frac{1}{\langle n_t^N, \psi_1\rangle\langle n_t^N, \psi_2\rangle}\int_{(\mathbb{T}^d)^2} \mathbb{1}_{\diagdown}(k_1, k_2)\psi_1(z_1)\psi_2(z_2)\rho_t^N(z_1, dk_1)\rho_t^N(z_2, dk_2)dz_1dz_2\right] \\ &= \mathbb{E}\left[\frac{\left\langle \rho_t^N \otimes \rho_t^N, (\psi_1 \otimes \psi_2)\mathbb{1}_{\diagdown}\right\rangle}{\langle n_t^N, \psi_1\rangle\langle n_t^N, \psi_2\rangle}\right]. \end{aligned} \tag{4.19}$$

We think of this sampling scheme as giving weight $\psi_1(z)$ to all individuals in location $z \in \mathbb{T}^d$, and then sampling an individual proportionally to these weights. The procedure is repeated with ψ_2 for the second sampled individual.

The question still remains how the probability of identity is related to the central limit theorem Theorem 4.3.3. The key observation, made in [For22], is that as Lebesgue measure does not charge the diagonal, we can utilize the relationship $Z_t^N := \sqrt{N\delta_N^{2-d}}\big(\rho_t^N - \lambda m_t^N\big)$, to rewrite (4.19) as

$$P_t^N(\psi_1, \psi_2) = \frac{1}{N\delta_N^{2-d}}\mathbb{E}\left[\frac{\left\langle Z_t^N \otimes Z_t^N, (\psi_1 \otimes \psi_2)\mathbb{1}_{\diagdown}\right\rangle}{\langle n_t^N, \psi_1\rangle\langle n_t^N, \psi_2\rangle}\right].$$

Rearranging,

$$N\delta_N^{2-d}P_t^N(\psi_1, \psi_2) = \mathbb{E}\left[\frac{\left\langle Z_t^N \otimes Z_t^N, (\psi_1 \otimes \psi_2)\mathbb{1}_{\diagdown}\right\rangle}{\langle n_t^N, \psi_1\rangle\langle n_t^N, \psi_2\rangle}\right]. \tag{4.20}$$

The above equation shows that the scaling factor $N\delta_N^{2-d}$ of the central limit theorem is precisely the scaling factor required for $P_t^N(\psi_1, \psi_2)$ to converge to a non-trivial limit.

To use the central limit theorem, we would like to take the limit $N \to \infty$ on both sides of (4.20) and exchange limit and expectation on the right-hand side to obtain

$$\lim_{N\to\infty} N\delta_N^{2-d}P_t^N(\phi, \psi) = \mathbb{E}\left[\frac{\left\langle Z_t \otimes Z_t, (\psi_1 \otimes \psi_2)\mathbb{1}_{\diagdown}\right\rangle}{\langle n_t, \psi_1\rangle\langle n_t, \psi_2\rangle}\right].$$

When the population size is constant, this can be achieved via an application of the dominated convergence theorem (see [For22]). As we will elaborate in Remark 4.3.8, in the general case there remains a technicality resulting from the

intrinsically varying population size, which prevented us from applying dominated convergence. Replacing the dependency of the growth function on the population size n^N in Definition 4.2.2 by a dependency on the deterministic approximation m^N would eliminate this technicality. In this case, bounding the population size would become more involved as

$$\Big(1 + r(\overline{m_t^N})\Big)\overline{n_t^N}$$

could potentially grow to infinity under our current assumption. As this would not add to the biological interpretation of this work, we refrain from dealing with this additional technicality.

Conjecture 4.3.5. *For any $t \geq 0$, any $\psi_1, \psi_2 \in \mathcal{D}(\mathbb{T}^d)$, we have*

$$\lim_{N\to\infty} N\eta_N P_t^N(\psi_1, \psi_2) = \mathbb{E}\left[\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown}\psi_1 \otimes \psi_2\rangle}{\langle n_t, \psi_1\rangle\langle n_t, \psi_2\rangle}\right].$$

Our aim is now to compute the right-hand side. To avoid dealing with the fraction $\langle n_t, \psi_i\rangle$, we choose

$$\vartheta_i := \frac{n_t\psi_i}{\langle n_t, \psi_i\rangle},$$

which transforms the quantity to

$$\mathbb{E}\left[\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown}\psi_1 \otimes \psi_2\rangle}{\langle n_t, \psi_1\rangle\langle n_t, \psi_2\rangle}\right] = \mathbb{E}\Big[\Big\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown}\Big(\frac{\vartheta_1}{n_t} \otimes \frac{\vartheta_2}{n_t}\Big)\Big\rangle\Big].$$

In Section 4.7, we will use this expression to prove the following theorem.

Theorem 4.3.6. *Let $G_{s,t}$ be the fundamental solution associated to*

$$-\frac{d}{ds}G_{s,t}\psi = A_s^* G_{s,t}\psi,$$

where A_s^* is the adjoint of

$$A_s\phi = \frac{uV_R R^2}{d+2}\Big(\Delta\phi + 2\frac{\nabla n_s}{n_s}\cdot\nabla\phi\Big) \tag{4.21}$$

and let ϑ_i be as above. Then the Wright-Malécot formula for two probability densities ψ_1, ψ_2 is given by

$$\mathbb{E}\left[\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown}\psi_1 \otimes \psi_2\rangle}{\langle n_t, \psi_1\rangle\langle n_t, \psi_2\rangle}\right] = u^2V_R^2\int_0^t\int_{\mathbb{T}^d}\frac{e^{-2\mu(t-s)}}{n_s(z)+1}G_{s,t}\vartheta_1(z)G_{s,t}\vartheta_2(z)dzds. \tag{4.22}$$

The function $\vartheta_i = \frac{n_t \psi_i}{\langle n_t, \psi_i \rangle}$ can be seen as being the limiting density according to which the sampling of positions occurs.

Remark 4.3.7. There is at least one more way in which one could define the probability of identity. Informally, we could instead sample two positions z_1, z_2 according to two the densities ψ_1, ψ_2. Then, the probability that two individuals sampled according to the probability measures $(n_t^N(z_1))^{-1}\rho_t^N(z_1)$ and $(n_t^N(z_2))^{-1}\rho_t^N(z_2)$ are of the same type is again given by the integral with respect to $\mathbb{1}_\diagdown(k_1, k_2) = \mathbb{1}_{k_1=k_2}$. Formally, this would correspond to the integral

$$\begin{aligned} P_t^N(\psi_1, \psi_2) &= \mathbb{E}\left[\int_{(\mathbb{T}^d)^2} \mathbb{1}_\diagdown(k_1, k_2)\psi_1(z_1)\psi_2(z_2)\frac{\rho_t^N(z_1, dk_1)}{n_t^N(z_1)}\frac{\rho_t^N(z_2, dk_2)}{n_t^N(z_2)}dz_1 dz_2\right] \\ &= \mathbb{E}\left[\left\langle \frac{\rho_t^N}{n_t^N} \otimes \frac{\rho_t^N}{n_t^N}, (\psi_1 \otimes \psi_2)\mathbb{1}_\diagdown \right\rangle\right]. \end{aligned}$$

Our choice of sampling circumvents some problems with this approach arising from the regularity of ρ_t^N/n_t^N, whose fractional character was leading to difficulties.

Remark 4.3.8. Here, we wish to elaborate on the technicalities which prevented us from proving Conjecture 4.3.5. We will represent the fluctuations $\langle Z^N, \phi \rangle$ as a stochastic integral plus error terms, which result from the varying population size and vanish in the limit (see (4.31)). The martingale problem associated to $Z^N \otimes Z^N$ would contain very similar error terms. Basically each error term would involve an additional copy of Z^N. The additional Z^N becomes problematic as it is not spatially averaged, preventing an application of Lemma 4.5.5 or an appropriate generalization. One might consider the test function $\Phi' = \mathbb{1}_\diagdown \overline{\phi_1} \otimes \overline{\phi_2}$ to gain the additional regularity in form of a spatial average. In the probability of identity, this would correspond to sampling a position and then adding a uniformly distributed vector from the unit ball. Unfortunately, most of the representations of the fluctuations used in this work involve choosing solutions to differential equations as test function to eliminate parts of the generator. If the growth term is among the eliminated ones, so is featured in the differential equation, the solution of the differential equation with averaged initial condition is not the same as the average of the original solution. This prevents the average from being transferred from the test function onto Z^N.

4.3.4 Interpretation and simulation

Next, we wish elaborate on how Theorem 4.3.6 reflects the behaviour of the ancestral lineages in the process and provide visualizations of the corresponding Wright-Malécot formulae.

Behaviour of two ancestral lineages

Remember that in Theorem 4.3.6 we describe a probability of identity: in order for two individuals sampled from two different locations at time t to have the same type, we need to track both their ancestral lineages to a common ancestor at some time $s < t$ and require that neither lineage acquires a mutation up to this point. The formula (4.22) can then be interpreted as follows:

Mutation From an individual-based perspective each individual mutates at rate μ. Hence, backwards in time starting from time t the probability that neither ancestral lineage has seen a mutation at time s in the past is given by $e^{-2\mu(t-s)}$.

Motion We can write

$$G_{s,t}\vartheta(z) = \int_{\mathbb{T}^d} p_{s,t}(z,x)\vartheta(x)dx,$$

where $z \mapsto p_{t-u,t}(z,x)$ is the transition density of $(X_u)_{u\in[0,t]}$, started from $X_0 = x$ and with the generator $u \mapsto A_{t-u}$ comprising the Laplacian and the drift according to twice the gradient of the population profile from (4.21). The quantity $G_{s,t}\vartheta_1(z)G_{s,t}\vartheta_2(z)$ can be seen as the 'probability' for both lineages to be at the same position z a time $t-s$ in the past. This illustrates that the lineages behave according to Brownian motions with a drift according to the gradient given by the population profile.

Coalescence The term $(n_s(z)+1)^{-1}$ captures the fact that once both lineages are at the same position z, coalescence occurs at a rate inversely proportional to the population size $n_s(z)$. This fits the intuition that with a larger

local population size, the time to the most recent common ancestor while staying in the same deme should be longer.

Notably, the motion of ancestral lineages is very much comparable to [Eth+24, Theorem 2.24, Corrollary 2.28 and Examples 3.2]. For example, in the case of profile w of the population behaving according to a travelling wave solution to the Fisher-KPP equation the authors obtain for the generator of ancestral lineages

$$\mathcal{L}f = \partial_{xx} f + 2\frac{\partial_x w}{w}\partial_x + 2\partial_x f.$$

The $2\partial_x f$ originates from the moving frame of the travelling wave with the remainder precisely equalling (4.21). Their work is based on the lookdown framework and we arrive at a similar expression from a central limit theorem perspective. We can even go further and describe the behaviour of two ancestral lineages with coalescence happening inversely proportional to the population size. A hint that this is heuristically similar in their framework can be found in the discussion of Terence Tsui Ho Lung's Phd thesis [Tsu23, Section 8.1.1].

Simulation and analytical prediction

Aim In this section, we visualize the accuracy of the analytical result Theorem 4.3.6 by comparing it to simulated probabilities of identity in Figure 4.2. We consider a scenario in which the population density is much lower at a central location (see Figure 4.1). In nature, this might be caused by, for example, a mountain range and will present in our model a barrier to gene flow. Ancestral lineages are attracted by higher population density, as we sample the parental type uniformly among the individuals in the affected region, so more likely from a densily populated area.

We will need to discretize space in both cases and will consider the interval $\{0, 1, 2, \ldots, 99, 100\}$. The analytical result Theorem 4.3.6 was telling us the motion and coalescence behaviour of ancestral lineages. Discretizing this behaviour, we can calculate the analytical prediction Θ of the Wright-Malécot-Formula given by (4.22). We will compare this prediction with the probability of identity $\mathcal{P}$ resulting from forwards in time simulations of the typed process from Definition 4.2.2. To compare

the analytical (limiting) result Θ with the simulated (prelimiting) probabilities, we 'undo' the scaling. Our analyis predicts

$$N\delta_N \mathcal{P}^N(x) \simeq \Theta\Big(\frac{N}{\delta_N^2}\mu_N, Nu_N, \delta_N x\Big).$$

The full details of prediction and simulation will be given in Appendix B.1.

Setting We discretize space and consider the interval $\{0, 1, 2, ..., 99, 100\}$. As illustrated by Figure 4.1, the population size profile will be at stationarity for the following choice of growth function

$$r_x(n) = \max\left\{\min\left\{14\frac{|x-50|}{50}, 7\right\} + 1 - n, -1\right\}. \tag{4.23}$$

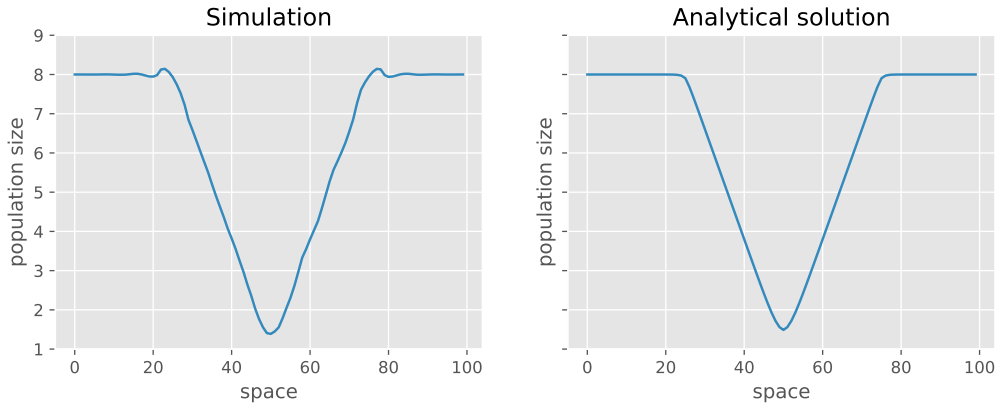


Figure 4.1: Comparison between the population size of the process run for a time of 125 on the left and the analytical solution on the right, both corresponding to (4.23). As the radius of events is positive, in events which fall right next to the edge of the valley, the growth function is positive leading to folds on the rim of the population density valley. The analytical solution depicts the solution to a discretized version of (4.13).

To simulate the probability of identity, we consider a discretization of the dynamics given in Definition 4.2.2 and follow the proportions of up to 2000 types. For more details, please see Appendix B.1.

The analytical prediction Θ of Figure 4.1 is based on calculating the transition probabilities of the lineages of sampled individuals backwards in time and incorporating the coalescence at a rate inversely proportional to the population

size plus one. The lineages are diffusive with a drift according to the population size profile given in (Figure 4.1).

In Figure 4.2 we consider three reference positions $45, 60$ and 75. For each location x_0 on the x-axis, representing space, we visualize probability of identity of a sample from x_0 and a sample from the reference position. The prediction from the analytical solution is shown by the continuous line. The dashed line represents the mean of the probability of identity in 2000 simulations of the process from Definition 4.2.2, whereas the darker and lighter area of the same colour show the 50-th and 90-th percentiles.

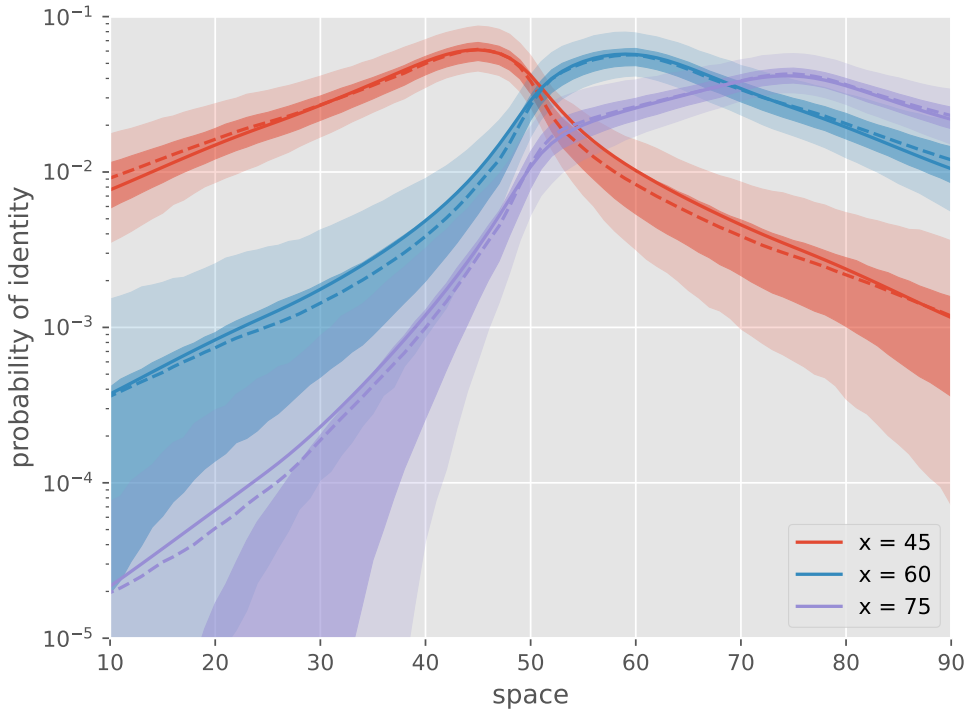


Figure 4.2: Comparison of the probabilities of identity (on a logarithmic scale) from three reference positions $45, 60, 75$ relative to all other locations. For example, red captures the probability of identity sampling one individual from 45 and one from the population at the corresponding position on the x-axis. The continuous line shows the prediction Θ from the analytical calculation based on a calculation of the transition probabilities of ancestral lineages. The dotted lines illustrate the average probability of identity from the simulations of the process, with the lighter regions representing the 50-th and 90-th percentiles after 2000 simulations.

Interpretation The analytical calculation predicts the simulated probabilities of identities well.

The valley in the population density leads to ancestral lineages on the left of the central position 50 being attracted to the left and lineages on the right being attracted to the right. Consequently, we would expect a rapid decorrelation of the probabilities of identity by descent around position 50, which is confirmed by both a simulation of the process and the analytical prediction (see Figure 4.2).

The mean of the simulated probabilities of identity, represented by the dashed line, is in the centre of the 50-th and 90-th percentiles close to reference positions $45, 60$ and 75. This illustrates that in fact even without the expectation

$$\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown} \psi_1 \otimes \psi_2 \rangle}{\langle n_t, \psi_1 \rangle \langle n_t, \psi_2 \rangle} \simeq \mathbb{E}\left[\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown} \psi_1 \otimes \psi_2 \rangle}{\langle n_t, \psi_1 \rangle \langle n_t, \psi_2 \rangle} \right]$$

for sampling positions described by ψ_1 and ψ_2 being not too far apart. For locations further away from the reference positions a divergence of the mean from the median of the probabilities of identity occurs (particularly visible in the purple part on the left-hand side).

4.4 Convergence of the population size process

In this section, we will explain how the law of large numbers/Proposition 4.3.2 for the population size can be proven by adapting the strategy of [FP17, Section 3.2-4.3]. We consider test functions $\phi : \mathbb{T}^d \to \mathbb{R}$ depending only on space, so that the process evaluated against them describes the evolution of the population size. In other words, taking $\phi(x, k) = \phi(x)$, we define $\left\langle n_t^N, \phi \right\rangle := \left\langle \rho_t^N, \phi \right\rangle$ under the scaling (4.7) and (4.8). We start with the characterisation via a martingale problem.

Proof of Lemma 4.3.1. If we apply the generator (4.6) to test functions only acting on space, we get

$$\mathcal{G}F_{1,\phi}(n) = \int_{\mathbb{T}^d} (\overline{n} + 1) \left(\left\langle n + \frac{u}{\overline{n} + 1} \mathbb{1}_{B(x,R)} \Big([1 + r_x(\overline{n})]\overline{n} - n \Big), \phi \right\rangle - \left\langle n, \phi \right\rangle \right) dx.$$

The scaling (4.7) and (4.8) amounts to considering

$$\left\langle \rho_t^N, \phi \right\rangle = \left\langle \mathfrak{p}^N_{Nt/\delta_N^2}(\delta_N \cdot), \phi^N \right\rangle$$

for $\phi^N = \delta_N^d \phi(\delta_N \cdot)$. Recall that $\mathfrak{p}^N$ denotes the mSLFV with $u^N = u/N$ and $r_\cdot^N = \delta_N^2 r_{\delta_N \cdot}$ on $\delta_N^{-1}\mathbb{T}^d$. In the following calculation, we will apply the formula for multidimensional substitution

$$\int_{\mathbb{T}^d} \phi(x)dx = \int_{\delta_N^{-1}\mathbb{T}^d} \phi(\delta_N x)\delta_N^d dx.$$

Indeed, we arrive at the desired expression:

$$\begin{aligned}
&\mathcal{G}^N F_{1,\phi}(n(\delta_N \cdot)) \\
&= \frac{N}{\delta_N^2} \mathcal{G} F_{1,\phi^N}(n(\delta_N \cdot)) \\
&= \frac{N}{\delta_N^2} \int_{\delta_N^{-1}\mathbb{T}^d} (\overline{n(\delta_N \cdot)}(x,R) + 1) \Bigg[\Bigg\langle n(\delta_N \cdot) + \frac{u^N}{\overline{n(\delta_N \cdot)}(x,R) + 1} \mathbb{1}_{B(x,R)} \\
&\qquad \Big([1 + r_x^N(\overline{n(\delta_N \cdot)}(x,R))] \overline{n(\delta_N \cdot)}(x,R) - n(\delta_N \cdot) \Big), \phi^N \Bigg\rangle - \left\langle n(\delta_N \cdot), \phi^N \right\rangle \Bigg] dx \\
&= \frac{u}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \Bigg\langle \mathbb{1}_{B(x/\delta_N,R)} \\
&\qquad \Big([1 + r^N_{x/\delta_N}(\overline{n(\delta_N \cdot)}(x/\delta_N,R))] \overline{n(\delta_N \cdot)}(x/\delta_N,R) - n(\delta_N \cdot) \Big), \phi^N \Bigg\rangle dx \\
&= \frac{u}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \int_{\delta_N^{-1}\mathbb{T}^d} \mathbb{1}_{B(x,\delta_N R)}(\delta_N z) \\
&\qquad \Big([1 + \delta_N^2 r_x(\overline{n(\delta_N \cdot)}(x/\delta_N,R))] \overline{n(\delta_N \cdot)}(x/\delta_N,R) - n(\delta_N z) \Big) \delta_N^d \phi(\delta_N z) dz dx \\
&= \frac{u}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \mathbb{1}_{B(x,\delta_N R)}(z) \\
&\qquad \Big([1 + \delta_N^2 r_x(\overline{n(\delta_N \cdot)}(x/\delta_N,R))] \overline{n(\delta_N \cdot)}(x/\delta_N,R) - n(z) \Big) \phi(z) dz dx \\
&= \frac{u}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \mathbb{1}_{B(x,\delta_N R)}(z) \Big([1 + \delta_N^2 r_x(\overline{n}(x,\delta_N R))] \overline{n}(x,\delta_N R) - n(z) \Big) \phi(z) dz dx \\
&= \frac{u V_{\delta_N R}}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \left(\frac{1}{V_{\delta_N R}} \int_{B(z,\delta_N R)} [1 + \delta_N^2 r_x(\overline{n}(x,\delta_N R))] \overline{n}(x,\delta_N R) dx - n(z) \right) \phi(z) dz \\
&= u V_R \delta_N^{-2} \left\langle \overline{[1 + \delta_N^2 r(\overline{n})]\overline{n}} - n, \phi \right\rangle \\
&= u V_R \left\langle n, (\mathcal{L}^N + \mathcal{R}_n^N)\phi \right\rangle.
\end{aligned}$$

The last step uses that both $\mathcal{L}^N$ and $\mathcal{R}_n^N$ are self-adjoint operators (which we show in the proof of Lemma 4.5.2). Similarly, for the quadratic variation

$$\begin{aligned}
&\frac{N}{\delta_N^{d+2}} \int_{\mathbb{T}^d} (\overline{n}(x, \delta_N R) + 1) \\
&\qquad \left(\left\langle n + \frac{u}{(\overline{n}(x, \delta_N R) + 1)N} \mathbb{1}_{B(x, \delta_N R)} \Big([1 + \delta_N^2 r_x(\overline{n}(x, \delta_N)] \overline{n}(x, \delta_N) - n \Big), \phi \right\rangle \right. \\
&\qquad\qquad \left. - \left\langle n, \phi \right\rangle \right)^2 dx \\
&= \frac{u^2}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \frac{1}{(\overline{n}(x, \delta_N R) + 1)N} \\
&\qquad\qquad \left(\int_{\mathbb{T}^d} \mathbb{1}_{B(x, \delta_N R)}(z) \Big([1 + \delta_N^2 r_x(\overline{n}(x, \delta_N)] \overline{n}(x, \delta_N R) - n(z) \Big) \phi(z) dz \right)^2 dx. \quad \square
\end{aligned}$$

The main idea of the proof of Proposition 4.3.2 is to separate the convergence of the population size into steps:

1. The deterministic approximation m^N of (4.16) given by

$$\left\langle m_t^N, \phi \right\rangle := \left\langle m_0^N, \phi \right\rangle + \int_0^t \left\langle m_s^N, (\mathcal{L}^N + \mathcal{R}_{m_s^N}^N) \phi \right\rangle ds,$$

   can be seen to converge to the solution of the reaction-diffusion equation.

2. Then, we will prove $|n^N - m^N| \to 0$ by defining fluctuations

$$F_t^N = \sqrt{N \delta_N^{2-d}} (n_t^N - m_t^N)$$

   and showing that they are sufficiently bounded. Note that

$$F_t^N = \sqrt{N \delta_N^{2-d}} (n_t^N - m_t^N) = \int_0^1 \sqrt{N \delta_N^{2-d}} (\rho_t^N - m_t^N \lambda)(dk) = \mathcal{H} Z_t^N,$$

   which will become crucial in later sections.

The first step is established by the next lemma. As its proof is similar to several results in Appendix B.2, which we will require for the proof of the main Theorem 4.3.3, we only provide a sketch.

Lemma 4.4.1 (cf. [FP17, Proposition 4.6])**.** *There exist constants $C, C' > 0$, such that for any multi-index $0 \leq |\kappa| \leq 4$*

$$\sup_{0 \leq t \leq T} \|\partial_\kappa m_t^N\|_\infty \leq C,$$

and uniform convergence holds

$$\sup_{0\le t\le T} \|m_t^N - n_t\|_\infty \le \delta_N C'.$$

Proof. Let G_t be the Gaussian kernel with variance $\frac{uV_R R^2}{d+2}$. Then the population size can be represented as

$$n_t(x) = G_t * n_0(x) + \int_0^t (G_{t-s} * r(n_s)n_s)(x)ds.$$

Similarly, if G_t^N denotes the semigroup associated to the generator $\mathcal{L}^N$, we get

$$m_t^N = G_t^N * m_0^N + \int_0^t (G_{t-s}^N * \overline{r(\overline{m_s^N})\overline{m_s^N}})(x)ds.$$

The difference between the two terms can now be bounded using Grönwall's inequality. Using Assumption 4.2.1 the argument of [FP17, Section B.1, Proof of Proposition 4.6] carries over. □

To perform the second step first note that if we assume the same initial conditions $n_0^N = m_0^N a.s.$ the fluctuations satisfy the following semimartingale decomposition

$$\begin{aligned}\left\langle F_t^N, \phi \right\rangle &= uV_R\sqrt{N\delta_N^{2-d}} \int_0^t \left\langle (\mathcal{L}^N + \mathcal{R}_{n_s^N})(n_s^N) - (\mathcal{L}^N + \mathcal{R}_{m_s^N})(m_s^N), \phi \right\rangle ds \\ &\quad + \sqrt{N\delta_N^{2-d}} Y_t^N(\phi) \\ &= uV_R \int_0^t \left\langle \mathcal{L}^N F_s^N + \sqrt{N\delta_N^{2-d}}\Big(r(\overline{n_s^N})\overline{n_s^N} - r(\overline{m_s^N})\overline{m_s^N}\Big)(\delta_N R), \phi \right\rangle ds \\ &\quad + \sqrt{N\delta_N^{2-d}} Y_t^N(\phi).\end{aligned}$$

The martingale $Y_t^N(\phi)$ had been defined in Lemma 4.3.1. This equation is of the same form as [FP17, Equation (45)]. The next lemma controls the martingale part.

Lemma 4.4.2 (cf. [FP17, Lemma 4.4])**.** *There exists $C > 0$ such that for all $\phi : \mathbb{R}_+ \to L^2(\mathbb{T}^d)$ we have*

$$\mathbb{E}\left[\left(\sqrt{N\delta_N^{2-d}} \int_{[0,t]\times\mathbb{T}^d} \phi_s(x) Y^N(dx\ ds)\right)^2\right] \le C \int_0^t \|\phi_s\|_2^2\ ds. \tag{4.24}$$

Proof. Let $V_R(z_1, z_2)$ denote the volume of the intersection of two balls of radius R centred at positions z_1 and z_2. From Lemma 4.3.1 we know that the left-hand side is equal to

$$\begin{aligned}
&\mathbb{E}\Bigg[\int_0^t \frac{u^2}{\delta_N^{2d}} \int_{\mathbb{T}^d} \frac{1}{(\overline{n}(x, \delta_N R) + 1)} \\
&\quad \left(\int_{\mathbb{T}^d} \mathbb{1}_{B(x,\delta_N R)}(z) \Big([1 + \delta_N^2 r_x(\overline{n^N}(x, \delta_N R))] \overline{n^N}(x, \delta_N R) - n^N(z) \Big) \phi_s(z) dz \right)^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad dxds \Bigg] \\
&\leq \mathbb{E}\left[\int_0^t \frac{u^2}{\delta_N^{2d}} \int_{\mathbb{T}^d} \left(\int_{\mathbb{T}^d} \mathbb{1}_{B(x,\delta_N R)}(z)\ 2n_{\max}\ |\phi_s(z)| dz \right)^2 dxds \right] \\
&\leq 4n_{\max}^2 u^2 V_R^2 \mathbb{E}\left[\int_0^t \int_{\mathbb{T}^d} \left(\frac{1}{V_{\delta_N R}} \int_{B(x,\delta_N R)} |\phi_s(z)| dz \right)^2 dxds \right] \\
&\leq 4n_{\max}^2 u^2 V_R^2 \int_0^t \|\phi\|_2^2\, ds,
\end{aligned}$$

where we have used $\|\overline{\phi}\|_2 \leq \|\phi\|_2$ in the last line. □

The crucial idea to control the fluctuations F^N is that the square of the spatial average is uniformly controlled. To prove the convergence of Z^N, we will use Lemma 4.5.5, which can be seen of a generalisation of the next result to involve types.

Lemma 4.4.3 (cf. [FP17, Lemma 4.5])**.** *There exists a constant $C > 0$ such that*

$$\sup_{0 \leq t \leq T} \sup_{z \in \mathbb{T}^d} \mathbb{E}\left[\overline{F_t^N}(z, \delta_N R)^2 \right] \leq \frac{C}{V_{\delta_N R}}.$$

Proof. The remainder of Taylor's expansion to the k-th order of a function $\mathfrak{r} \in C^k(\mathbb{R})$ about the point x is given by

$$\mathfrak{T}_k(x, y, \mathfrak{r}) = \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \mathfrak{r}^{(k)}(x + t(y - x)) dt. \tag{4.25}$$

Applying Taylor's expansion to $\mathfrak{r} = r(x)x$ yields

$$\left\langle F_t^N, \phi \right\rangle = uV_R \int_0^t \left\langle \mathcal{L}^N F_s^N + \overline{F_s^N \mathfrak{T}_1(\overline{n_s^N}, \overline{m_s^N}, \mathfrak{r})}, \phi \right\rangle ds + \sqrt{N \delta_N^{2-d}} Y_t^N(\phi). \tag{4.26}$$

We wish to remove $\mathcal{L}^N$ by using an appropriate test function - a technique which we will use frequently. Let G_t^N be again the semigroup associated to the operator $\mathcal{L}^N$. Substituting the time-dependent test function $G_{t-s}^N * \phi$ in (4.26) yields

$$\begin{aligned}\left\langle F_t^N, \phi \right\rangle = uV_R \int_0^t \left\langle G_{t-s}^N * \left(\overline{F_s^N \mathfrak{T}_1(\overline{n_s^N}, \overline{m_s^N}, \mathfrak{r})}\right), \phi \right\rangle ds \\ + \int_0^t \int_{\mathbb{T}^d} (G_{t-s}^N * \phi)(z) \sqrt{N\delta_N^{2-d}} Y^N(dz\ ds).\end{aligned}$$

Choosing $\phi = \frac{1}{V_R} \mathbb{1}_{|\cdot - z| < \delta_N R}$, the left-hand side $\langle F_t^N, \phi \rangle$ equals $\overline{F_t^N}(z, \delta_N R)$. Squaring the equation, and using Jensen's inequality as well as Lemma 4.4.2 to bound the martingale part, one can set up an application of Grönwall's inequality. We provide details of these steps in the more general setting of Lemma 4.5.5 (in which we include genetic types) and so we omit the details here. □

Lemma 4.4.4 (cf. [FP17, Lemma 4.7])**.** *There exists a constant $C > 0$ such that for any twice continuously differentiable function ϕ with*

$$\max_{p \in \{1,2\}} \max \left\{ \|\phi\|_p, \max_{0 \le |\kappa| \le 2} \|\partial_\kappa \phi\|_p \right\} \le 1, \tag{4.27}$$

we have

$$\mathbb{E}\left[\sup_{0 \le t \le T} \left| \left\langle F_t^N, \phi \right\rangle \right| \right] \le C. \tag{4.28}$$

Proof. Similar to [FP17, Section 3.2], using Taylor's expansion of $\mathfrak{r} = r(x)x$, this time to the second order

$$\begin{aligned}\left\langle F_t^N, \phi \right\rangle = uV_R \int_0^t \left\langle F_s^N, \mathcal{L}^N \phi + \overline{\mathfrak{r}'(\overline{m_s^N})\overline{\phi}}(\delta_N R) \right\rangle ds \\ + \frac{1}{\sqrt{N\delta_N^{2-d}}} \int_0^t \left\langle (\overline{F_s^N})^2, \mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, \mathfrak{r})\overline{\phi} \right\rangle ds + \sqrt{N\delta_N^{2-d}} Y_t^N(\phi)\end{aligned} \tag{4.29}$$

where the last line is of the same form as [FP17, Equation (48)]. To obtain (4.28), we would like to take absolute values, supremum over time and expectation on both sides of (4.29). Inescapably, this will lead to $\mathbb{E}\left| \left\langle F_s^N, \mathcal{L}^N \phi + \overline{\mathfrak{r}'(\overline{m_s^N})\overline{\phi}}(\delta_N R) \right\rangle \right|$ on the right-hand side. We therefore require first a bound on $\mathbb{E}|\langle F_s^N, \phi \rangle|$. Indeed, choosing ϕ^N to be the time-dependent solution to

$$\begin{cases} \partial_s \phi^N(x,s,t) + \mathcal{L}^N \phi^N(x,s,t) - \overline{\mathfrak{r}'(\overline{m_s^N})\overline{\phi^N(s,t)}}(x, \delta_N R) = 0 \\ \phi^N(x,t,t) = \phi(x) \end{cases},$$

one obtains instead the equation (see [FP17, Section 4.1])

$$\begin{aligned}\left\langle F_t^N, \phi \right\rangle = \frac{1}{\sqrt{N\delta_N^{2-d}}} \int_0^t \left\langle (\overline{F_s^N})^2, \mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, \mathfrak{r}) \overline{\phi^N(s,t)} \right\rangle ds \\ + \int_0^t \int_{\mathbb{T}^d} \phi^N(x,s,t) \sqrt{N\delta_N^{2-d}} Y^N(dx\, ds),\end{aligned} \tag{4.30}$$

which is of the same form as [FP17, Equation (58)]. The first term is of the shape necessary to apply Lemma 4.4.3, whereas the second can again be controlled by Lemma 4.4.2. Returning to (4.29) above, one can then apply this bound to the function $\mathcal{L}^N\phi + \overline{\mathfrak{r}'(\overline{m_s^N})\overline{\phi}}(\delta_N R)$. Using in addition to Lemma 4.4.2 and Lemma 4.4.3 bounds on $\mathcal{L}^N$, the result follows. □

We can now conclude the law of large numbers for the population size.

Proof of Proposition 4.3.2. We proceed with the same argument as [FP17, Section 4.3] by combining Lemma 4.4.1 and Lemma 4.4.4. Without loss of generality, we assume the existence of a constant C s.t. the sequence of test functions satisfies $\max_{0\le|\kappa|\le 2} \|\partial_\kappa \phi_n\|_p < C$ uniformly for all $n \ge 1$ and $p \in \{1,2\}$ (see (4.3)). We can now bound

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\le t\le T} d(n_t^N, n_t)\right] \\ &\le \sum_{i=1}^\infty \frac{1}{2^n}\left(\mathbb{E}\Big[\sup_{0\le t\le T}\Big|\left\langle n_t^N, \phi_i\right\rangle - \left\langle m_t^N, \phi_i\right\rangle\Big|\Big] + \sup_{0\le t\le T}\Big|\left\langle m_t^N, \phi_i\right\rangle - \left\langle n_t, \phi_i\right\rangle\Big|\right) \\ &\le \sum_{i=1}^\infty \frac{1}{2^n}\left((N\delta_N^{2-d})^{-1/2}\mathbb{E}\Big[\sup_{0\le t\le T}\Big|\left\langle F_t^N, \phi_i\right\rangle\Big|\Big] + \sup_{0\le t\le T}\Big\|m_t^N - n_t\Big\|_\infty \|\phi_n\|_1\right),\end{aligned}$$

where the right-hand side converges to zero as $N \to \infty$ due to Lemma 4.4.1 and Lemma 4.4.4. □

4.5 Semimartingale decomposition of the fluctuations

In this section, we will derive a representation of the fluctuations and prove some auxiliary results, in preparation for the convergence proofs in the next section.

4.5.1 Rescaled martingale problem for the process

We start by rescaling the martingale problem for the mSLFV process.

Lemma 4.5.1. *Let $(\rho_t, t \geq 0)$ be an mSLFV process and $\mathcal{G}$ be the generator given by (4.6). Then, for any test function $\phi \in L^1(\mathbb{T}^d \times [0,1])$,*

$$\mathcal{M}_t(F_{f,\phi}) := f(\langle \rho_t, \phi \rangle) - f\langle \rho_0, \phi \rangle) - \int_0^t \mathcal{G}F_{f,\phi}(\rho_s)ds$$

is a square integrable martingale with quadratic variation

$$\int_{\mathbb{T}^d} \frac{1}{\int_{B(x,R)\times[0,1]} \rho(y,dk)dy} \int_{B(x,R)\times[0,1]} (\overline{n}+1)\Big[f(\langle \rho + \frac{u}{\overline{n}+1}\mathbb{1}_{B(x,R)}(\delta_{k_0}[1+r(\overline{n})]\overline{n} - \rho), \phi\rangle) - f(\langle \rho, \phi\rangle)\Big]^2 \rho(y,dk_0)dydx$$

In very much the same way as for the population size, the dynamics can be expressed with the operators $\mathcal{L}^N$ and $\mathcal{R}^N_n$. Again, we will abbreviate notation by shortening $\overline{n}(x, \delta_N R)$ to $\overline{n}$.

Lemma 4.5.2. *Let $(\rho^N_t, t \geq 0)$ be an mSLFV process under the scaling (4.7) and (4.8). Then, for $\phi \in L^1(\mathbb{T}^d \times [0,1])$,*

$$M^N_t(\phi) := \Big\langle \rho^N_t, \phi \Big\rangle - \Big\langle \rho^N_0, \phi \Big\rangle - \int_0^t \Big[\mu \Big\langle n^N_s \lambda - \rho^N_s, \phi \Big\rangle + \Big\langle \mathcal{L}^N\Big(\rho^N_s\Big) + \mathcal{R}^N_{n^N_s}\Big(\rho^N_s\Big), \phi \Big\rangle\Big]ds$$

is a martingale with predictable variation process

$$\int_0^t \frac{N}{\delta_N^{2+d}} \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n^N_s(y)dy} \int_{B(x,\delta_N R)\times[0,1]} (\overline{n^N_s}(x,\delta_N R)+1) \Bigg[\Big\langle \rho^N_s + \frac{u}{(\overline{n^N_s}(x,\delta_N R)+1)}\mathbb{1}_{B(x,\delta_N R)}(\cdot) (\delta_{k_0}[1+\delta_N^2 r_x(\overline{n^N_s}(x,\delta_N))]\overline{n^N_s}(x,\delta_N R) - \rho^N_s), \phi\Big\rangle - \Big\langle \rho^N_s, \phi\Big\rangle\Bigg]^2 \rho^N_s(y,dk_0)dydxds.$$

Proof. We have

$$\begin{aligned}
&\lim_{\delta t\downarrow 0}\frac{1}{\delta t}\mathbb{E}\Big[\Big\langle\,\rho_{t+\delta t},\phi\,\Big\rangle-\Big\langle\,\rho_t,\phi\,\Big\rangle\Big|\rho_t=\rho\Big]\\
&=\int_{\mathbb{T}^d}\Bigg(\frac{1}{\int_{B(x,R)}n(y)dy}\int_{B(x,R)\times[0,1]}\\
&\qquad(\overline{n}+1)\Big[\Big\langle\rho+\frac{u}{(\overline{n}+1)}\mathbb{1}_{B(x,R)}(\delta_{k_0}[1+r(\overline{n})]\overline{n}-\rho),\phi\Big\rangle-\Big\langle\rho,\phi\Big\rangle\Big]\\
&\qquad\qquad\rho(y,dk_0)dy\Bigg)dx+\mu\Big\langle\lambda\cdot n-\rho,\phi\Big\rangle\\
&=u\int_{\mathbb{T}^d}\Bigg(\frac{1}{\int_{B(x,R)}n(y)dy}\int_{B(x,R)\times[0,1]}\Big\langle\mathbb{1}_{B(x,R)}(\delta_{k_0}[1+r(\overline{n})]\overline{n}-\rho),\phi\Big\rangle\rho(y,dk_0)dy\Bigg)dx\\
&\qquad+\mu\Big\langle\lambda\cdot n-\rho,\phi\Big\rangle.
\end{aligned}$$

Again we wish to apply (4.7) and (4.8). Defining the rescaled test functions $\phi_N(x)=\delta_N^d\phi(\delta_N x)$, the rescaled generator will be

$$\mathcal{G}^N F_{f,\phi}(\rho)=\frac{N}{\delta_N^2}\mathcal{G}F_{f,\phi_N}(\rho(\delta_N\cdot)).$$

Recall that we scale impact by $u^N=u/N$ and mutation rate by $\mu_N=\delta_N^2\mu/N$. The rescaled growth function was defined as $r_\cdot^N=\delta_N^2 r_{\delta_N\cdot}$. For the mutation part, we obtain

$$\begin{aligned}
&\frac{N}{\delta_N^2}\frac{\delta_N^2\mu}{N}\int_{\delta_N^{-1}\mathbb{T}^d}\Big(\lambda\cdot n(\delta_N z)-\rho(\delta_N z)\Big)\phi(\delta_N z)\delta_N^d dz\\
&=\mu\int_{\mathbb{T}^d}\Big(\lambda\cdot n(z)-\rho(z)\Big)\phi(z)dz.
\end{aligned}$$

In a similar way to our calculations for the population size, we can calculate, abbreviating $\overline{n}(x,\delta_N R)$ by $\overline{n}$,

$$\begin{aligned}
&\frac{u}{N}\frac{N}{\delta_N^{d+2}}\int_{\mathbb{T}^d}\frac{1}{\int_{B(x,\delta_N R)}n(y)dy}\int_{B(x,\delta_N R)\times[0,1]}\int_{B(x,\delta_N R)}\Bigg[[1+\delta_N^2 r_x(\overline{n})]\overline{n}\phi(z,k_0)\\
&\qquad-\int_{[0,1]}\phi(z,k)\rho(z,dk)\Bigg]dz\rho(y,dk_0)dydx\\
&=\frac{u}{\delta_N^{d+2}}\int_{\mathbb{T}^d}\frac{1}{\int_{B(x,\delta_N R)}n(y)dy}\int_{B(x,\delta_N R)\times[0,1]}\int_{B(x,\delta_N R)}\\
&\qquad\qquad\Big[[1+\delta_N^2 r_x(\overline{n})]\overline{n}\phi(z,k_0)\Big]dz\rho(y,dk_0)dydx
\end{aligned}$$

$$
\begin{aligned}
&\quad - \int_{\mathbb{T}^d} \int_{B(x,\delta_N R)} \int_{[0,1]} \phi(z,k)\rho(z,dk)dzdx \\
&= \frac{u}{\delta_N^{d+2}} \int_{\mathbb{T}^d\times[0,1]} \int_{B(y,\delta_N R)} \frac{1}{\int_{B(x,\delta_N R)} n(y)dy} \int_{B(x,\delta_N R)} \\
&\qquad\qquad [1+\delta_N^2 r_x(\overline{n})]\overline{n}\phi(z,k_0)dzdx\rho(y,dk_0)dy \\
&\quad - V_{\delta_N R} \int_{\mathbb{T}^d\times[0,1]} \phi(z,k)\rho(z,dk)dz \\
&= \frac{u}{\delta_N^{d+2}} \int_{\mathbb{T}^d\times[0,1]} \Bigg[\int_{B(z,\delta_N R)} \frac{1}{V_{\delta_N R}} \int_{B(x,\delta_N R)} [1+\delta_N^2 r_x(\overline{n})]\phi(w,k)dwdx \\
&\qquad\qquad - V_{\delta_N R}\phi(z,k) \Bigg] \rho(z,dk)dz \\
&= \Big\langle \rho(\cdot,dk), (\mathcal{L}^N + \mathcal{R}_n^N)(\phi(\cdot,k)) \Big\rangle.
\end{aligned}
$$

Additionally, together with the corresponding calculation for the population size in the proof of Lemma 4.3.1, this shows that the operators $\mathcal{L}^N$ and $\mathcal{R}_n^N$ are self-adjoint. □

4.5.2 Representations of the fluctuations

Finally, let us return to the fluctuations defined by

$$Z_t^N = \sqrt{N\delta_N^{2-d}}\Big(\rho_t^N - m_t^N\lambda\Big).$$

The process ρ_t^N and the deterministic approximation to the population size $m_t^N\lambda$ have the following representations

$$
\begin{aligned}
\langle \rho_t^N, \phi\rangle &= \langle \rho_0^N, \phi\rangle + \int_0^t \Big[\mu\langle n_s^N - \rho_s^N, \phi\rangle + \langle \rho_s^N, \mathcal{L}^N(\phi) + \mathcal{R}_{n_s^N}(\phi)\rangle\Big]ds + M_t^N(\phi), \\
\langle m_t^N\lambda, \phi\rangle &= \langle m_0^N\lambda, \phi\rangle + \int_0^t \langle m_s^N\lambda, \mathcal{L}^N(\phi) + \mathcal{R}_{m_s^N}\phi\rangle ds.
\end{aligned}
$$

The are now different options for how to summarize the resulting terms for the fluctuations. Let us define

$$\mathfrak{M}_t^N(\phi) := \sqrt{N\delta_N^{2-d}} M_t^N(\phi).$$

We will use

$$\langle Z_t^N, \phi\rangle$$

$$
\begin{aligned}
&= \sqrt{N\delta_N^{2-d}}\Bigg[\langle \rho_0^N, \phi\rangle + \int_0^t \Big[\mu\langle n_s^N\lambda - \rho_s^N, \phi\rangle + \langle \rho_s^N, \mathcal{L}^N(\phi) + \mathcal{R}^N_{n_s^N}(\phi)\rangle\Big]ds + M_t^N(\phi) \\
&\qquad - \langle m_0^N\lambda, \phi\rangle - \int_0^t \langle m_s^N\lambda, \mathcal{L}^N(\phi) + \mathcal{R}^N_{m_s^N}\phi\rangle ds\Bigg] \\
&= \langle Z_0^N, \phi\rangle + \mathfrak{M}_t^N(\phi) + \int_0^t \langle Z_s^N, \mathcal{L}^N(\phi)\rangle ds \\
&\qquad + \int_0^t \Big[- \mu\langle Z_s^N, \phi\rangle + \sqrt{N\delta_N^{2-d}}\mu\Big\langle (n_s^N - m_s^N)\lambda, \phi\Big\rangle\Big]ds \\
&\qquad + \int_0^t \Big[\langle Z_s^N, \mathcal{R}_{n_s^N}(\phi)\rangle + \sqrt{N\delta_N^{2-d}}\Big\langle m_s^N\lambda, \big(\mathcal{R}^N_{n_s^N} - \mathcal{R}^N_{m_s^N}\big)\phi\Big\rangle\Big]ds \\
&= \langle Z_0^N, \phi\rangle + \mathfrak{M}_t^N(\phi) + \int_0^t \Big\langle Z_s^N, \mathcal{L}^N(\phi) + \mathcal{R}^N_{n_s^N}(\phi) - \mu\phi\Big\rangle ds \\
&\qquad + \int_0^t \mu\Big\langle Z_t^N, \mathcal{H}\phi\Big\rangle ds \\
&\qquad + \sqrt{N\delta_N^{2-d}}uV_R\int_0^t \Big\langle m_s^N\lambda, \overline{r(\overline{n_s^N})\overline{\phi}}(\cdot, \delta_N R) - \overline{r(\overline{m_s^N})\overline{\phi}}(\cdot, \delta_N R)\Big\rangle ds \\
&= \langle Z_0^N, \phi\rangle + \mathfrak{M}_t^N(\phi) + \int_0^t \Big\langle Z_s^N, \mathcal{L}^N(\phi) + \mathcal{R}^N_{n_s^N}(\phi) + \mu(\mathcal{H}\phi - \phi)\Big\rangle ds \\
&\qquad + \sqrt{N\delta_N^{2-d}}uV_R\int_0^t \Big\langle m_s^N\lambda, \overline{[r(\overline{n_s^N}) - r(\overline{m_s^N})]\overline{\phi}}(\cdot, \delta_N R)\Big\rangle ds,
\end{aligned}
$$

where again $\mathcal{H}\phi(x,k) = \int_0^1 \phi(x,k)dk$. The last summand will make a non-trivial contribution to the central limit theorem. To see this, we apply Taylor's expansion to obtain

$$
\begin{aligned}
&\sqrt{N\delta_N^{2-d}}uV_R\int_0^t \Big\langle \overline{m_s^N}\lambda, \Big[r'(\overline{m_s^N})(\overline{n_s^N} - \overline{m_s^N}) + \mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, r)(\overline{n_s^N} - \overline{m_s^N})^2\Big]\overline{\phi}\Big\rangle ds \\
&= uV_R\int_0^t \Big\langle \overline{m_s^N}r'(\overline{m_s^N})\overline{Z}_s^N, \overline{\mathcal{H}\phi}\Big\rangle ds \\
&\qquad + uV_R\int_0^t \Big\langle \overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, r)(\overline{n_s^N} - \overline{m_s^N})\overline{Z}_s^N, \overline{\mathcal{H}\phi}\Big\rangle ds.
\end{aligned}
$$

Returning to the complete expression, this yields

$$
\begin{aligned}
&\langle Z_t^N, \phi\rangle \\
&= \langle Z_0^N, \phi\rangle + \int_0^t \Big\langle Z_s^N, \mathcal{L}^N(\phi) + \mathcal{R}^N_{n_s^N}(\phi) + \mu(\mathcal{H}\phi - \phi) + uV_R\overline{\overline{m_s^N}r'(\overline{m_s^N})\overline{\mathcal{H}\phi}}\Big\rangle ds \\
&\qquad + \mathfrak{M}_t^N(\phi) + uV_R\int_0^t \Big\langle \overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, r)(\overline{n_s^N} - \overline{m_s^N})\overline{Z}_s^N, \overline{\mathcal{H}\phi}\Big\rangle ds.
\end{aligned}
$$

From now on, we will choose the initial condition $\rho_0 = \lambda$ and $m_0 = n_0$ leading to $Z_0 = 0$ almost surely. At this stage, we are already able to read off the familiar terms of the SPDE in Theorem 4.3.3. Again, for later analysis, we would like to

eliminate the term involving $\mathcal{L}^N$ by choosing an appropriate test function. For a time-dependent test function

$$g : \{(s,t) : 0 \le s \le t \le T\} \to L^1(\mathbb{T}^d \times [0,1])$$

the representation becomes

$$\begin{aligned}
&\langle Z_t^N, g_{t,t}\rangle \\
&= \int_0^t \Big\langle Z_s^N, \partial_s g_{s,t} + \mathcal{L}^N(g) + \mathcal{R}^N_{n_s^N}(g_{s,t}) + \mu(\mathcal{H}g_{s,t} - g_{s,t}) + uV_R\overline{\overline{m_s^N} r'(\overline{m_s^N})\overline{\mathcal{H}g_{s,t}}}\Big\rangle ds \\
&\quad + \int_0^t \Big\langle\, g_{s,t}, d\mathfrak{M}_s^N \,\Big\rangle + uV_R \int_0^t \Big\langle \overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, r)(\overline{n_s^N} - \overline{m_s^N})\overline{Z}^N_s, \overline{\mathcal{H}g_{s,t}}\Big\rangle ds.
\end{aligned}$$

We now take g^N to be the solution to

$$\begin{cases}
-\partial_s g^N_{s,t} = \mathcal{L}^N(g^N_{s,t}) + \mathcal{R}^N_{m_s^N}(g^N_{s,t}) + \mu\mathcal{H}g^N_{s,t} \\
\qquad\qquad - \mu g^N_{s,t} + uV_R\overline{\overline{m_s^N} r'(\overline{m_s^N})\overline{\mathcal{H}g^N_{s,t}}}, \\
\quad g^N_{t,t} = \phi,
\end{cases}$$

where we have replaced $\mathcal{R}^N_{n_s^N}$ by $\mathcal{R}^N_{m_s^N}$. Then we get

$$\begin{aligned}
\langle Z_t^N, g^N_{t,t}\rangle = \int_0^t \Big\langle\, g^N_{s,t}, d\mathfrak{M}_s^N \,\Big\rangle + \int_0^t \Big\langle Z_s^N, \mathcal{R}_{n_s^N}(g^N_{s,t}) - \mathcal{R}_{m_s^N}(g^N_{s,t})\Big\rangle ds \\
+ uV_R \int_0^t \Big\langle \overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, r)(\overline{n_s^N} - \overline{m_s^N})\overline{Z}^N_s, \overline{\mathcal{H}g^N_{s,t}}\Big\rangle ds.
\end{aligned} \tag{4.31}$$

We need to use $\mathcal{R}^N_{n_s^N}$ instead of $\mathcal{R}^N_{m_s^N}$ in the definition of g^N, so that the stochastic integral with respect to $\mathfrak{M}^N$ is well-defined.

We can regard the martingales $\mathfrak{M}_t^N(\phi)$ as evaluations of a single martingale measure $\mathfrak{M}_t^N$. In [Wal86] a theory of stochastic integration with respect to worthy martingale measures has been developed, which we will use to prove the central limit theorem. The definition of worthy martingale measures can be found in Definition A.2.1 and Definition A.2.2.

Lemma 4.5.3. *The martingale $\mathfrak{M}_t^N(\phi) = \sqrt{N\delta_N^{2-d}} M_t^N(\phi)$ is square-integrable with*

quadratic variation

$$\begin{aligned}
&\Big\langle \mathfrak{M}^N(\phi) \Big\rangle_t \\
&= \frac{u^2}{\delta_N^{2d}} \int_0^t \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_s^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} \\
&\qquad \frac{1}{\overline{\overline{n_s^N}(x,\delta_N R)+1}} \Bigg[\int_{B(x,\delta_N R)} \Big[[1+\delta_N^2 r_x(\overline{n_s^N}(x,\delta_N R))]\overline{n_s^N}(x,\delta_N R)\phi(z,k_0) \\
&\qquad\qquad - \int_{[0,1]} \phi(z,k)\rho_s^N(z,dk) \Big] dz \Bigg]^2 \rho_s^N(y,dk_0)dydxds
\end{aligned}$$

and $\mathfrak{M}^N$ is a worthy martingale measure.

Proof. The quadratic variation is given by

$$\begin{aligned}
&\frac{N\delta_N^{2-d}N}{\delta_N^{d+2}} \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_s^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} (\overline{n_s^N}(x,\delta_N R)+1) \\
&\qquad \Bigg[\Big\langle \rho_s^N + \frac{u}{(\overline{n_s^N}(x,\delta_N R)+1)N} \mathbb{1}_{B(x,\delta_N R)}(\cdot) \\
&\qquad\qquad (\delta_{k_0}[1+\delta_N^2 r_x(\overline{n_s^N}(x,\delta_N R))]\overline{n_s^N}(x,\delta_N R) - \rho), \phi \Big\rangle \\
&\qquad\qquad - \Big\langle \rho_s^N, \phi \Big\rangle \Bigg]^2 \rho_s^N(y,dk_0)dydx \\
&= \frac{u^2}{\delta_N^{2d}} \int_0^t \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_s^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} \frac{1}{\overline{\overline{n_s^N}(x,\delta_N R)+1}} \\
&\qquad \Bigg[\int_{B(x,\delta_N R)} \Big[[1+\delta_N^2 r_x(\overline{n_s^N}(x,\delta_N R))]\overline{n_s^N}(x,\delta_N R)\phi(z,k_0) \\
&\qquad\qquad - \int_{[0,1]} \phi(z,k)\rho_s^N(z,dk) \Big] dz \Bigg]^2 \rho_s^N(y,dk_0)dydxds.
\end{aligned}$$

Using Assumption 4.2.1 a dominating measure D^N satisfying Definition A.2.2 is given by

$$\begin{aligned}
&\Big\langle D^N, \phi\otimes\phi \Big\rangle \\
&= \frac{u^2 n_{\max}^2}{\delta_N^{2d}} \int_0^t \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_{s-}^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} \\
&\qquad \Bigg[\int_{B(x,\delta_N R)} \Big(\phi(z,k_0) + \int_{[0,1]} \phi(z,k)\rho_{s-}^N(z,dk) \Big) dz \Bigg]^2 \rho_{s-}^N(y,dk_0)dydxds.
\end{aligned}$$

Here, $n_{\max}$ is the upper bound on the population size guaranteed by Assumption 4.2.1.

□

Lemma 4.5.4. *The sequence of dominating measures D^N of Lemma 4.5.3 are uniformly bounded in the sense that there exists a constant C such that for any $0 \le s \le t$ and $\phi \in \mathcal{D}(\mathbb{T}^d \times [0,1])$*

$$\left\langle D^N, \mathbb{1}_{[s,t]}\phi \otimes \phi \right\rangle \le C|t-s|\|\phi\|_2^2 \ a.s. \tag{4.32}$$

Proof. Indeed, we can bound the dominating measure by

$$\begin{aligned}
&\left\langle D^N, \phi \otimes \phi \right\rangle \\
&= \frac{u^2 n_{\max}^2}{\delta_N^{2d}} \int_0^t \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_{s-}^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} \\
&\qquad \left[\int_{B(x,\delta_N R)} \left(\phi(z,k_0) + \int_{[0,1]} \phi(z,k)\rho_{s-}^N(z,dk) \right) dz \right]^2 \rho_{s-}(y,dk_0)dydxds \\
&\le \frac{4u^2 n_{\max}^2}{\delta_N^{2d}} \int_0^t \int_{\mathbb{T}^d} \left[\int_{B(x,\delta_N R)} \sup_{k\in[0,1]} |\phi(z,k)| dz \right]^2 dxds \\
&\le 4u^2 n_{\max}^2 \|\phi\|_2^2,
\end{aligned}$$

where the last step used the identity (see [For22, Proof of Lemma 3.4])

$$\int_{(\mathbb{T}^d)^2} V_{\delta_N R}(z_1,z_2) \sup_{k\in[0,1]} |\phi(z_1,k)| \sup_{k\in[0,1]} |\phi(z_2,k)| dz_1 dz_2 \le V_{\delta_N R}^2 \|\phi\|_2^2. \qquad \square$$

4.5.3 Bound on the square of local averages of the fluctuations

We end this section with a result, which will be the crucial ingredient for proving tightness of the sequence Z^N.

Lemma 4.5.5. *Let $h : [0,1] \to \mathbb{R}$ be a measurable test function acting only on types and assume*

$$\|h\|_\infty = \sup_{k\in[0,1]} |h(k)| < \infty.$$

Then there exists a constant $K > 0$, independent of h, such that

$$\sup_{0\le t\le T} \sup_{x\in\mathbb{T}^d} \mathbb{E}\left[\left(\frac{1}{V_{\delta_N R}} \int_{B(x,\delta_N R)\times[0,1]} h(k) Z_t^N(y,dk)dy \right)^2 \right] \le K \frac{\|h\|_\infty^2}{V_{\delta_N R}}.$$

Corollary 4.5.6. *Suppose* $h : \mathbb{T}^d \times [0,1] \to \mathbb{R}$ *and assume*

$$\|h\|_2 = \left(\int_{\mathbb{T}^d} \sup_{k\in[0,1]} |h(x,k)|^2 dx \right)^{\frac{1}{2}} < \infty.$$

Then there exists a constant $K > 0$*, independent of* h*, such that*

$$\sup_{0\le t\le T} \int_{\mathbb{T}^d} \mathbb{E}\left[\left(\frac{1}{V_{\delta_N R}} \int_{B(x,\delta_N R)\times[0,1]} h(x,k) Z_t^N(y,dk)dy \right)^2 \right] dx \le K \frac{\|h\|_2^2}{V_{\delta_N R}}.$$

Proof of Lemma 4.5.5. This proof can be seen as a generalization of [FP17, Lemma 4.5] to include types and to accomodate our setting. We will be using the representation

$$\begin{aligned} \langle Z_t^N, \phi\rangle = \mathfrak{M}_t^N(\phi) + \int_0^t \Big\langle Z_s^N, \mathcal{L}^N(\phi) + \mathcal{R}_{n_s^N}^N(\phi) + \mu(\mathcal{H}\phi - \phi)\Big\rangle ds \\ + \sqrt{N\delta_N^{2-d}} u V_R \int_0^t \Big\langle m_s^N \lambda, \overline{[r(\overline{n_s^N}) - r(\overline{m_s^N})]\bar{\phi}}(\cdot, \delta_N R)\Big\rangle ds, \end{aligned} \tag{4.33}$$

which we deduced in Section 4.5.2. Ultimately, we wish to set up an application of Grönwall's inequality. We would like to substitute the solution to

$$\begin{cases} -\partial_s \zeta_{s,t}^N = \mathcal{L}^N(\zeta_{s,t}^N) - \mu\zeta_{s,t}^N + \mu\mathcal{H}\zeta_{s,t}^N, \\ \quad \zeta_{t,t}^N = \phi, \end{cases}$$

to remove the corresponding operators from the first integral. If we let G_t^N be the semigroup of the operator $\mathcal{L}^N$, one can see that the differential equation is solved by

$$\zeta_{s,t}^N = G_{t-s}^N * \Big(e^{-\mu(t-s)}\phi + \big(1 - e^{-\mu(t-s}\big)\mathcal{H}\phi \Big).$$

We intend to choose the test function $\phi^N(z,k) = h(k)\frac{1}{V_{\delta_N R}} \mathbb{1}_{|x-z|<\delta_N R}$, so that the evaluation $\langle Z_t^N, \phi\rangle$ will lead to the average in the statement of the lemma. This leads to

$$\begin{aligned} & G_{t-s}^N * \Big(e^{-\mu(t-s)}\phi^N + \big(1 - e^{-\mu(t-s}\big)\mathcal{H}\phi^N \Big)(z,k) \\ &= \left(G_{t-s}^N * \frac{1}{V_{\delta_N R}} \mathbb{1}_{|x-\cdot|<\delta_N R} \right)(z) \Big(e^{-\mu(t-s)} h(k) + \big(1 - e^{-\mu(t-s)}\big)\mathcal{H}h \Big). \end{aligned}$$

Let us define

$$h_{t-s}(k) := \Big(e^{-\mu(t-s)} h(k) + \big(1 - e^{-\mu(t-s}\big)\mathcal{H}h \Big)$$

and for $n, m \geq 0$ the quotient

$$q_z(n,m) := \begin{cases} \dfrac{r_z(n) - r_z(m)}{n-m} & n \neq m, \\ r_z'(n) & n = m. \end{cases} \tag{4.34}$$

Inserting $\zeta_{s,t}^N$ into (4.33) and using $\mathcal{R}_{m_s^N}^N(\zeta_{s,t}^N) = uV_R\overline{\overline{r(\overline{m_s^N})}\,\overline{\zeta_{s,t}^N}}(\cdot, \delta_N R)$ yields

$$\begin{aligned}\langle Z_t^N, \phi\rangle = \int_0^t \int_{[0,t]\times\mathbb{T}^d\times[0,1]} \zeta_{s,t}^N(z,k)\mathfrak{M}^N(ds\ dz\ dk) + uV_R\int_0^t \Big\langle \overline{Z_s^N}, r(\overline{n_s^N})\overline{\zeta_{s,t}^N}\Big\rangle ds \\ + uV_R\int_0^t \Big\langle \overline{m_s^N\lambda}, \overline{F_s^N} q(\overline{n_s^N}, \overline{m_s^N})\overline{\zeta_{s,t}^N}\Big\rangle ds.\end{aligned}$$

If we now choose $\phi^N(z,k) = h(k)\frac{1}{V_{\delta_N R}}\mathbb{1}_{|x-z|<\delta_N R}$, we obtain

$$\begin{aligned}\langle Z_t^N, \phi^N\rangle = \int_{[0,t]\times\mathbb{T}^d\times[0,1]} \overline{G_{t-s}^N}(x-z)h_{t-s}(k)\mathfrak{M}^N(ds\ dz\ dk) \\ + uV_R\int_0^t \Big\langle \overline{Z_s^N}, r(\overline{n_s^N})\overline{\overline{G_{t-s}^N}}(x-\cdot)h_{t-s}\rangle ds \\ + uV_R\int_0^t \Big\langle \overline{m_s^N\lambda}, \overline{F_s^N} q(\overline{n_s^N}, \overline{m_s^N})\overline{\overline{G_{t-s}^N}}(x-\cdot)h_{t-s}\Big\rangle ds.\end{aligned}$$

Applying the inequality

$$\Big(\sum_{i=1}^n a_i\Big)^2 \leq n\Big(\sum_{i=1}^n a_i^2\Big),$$

we arrive at

$$\begin{aligned}&\langle Z_t^N, \phi^N\rangle^2 \\ &\leq 3\left(\int_{[0,t]\times\mathbb{T}^d\times[0,1]} \overline{G_{t-s}^N}(x-z)h_{t-s}(k)\mathfrak{M}^N(ds\ dz\ dk)\right)^2 \\ &\quad + 3u^2V_R^2\left(\int_0^t\int_{\mathbb{T}^d\times[0,1]} \overline{Z_s^N}(z,dk)r(\overline{n_s^N}(z))\overline{\overline{G_{t-s}^N}}(x-z)h_{t-s}(k)dzds\right)^2 \\ &\quad + 3u^2V_R^2\left(\int_0^t\int_{\mathbb{T}^d} \overline{m_s^N}(z)\overline{F_s^N}(z)q(\overline{n_s^N}(z), \overline{m_s^N}(z))\overline{\overline{G_{t-s}^N}}(x-z)\mathcal{H}h\ dzds\right)^2.\end{aligned} \tag{4.35}$$

One is able to apply Jensen's inequality as $\int_{\mathbb{T}^d} G_{t-s}^N(z)dz = 1$, which we will use to treat the second and third term. For the second term we get the bound

$$3u^2V_R^2\|r\|_\infty^2 t\int_0^t\int_{\mathbb{T}^d}\left(\int_{[0,1]} h_{t-s}(k)\overline{Z_s^N}(z,dk)\right)^2 \overline{\overline{G_{t-s}^N}}(x-z)dzds.$$

In turn, using Jensen's inequality

$$\begin{aligned}
&\left(\int_{[0,1]} h_{t-s}(k)\overline{Z_s^N}(z,dk)\right)^2 \\
&= \left(e^{-\mu(t-s)}\int_{[0,1]} h(k)\overline{Z_s^N}(z,dk) + \left(1-e^{-\mu(t-s)}\right)\mathcal{H}h\cdot\overline{F_s^N}(z)\right)^2 \\
&\leq \left(\int_{[0,1]} h(k)\overline{Z_s^N}(z,dk)\right)^2 + \|h\|_\infty^2\overline{F_s^N}(z)^2,
\end{aligned}$$

where we used $\|h\|_1 \leq \|h\|_\infty$. The third term of (4.35) results in

$$3u^2V_R^2\|h\|_\infty^2 n_{\max}^2\|r'\|_\infty^2 t\int_0^t\int_{\mathbb{T}^d}\overline{F_s^N}(z)^2\overline{\overline{G_{t-s}^N}}(x-z)dzds,$$

where the square of the quotient (4.34) was bounded by $\sup_{z\in\mathbb{T}^d}\|r_z'\|_\infty^2$. Taking expectations on both sides of (4.35) and exchanging integral and expectation, we arrive at

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z_t^N}(x,dk)\right)^2\right] \\
&\leq 3\mathbb{E}\left[\left(\int_{[0,t]\times\mathbb{T}^d\times[0,1]}\overline{G_{t-s}^N}(x-z)h_{t-s}(k)\mathfrak{M}^N(ds\,dz\,dk)\right)^2\right] \\
&\quad + 3u^2V_R^2\|r\|_\infty^2 t\int_0^t\int_{\mathbb{T}^d}\mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z_s^N}(z,dk)\right)^2\right]\overline{\overline{G_{t-s}^N}}(x-z)dzds \\
&\quad + 3u^2V_R^2\|h\|_\infty^2\|r\|_\infty^2 t\int_0^t\int_{\mathbb{T}^d}\mathbb{E}\left[\overline{F_s^N}(z)^2\right]\overline{\overline{G_{t-s}^N}}(x-z)dzds \\
&\quad + 3u^2V_R^2\|h\|_\infty^2 n_{\max}^2\|r'\|_\infty^2 t\int_0^t\int_{\mathbb{T}^d}\mathbb{E}\left[\overline{F_s^N}(z)^2\right]\overline{\overline{G_{t-s}^N}}(x-z)dzds.
\end{aligned}$$

By Lemma 4.5.4, the term involving the martingale can be bounded by

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t\int_{\mathbb{T}^d\times[0,1]}\overline{G_{t-s}^N}(x-z)h_{t-s}(k)\mathfrak{M}^N(dz,dk)\right)^2\right] \\
&\leq C\int_0^t\int_{\mathbb{T}^d}\sup_{k\in[0,1]}\left|\overline{G_{t-s}^N}(x-z)h_{t-s}(k)\right|^2 dzds \\
&\leq C\|h\|_\infty^2\int_0^t\left\|\overline{G_{t-s}^N}(x-z)\right\|_2^2 ds.
\end{aligned}$$

If we let $(X_{t-s}^N)_{0\leq s\leq t}$ be a random walk started in the origin with generator $\mathcal{L}^N$, we

can write

$$\begin{aligned}\left\|\overline{G^N_{t-s}}(x-z)\right\|_2^2 &= \int_0^t \int_{\mathbb{T}^d} \mathbb{E}_0\left[\frac{1}{V_{\delta_N R}} \mathbb{1}_{|X^N_{t-s}-(x-z)|<\delta_N R}\right]^2 dz ds \\ &\leq \int_0^t \mathbb{E}_0\left[\int_{\mathbb{T}^d}\left(\frac{1}{V_{\delta_N R}} \mathbb{1}_{|X^N_{t-s}-(x-z)|<\delta_N R}\right)^2 dz\right] ds \\ &= \frac{t}{V_{\delta_N R}}.\end{aligned}$$

Summarizing, we obtain

$$\begin{aligned}&\mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z^N_t}(x,dk)\right)^2\right] \\ &\leq 3C\|h\|_\infty^2 \frac{t^2}{V_{\delta_N R}} + 3u^2 V_R^2 \|r\|_\infty^2 t \int_0^t \sup_{z\in\mathbb{T}^d} \mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z^N_s}(z,dk)\right)^2\right] ds \\ &\quad + 3u^2 V_R^2 t^2 \|h\|_\infty^2 \left(\|r\|_\infty^2 + n_{\max}^2 \|r'\|_\infty^2\right) \sup_{s\in[0,t]} \sup_{z\in\mathbb{T}^d} \mathbb{E}\left[\overline{F^N_s}(z)^2\right].\end{aligned} \tag{4.36}$$

In turn, by Lemma 4.4.3, there exists a constant $C' > 0$ such that

$$\sup_{s\in[0,t]} \sup_{z\in\mathbb{T}^d} \mathbb{E}\left[\overline{F^N_s}(z)^2\right] \leq \frac{C'}{V_{\delta_N R}}.$$

As the right-hand side of (4.36) does not depend on x, we can take a supremum on the left-hand side

$$\begin{aligned}&\sup_{x\in\mathbb{T}^d} \mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z^N_t}(x,dk)\right)^2\right] \\ &\leq \frac{\|h\|_\infty^2}{V_{\delta_N R}} 3t^2\left(C + u^2 V_R^2\left(\|r\|_\infty^2 + n_{\max}^2\|r'\|_\infty^2\right)C'\right) \\ &\qquad + 3u^2 V_R^2 \|r\|_\infty^2 t \int_0^t \sup_{x\in\mathbb{T}^d} \mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z^N_s}(x,dk)\right)^2\right] ds.\end{aligned}$$

The function

$$t \mapsto \sup_{x\in\mathbb{T}^d} \mathbb{E}\left[\left(\int_{[0,1]} h(k)\overline{Z^N_t}(x,dk)\right)^2\right]$$

is bounded as

$$\left|\int_{[0,1]} h(k)\overline{Z^N_t}(x,dk)\right| \leq 2\|h\|_\infty \sqrt{N\delta_N^{2-d}} n_{\max}.$$

Hence, by Grönwall's inequality

$$\begin{aligned}
&\sup_{x\in\mathbb{T}^d}\mathbb{E}\left[\left(\int_{[0,1]}h(k)\overline{Z_t^N}(x,dk)\right)^2\right]\\
&\leq\frac{\|h\|_\infty^2}{V_{\delta_N R}}\left[3T^2\Big(C+u^2V_R^2\Big(\|r\|_\infty^2+n_{\max}^2\|r'\|_\infty^2\Big)C'\Big)\right]e^{3u^2V_R^2\|r\|_\infty^2Tt}\\
&\leq C''\frac{\|h\|_\infty^2}{V_{\delta_N R}}.
\end{aligned}$$

□

4.6 Proof of the central limit theorem

Recall the representation (4.31)

$$\begin{aligned}
\langle Z_t^N,\phi\rangle=\int_0^t\Big\langle g_{s,t}^N,d\mathfrak{M}_s^N\Big\rangle+\int_0^t\Big\langle Z_s^N,\mathcal{R}_{n_s^N}(g_{s,t}^N)-\mathcal{R}_{m_s^N}(g_{s,t}^N)\Big\rangle ds\\
+uV_R\int_0^t\Big\langle\overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N},\overline{m_s^N},r)(\overline{n_s^N}-\overline{m_s^N})\overline{Z}_s^N,\overline{\mathcal{H}g_{s,t}^N}\Big\rangle ds.
\end{aligned}$$

The first part will be the essential one, the second and third part are error terms which will vanish. The proof classically separates into:

Tightness We show compact containment and control of the modulus of continuity of the fluctuations to apply the Aldous-Rebolledo criterion.

Uniqueness We will show convergence of the finite-dimensional distributions by proving the convergence of g^N, as well as the convergence of the martingale measures $\mathfrak{M}^N$ to a Gaussian random field.

Our strategy will be an adaptation and synthesis of the ideas of [FP17] and [For22].

4.6.1 Properties of the functions g^N

An important prerequisite is the control of the behaviour of the functions g^N. Recall that g^N is the solution to

$$\begin{cases}
-\partial_s g_{s,t}^N=\mathcal{L}^N(g_{s,t}^N)+\mathcal{R}_{m_s^N}(g_{s,t}^N)-\mu g_{s,t}^N\\
\qquad\qquad+\mu\mathcal{H}g_{s,t}^N+uV_R\overline{\overline{m_s^N}r'(\overline{m_s^N})\overline{\mathcal{H}g_{s,t}^N}},\\
\quad g_{t,t}^N=\phi,
\end{cases}$$

where $\mathcal{H}g^N = \int_0^1 g^N(\cdot, k)dk$. Unsurprisingly, the limit g is the solution to

$$\begin{cases} -\partial_s g_{s,t} = uV_R\Big(\dfrac{R^2}{d+2}\Delta + r(n_s)\Big)g_{s,t} - \mu g_{s,t} \\ \qquad\qquad + \mu\mathcal{H}g_{s,t} + uV_R n_s r'(n_s)\mathcal{H}g_{s,t}, \\ g^N_{t,t} = \phi. \end{cases}$$

Let us remind ourselves that in (4.2) we defined the L^p-norm to involve the supremum over types. The proofs of the following lemmas can be found in Appendix B.2. The first result provides uniform bounds on the L^p-norms of the functions and derivatives.

Lemma 4.6.1. *For $p \in [1,2]$ and $\beta \in \mathbb{N}^d, |\beta|_1 \leq 4$, for any $T > 0$ there exists constants $C, C' > 0$ with C being independent of ϕ such that*

$$\sup_{0\leq s\leq t\leq T}\left\|g^N_{s,t}\right\|_p \leq C\|\phi\|_p,$$
$$\sup_{0\leq s\leq t\leq T}\left\|\partial_\beta g^N_{s,t}\right\|_p \leq C'.$$

The next lemma confirms our intuition that g should be the limit of g^N as $N \to \infty$.

Lemma 4.6.2. *For any $T > 0$, there exists a constant $C > 0$, depending on ϕ, such that for $p \in [1,2]$*

$$\sup_{0\leq s\leq t\leq T}\left\|g^N_{s,t} - g_{s,t}\right\|_p \leq C\delta_N.$$

We end this interlude with a characterisation of the continuity of g^N in the second time component.

Lemma 4.6.3. *For any $T > 0$, there exists $C > 0$ depending on ϕ such that for $0 \leq s, t, t' \leq T$ and $p \in [1,2]$*

$$\left\|g^N_{s,t'} - g^N_{s,t}\right\|_p \leq C|t' - t|.$$

4.6.2 Tightness

By Theorem 4.6.8, the tightness of the processes $(Z^N)_{N\geq 1}$ can be reduced to the tightness of $((\langle Z^N_t, \phi\rangle)_{t\geq 0}, N \geq 1)$ for any $\phi \in \mathcal{D}(\mathbb{T}^d \times [0,1])$. We will conclude by Aldous' criterion [Ald78]. The bounds on the error terms we develop in the compact containment proof will also show that these terms vanish in L^1 uniformly in time as $N \to \infty$.

Compact containment and vanishing error terms

Lemma 4.6.4. *There exists a constant $C > 0$, independent of $N \geq 1$, such that for any $\phi : \mathbb{T}^d \times [0,1] \to \mathbb{R}$ with $\sup_{p\in\{1,2\}} \|\phi\|_p < 1$*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left|\left\langle Z_t^N, \phi \right\rangle\right|\right] \leq C.$$

Proof of Lemma 4.6.4. Again, for convenience, let us recall the representation from (4.31) given by

$$\begin{aligned}\langle Z_t^N, \phi\rangle = &\int_0^t \left\langle g_{s,t}^N, d\mathfrak{M}_s^N \right\rangle + \int_0^t \left\langle Z_s^N, \mathcal{R}_{n_s^N}(g_{s,t}^N) - \mathcal{R}_{m_s^N}(g_{s,t}^N)\right\rangle ds \\ &+ uV_R \int_0^t \left\langle \overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N}, \overline{m_s^N}, r)(\overline{n_s^N} - \overline{m_s^N})\overline{Z}_s^N, \overline{\mathcal{H}g_{s,t}^N}\right\rangle ds.\end{aligned}$$

For the integral with respect to the martingale measure we can use Theorem A.3.1, whose prerequisites have been verified in Lemma 4.5.4, Lemma 4.6.1 and Lemma 4.6.3. To treat the second integral, recall that the remainder in Taylor expansion was defined in Equation (4.25). We can write for the integrand of the second integral for $\psi : \mathbb{T}^d \times [0,1] \to \mathbb{R}$

$$\begin{aligned}&\left\langle Z_s^N, \overline{\left[r(\overline{n_s^N}) - r(\overline{m_s^N})\right]\overline{\psi}}\right\rangle \\ &= \left\langle \overline{Z_s^N}, \left[r(\overline{n_s^N}) - r(\overline{m_s^N})\right]\overline{\psi}\right\rangle \\ &= \int_{\mathbb{R}^d\times[0,1]} \overline{Z_s^N}(x,dk)(\overline{n_s^N}(x) - \overline{m_s^N}(x))\mathfrak{T}_1(\overline{n_s^N}(x), \overline{m_s^N}(x), r_x)\overline{\psi}(x,k)dx \\ &= \frac{1}{\sqrt{N\delta_N^{2-d}}} \int_{\mathbb{T}^d}\int_{[0,1]} \overline{\psi}(x,k)\overline{Z_s^N}(x,dk)\overline{F_s^N}(x)\mathfrak{T}_1(\overline{n_s^N}(x), \overline{m_s^N}(x), r_x)dx \\ &\leq \frac{1}{\sqrt{N\delta_N^{2-d}}}\left(\int_{\mathbb{T}^d}\left(\int_{[0,1]} \overline{\psi}(x,k)\overline{Z_s^N}(x,dk)\right)^2 dx\right)^{1/2} \\ &\qquad\qquad \left(\int_{\mathbb{T}^d} \overline{F_s^N}(x)^2\mathfrak{T}_1(\overline{n_s^N}(x), \overline{m_s^N}(x), r_x)^2 dx\right)^{1/2}.\end{aligned}$$

Taking the expectation on both sides and using the Cauchy-Schwarz inequality

again

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\Big\langle Z_s^N, \overline{\Big[r(\overline{n_s^N}) - r(\overline{m_s^N})\Big]\overline{\psi}}\Big\rangle\Big|\Big] \\
&\leq \frac{1}{\sqrt{N\delta_N^{2-d}}}\mathbb{E}\Bigg[\int_{\mathbb{T}^d}\Bigg(\int_{[0,1]}\overline{\psi}(x,k)\overline{Z_s^N}(x,dk)\Bigg)^2 dx\Bigg]^{1/2} \\
&\qquad\qquad \mathbb{E}\Bigg[\int_{\mathbb{T}^d}\overline{F_s^N}(x)^2\mathfrak{T}_1(\overline{n_s^N}(x),\overline{m_s^N}(x),r_x)^2 dx\Bigg]^{1/2} \\
&\leq \frac{1}{\sqrt{N\delta_N^{2-d}}}\frac{\sqrt{K}\|\psi\|_2}{\sqrt{V_{\delta_N R}}}\cdot\sup_{z\in\mathbb{T}^d}\|r_z'\|_\infty Vol(\mathbb{T}^d)^{1/2}\Bigg(\sup_{x\in\mathbb{T}^d}\mathbb{E}\Big[\overline{F_s^N}(x)^2\Big]\Bigg)^{1/2} \\
&\leq \Big(N^{1/2}\delta_N^{1+\frac{d}{2}}\Big)^{-1}\frac{\sqrt{KC'}}{V_R}\sup_{z\in\mathbb{T}^d}\|r_z'\|_\infty Vol(\mathbb{T}^d)^{1/2}\|\psi\|_2.
\end{aligned}$$

Here, we used Corollary 4.5.6 and Lemma 4.4.3 and denote the constants in the corresponding bounds by K and C' respectively. Returning to the integral over time

$$\begin{aligned}
&\mathbb{E}\Bigg[\sup_{0\leq t\leq T}\Big|\int_0^t\Big\langle Z_s^N, \overline{\Big[r(\overline{n_s^N}) - r(\overline{m_s^N})\Big]\overline{g_{s,t}^N}}(\cdot,\delta_N R)\Big\rangle ds\Big|\Bigg] \\
&\leq \mathbb{E}\Bigg[\int_0^T\Big|\Big\langle Z_s^N, \overline{\Big[r(\overline{n_s^N}) - r(\overline{m_s^N})\Big]\overline{g_{s,t}^N}}(\cdot,\delta_N R)\Big\rangle\Big| ds\Bigg] \\
&\leq T\Big(N^{1/2}\delta_N^{1+\frac{d}{2}}\Big)^{-1}\frac{\sqrt{KC'}}{V_R}\sup_{z\in\mathbb{T}^d}\|r_z'\|_\infty Vol(\mathbb{T}^d)^{1/2}\sup_{0\leq s\leq T}\Big\|\overline{g_{s,t}^N}\Big\|_2.
\end{aligned}$$

We assumed in (4.15) that $(\delta_N)_{N\geq 1}$ is chosen in such a way that $N^{-1/2}\delta_N^{-(1+\frac{d}{2})}$ vanishes. The norm of the average of g^N can be bounded noting $\|\overline{\psi}\|_p \leq \|\psi\|_p$ and

$$\int_{\mathbb{T}^d}\sup_{k\in[0,1]}\Big|\overline{g_{s,t}^N}\Big|^2 dx = \int_{\mathbb{T}^d}\Big(\overline{\sup_{k\in[0,1]}\Big|g_{s,t}^N\Big|}\Big)^2 dx \leq \Big\|g_{s,t}^N\Big\|_2^2,$$

which is bounded by Lemma 4.6.1.

For the integrand of the third integral we can use

$$\begin{aligned}
&\Big\langle\overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N},\overline{m_s^N},r)(\overline{n_s^N}-\overline{m_s^N})\overline{Z}_s^N, \overline{\mathcal{H}g_{s,t}^N}\Big\rangle \\
&= \frac{1}{\sqrt{N\delta_N^{2-d}}}\int_{\mathbb{T}^d}\overline{m_s^N}\mathfrak{T}_2(\overline{n_s^N}(x),\overline{m_s^N}(x),r_x)\Big[\overline{F}_s^N(x)\Big]^2\overline{\mathcal{H}g_{s,t}^N}(x)dx,
\end{aligned}$$

which can be bounded similarly to the previous term. As $\mathcal{H}g^N = \int_0^1 g^N dk$ we have no dependence on the types, which simplifies the calculation. We are thus able

to bound the expectation of the supremum over time of the three parts of the representation (4.31), which concludes the proof. □

Bound on the modulus of continuity

Lemma 4.6.5. *For any function $\phi \in \mathcal{D}(\mathbb{T}^d \times [0,1])$, sequence of stopping times $(T_N)_{N\geq 1}$ in $[0,T]$ and sequence of positive reals $(\gamma_N)_{N\geq 1}$ converging to zero,*

$$\langle Z^N_{T_N+\gamma_N}, \phi\rangle - \langle Z^N_{T_N}, \phi\rangle \xrightarrow{N\to\infty} 0$$

in probability.

Proof. We will use a strategy from [Wal86, Theorem 7.13] as in [FP17, Proposition 4.8]. We extend g^N from $0 \leq s \leq t \leq T$ to arbitrary $s,t \in [0,T]$ by defining

$$g^N_{s,t}(x,k) := g^N_{s\wedge t,t}(x,k),$$

and define

$$V^N_t = \int_0^T \langle g^N_{s,t}, d\mathfrak{M}^N_s\rangle.$$

We can then write

$$\begin{aligned}
&\int_0^{T^N+\gamma^N} \left\langle\, g^N_{s,T^N+\gamma^N}, d\mathfrak{M}^N_s \,\right\rangle - \int_0^{T^N} \left\langle\, g^N_{s,T^N}, d\mathfrak{M}^N_s \,\right\rangle \\
&= (V^N_{T^N+\gamma^N} - V^N_{T^N}) + \int_{T^N}^{T^N+\gamma^N} \left\langle\, \phi, d\mathfrak{M}^N_s \,\right\rangle,
\end{aligned}$$

where we used $g^N(x,k,s,t) = \phi$ for $s \geq t$. Thus, the representation (4.31) results in

$$\begin{aligned}
&\langle Z^N_{T^N+\gamma^N}, \phi\rangle - \langle Z^N_{T^N}, \phi\rangle \\
&= (V^N_{T^N+\gamma^N} - V^N_{T^N}) + \int_{T^N}^{T^N+\gamma^N} \left\langle\, \phi, d\mathfrak{M}^N_s \,\right\rangle \\
&\quad + uV_R \int_0^{T^N+\gamma^N} \Big\langle Z^N_s, \overline{\big[r(\overline{n^N_s}) - r(\overline{m^N_s})\big]\overline{g^N_{s,T^N+\gamma^N}}}\Big\rangle ds \\
&\quad + uV_R \int_0^{T^N+\gamma^N} \Big\langle \overline{m^N_s}\mathfrak{T}_2(\overline{n^N_s}, \overline{m^N_s}, r)(\overline{n^N_s} - \overline{m^N_s})\overline{Z}^N_s, \overline{\mathcal{H}g^N_{s,T^N+\gamma^N}}\Big\rangle ds \\
&\quad - uV_R \int_0^{T^N} \Big\langle Z^N_s, \overline{\big[r(\overline{n^N_s}) - r(\overline{m^N_s})\big]\overline{g^N_{s,T^N}}}\Big\rangle ds \\
&\quad - uV_R \int_0^{T^N} \Big\langle \overline{m^N_s}\mathfrak{T}_2(\overline{n^N_s}, \overline{m^N_s}, r)(\overline{n^N_s} - \overline{m^N_s})\overline{Z}^N_s, \overline{\mathcal{H}g^N_{s,T^N}}\Big\rangle ds.
\end{aligned}$$

Note that the four error terms converge to zero in L^1 using the same bounds as in the compact containment proof. By Lemma 4.5.4 and Lemma 4.6.3 we obtain for any $t' \geq t$

$$\begin{aligned}\mathbb{E}\Big[\big|V^N_{t'} - V^N_t\big|^2\Big] &= \mathbb{E}\left[\left(\int_0^T \int_{\mathbb{R}^d} (g^N_{s,t'}(x,k) - g^N_{s,t}(x,k)) M^N(dxdkds)\right)^2\right] \\ &\leq C \int_0^T \left\| g^N_{s,t'} - g^N_{s,t} \right\|_2^2 ds \\ &\leq C' T |t' - t|^2.\end{aligned}$$

Hence, by Kolomogorov's continuity criterion, there exists a random variable B^N such that for any $0 \leq t, t' \leq T$

$$|V^N_{t'} - V^N_t| \leq B^N |t' - t|$$

and $\mathbb{E}\Big[(B^N)^2\Big] < C''$, where the constant $C'' \geq 0$ is independent of $N \geq 1$. Therefore, we get

$$\mathbb{E}\Big[\big|V^N_{T^N+\gamma^N} - V^N_{T^N}\big|^2\Big] \leq \mathbb{E}\Big[(B^N)^2\Big]\gamma^N \leq C''\gamma^N.$$

Additionally, again by Lemma 4.5.4

$$\begin{aligned}\mathbb{E}\left[\left(\int_{T^N}^{T^N+\gamma^N} \left\langle \phi, d\mathfrak{M}^N_s \right\rangle\right)^2\right] &\leq C\mathbb{E}\left[\int_{T^N}^{T^N+\gamma^N} \|\phi\|_2^2 ds\right] \\ &\leq C\gamma^N \|\phi\|_2^2.\end{aligned}$$

Both terms converge to zero in L^2 and thus in probability. □

Lemma 4.6.4 gives, by the Markov inequality, the tightness of fixed time marginals. Together with Lemma 4.6.5 this amounts precisely to the requisites of Aldous' criterion [Ald78] and tightness follows.

4.6.3 Convergence of the martingale measures

The convergence of $(\mathfrak{M}^N)_{N\geq 1}$ is the last missing ingredient for the convergence of the finite-dimensional distributions, as we have already convergence of the integrands g^N of the stochastic integrals with respect to $\mathfrak{M}^N$.

Theorem 4.6.6 ([JS03, VIII Theorem 3.11]). *Suppose* $((X^N_t)_{t\in[0,T]}, N \geq 1)$ *is a sequence of local martingales in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ *such that*

1. *there exists* $C > 0$, *such that for any* $N \geq 1$, $\left|X_t^N - X_{t-}^N\right| \leq C$,

2. *the jumps vanish, i.e.* $\sup_{t\in[0,T]} \left|X_t^N - X_{t-}^N\right| \xrightarrow[N\to\infty]{P} 0$,

3. *there exists a continuous Gaussian martingale* $(X_t)_{t\geq 0}$ *such that for any rational time* t $\left\langle X^{N,i}, X^{N,j}\right\rangle_t \xrightarrow[N\to\infty]{} \left\langle X^i, X^j\right\rangle_t$ *in probability.*

Then X^N *converges to* X *in distribution with limit* X *in* $C(\mathbb{R}_+, \mathbb{R}^d)$.

Lemma 4.6.7. *In* $\mathbb{D}(\mathbb{R}_+, \mathcal{D}'(\mathbb{T}^d \times [0,1])$, *the sequence of martingale measures* $\mathfrak{M}^N$ *converges in distribution to a continuous Gaussian martingale limit* $\mathfrak{M}$, *such that for any* $\phi \in \mathcal{D}(\mathbb{T}^d \times [0,1])$

$$\left\langle \mathfrak{M}_t(\phi) \right\rangle = t\left\langle \mathcal{Q}, \phi \otimes \phi \right\rangle, \tag{4.37}$$

where $\mathcal{Q}$ *is given in* (4.18).

Proof. The uniform upper and lower bounds $n_{\min}, n_{\max}$ on the population size allow one to check the first two assumptions of Theorem 4.6.6 for $M_t(\phi)$ in a similar way to [For22, Lemma 3.10, fixed radius case], which also inspired the strategy of the rest of this proof. Let us recall the expression for the covariation of $\mathfrak{M}^N$ from Lemma 4.5.3:

$$\begin{aligned}
&\left\langle \mathfrak{M}^N(\phi) \right\rangle_t \\
&= \frac{u^2}{\delta_N^{2d}} \int_0^t \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_s^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} \\
&\qquad \frac{1}{\overline{\overline{n_s^N}}(x,\delta_N R) + 1} \Bigg[\int_{B(x,\delta_N R)} \Big[[1 + \delta_N^2 r_x(\overline{n_s^N}(x,\delta_N R))] \overline{n_s^N}(x,\delta_N R)\phi(z,k_0) \\
&\qquad\qquad - \int_{[0,1]} \phi(z,k)\rho_s^N(z,dk) \Big] dz \Bigg]^2 \rho_s^N(y,dk_0)dydxds
\end{aligned}$$

We start by proving that the term involving $\delta_N^2 r(\overline{n_s^N})$ does not contribute in the

limit. Bounding the term in question, we can see

$$\Bigg|\frac{u^2}{\delta_N^{2d}}\int_{\mathbb{T}^d}\frac{1}{\int_{B(x,\delta_N R)}n_s^N(y)dy}\int_{B(x,\delta_N R)\times[0,1]}\frac{1}{\overline{\overline{n_s^N}+1}}\Bigg[\int_{B(x,\delta_N R)}\delta_N^2 r(\overline{n_s^N})\overline{n_s^N}\phi(z,k_0)dz\Bigg]$$
$$\Bigg[\int_{B(x,\delta_N R)}\Big[[1+\delta_N^2 r(\overline{n_s^N})]\overline{n_s^N}\phi(z,k_0)-\int_{[0,1]}\phi(z,k)\rho_s^N(z,dk)\Big]dz\Bigg]$$
$$\rho_s^N(y,dk_0)dydx\Bigg|$$
$$\leq\frac{u^2}{\delta_N^{2d}}\int_{\mathbb{T}^d}\frac{1}{\int_{B(x,\delta_N R)}n_s^N(y)dy}\int_{B(x,\delta_N R)\times[0,1]}\frac{1}{\overline{\overline{n_s^N}+1}}\delta_N^2\|r\|_\infty n_{\max}$$
$$\int_{B(x,\delta_N R)}\sup_{k\in[0,1]}\Big|\phi(z,k)\Big|dz(2n_{\max}+\delta_N^2\|r\|_\infty)$$
$$\int_{B(x,\delta_N R)}\sup_{k\in[0,1]}\Big|\phi(z,k)\Big|dz\;\rho_s^N(y,dk_0)dydx.$$

The integrand no longer depends on k_0 anymore and the integral with respect to $\rho_s^N(y,dk_0)dy$ cancels with the average of the population size over the ball $B(x,\delta_N R)$. As before, let $V_{\delta_N R}(z_1,z_2)$ denote the volume of the intersection of two balls of radius $\delta_N R$ centreed at z_1 and z_2. Then, using the Cauchy-Schwarz inequality and $\int_{\mathbb{R}^d}V_r(x_1,x_2)dr=V_r^2$ as in Lemma 4.5.4, we obtain the upper bound

$$\begin{aligned}
&f(\delta_N)\frac{u^2}{\delta_N^{2d}}\int_{\mathbb{T}^d}\int_{B(x,\delta_N R)}\sup_{k\in[0,1]}\Big|\phi(z,k)\Big|dz\int_{B(x,\delta_N R)}\sup_{k\in[0,1]}\Big|\phi(z,k)\Big|dz\;dx\\
&\leq f(\delta_N)\frac{u^2}{\delta_N^{2d}}\int_{(\mathbb{T}^d)^2}V_{\delta_N R}(z_1,z_2)\sup_{k\in[0,1]}\Big|\phi(z_1,k)\Big|\sup_{k\in[0,1]}\Big|\phi(z_2,k)\Big|dz_1dz_2\\
&\leq f(\delta_N)u^2V_R^2\|\phi\|_2^2,
\end{aligned}$$

where we set $f(\delta_N):=\delta_N^2\|r\|_\infty n_{\max}(2n_{\max}+\delta_N^2\|r\|_\infty)$. We can therefore focus our analysis on

$$\begin{aligned}
&\frac{u^2}{\delta_N^{2d}}\int_0^t\int_{\mathbb{T}^d}\frac{1}{\int_{B(x,\delta_N R)}n_s^N(y)dy}\int_{B(x,\delta_N R)\times[0,1]}\\
&\frac{1}{\overline{\overline{n_s^N}+1}}\int_{(\mathbb{T}^d)^2}\mathbb{1}_{B(x,\delta_N R)}(z_1)\mathbb{1}_{B(x,\delta_N R)}(z_2)\Big(\overline{n_s^N}\phi(z_1,k_0)-\int_{[0,1]}\phi(z_1,k)\rho_s^N(z_1,dk)\Big)\\
&\Big(\overline{n_s^N}\phi(z_2,k_0)-\int_{[0,1]}\phi(z_2,k)\rho_s^N(z_2,dk)\Big)dz_1dz_2\rho_s^N(y,dk_0)dydxds.
\end{aligned}\tag{4.38}$$

The product of the two brackets in (4.38) will result in four terms, which we will treat separately. Introducing the average

$$\overline{\rho}(x,dk,\delta_N R)=\frac{1}{\int_{B(x,\delta_N R)}n(y)dy}\int_{B(x,\delta_N R)}\rho(y,dk)dy,$$

the first term can be written as

$$\begin{aligned}
&\frac{u^2}{\delta_N^{2d}} \int_{\mathbb{T}^d} \frac{1}{\int_{B(x,\delta_N R)} n_s^N(y)dy} \int_{B(x,\delta_N R)\times[0,1]} \\
&\qquad \frac{1}{\overline{n_s^N}+1} \int_{(\mathbb{T}^d)^2} \mathbb{1}_{B(x,\delta_N R)}(z_1)\mathbb{1}_{B(x,\delta_N R)}(z_2)\overline{n_s^N}(x)\phi(z_1,k_0)\overline{n_s^N}(x)\phi(z_2,k_0) \\
&\qquad\qquad dz_1 dz_2 \rho_s^N(y,dk_0)dydx \\
&= u^2 V_R^2 \left\langle \overline{\rho_s^N}(\cdot,\delta_N R), \frac{\overline{n_s^N}^2}{\overline{n_s^N}+1}\overline{\phi}(\cdot,\delta_N R)^2 \right\rangle.
\end{aligned}$$

The second term is given by

$$\begin{aligned}
&\frac{u^2}{\delta_N^{2d}} \int_{\mathbb{T}^d} \frac{1}{\overline{n_s^N}+1} \int_{(\mathbb{T}^d)^2} \mathbb{1}_{B(x,\delta_N R)}(z_1)\mathbb{1}_{B(x,\delta_N R)}(z_2) \\
&\qquad \times \int_{[0,1]} \overline{n_s^N}(x)\phi(z_1,k_0)\overline{\rho_s^N}(x,dk_0,\delta_N R) \int_{[0,1]} \phi(z_2,k)\rho_s^N(z_2,dk)dz_1 dz_2 dx.
\end{aligned} \tag{4.39}$$

Replacing $\phi(z_i,k)$ by $\phi(x,k)$ in the second term results in

$$\begin{aligned}
&u^2 V_R^2 \int_{\mathbb{T}^d\times[0,1]^2} \frac{1}{\overline{n_s^N}(x)+1} \overline{n_s^N}(x)^2 \phi(x,k_0)\phi(x,k)\overline{\rho_s^N}(x,dk_0,\delta_N R)\overline{\rho_s^N}(x,dk,\delta_N R)dx \\
&= u^2 V_R^2 \left\langle \overline{\rho_s^N}(\cdot,\delta_N R)\otimes\overline{\rho_s^N}(\cdot,\delta_N R), \frac{\overline{n_s^N}^2}{\overline{n_s^N}+1}\phi\otimes\phi \right\rangle.
\end{aligned} \tag{4.40}$$

We can control the error made by the replacement as

$$\left| \int_{B(x,\delta_N R)\times[0,1]} \Big(\phi(z_2,k)-\phi(x,k)\Big)\rho_s^N(z_2,dk)dz_2 \right| \le V_{\delta_N R} n_{\max}\cdot\delta_N R \max_{|\kappa|=1}\|\partial_\kappa\phi\|_\infty.$$

To bound the distance between the original second term (4.39) and the modified one (4.40), we can add and subtract an intermediate term with only one of the replacements. This allows us to bound the distance by

$$\begin{aligned}
&2\frac{u^2}{\delta_N^{2d}}\cdot n_{\max}V_{\delta_N R}\|\phi\|_1 \cdot V_{\delta_N R} n_{\max}\delta_N R \max_{|\kappa|=1}\|\partial_\kappa\phi\|_\infty \\
&= \delta_N 2u^2 n_{\max}^2 R V_R^2 \max_{|\kappa|=1}\|\partial_\kappa\phi\|_\infty\|\phi\|_1.
\end{aligned}$$

The third term is of the same form as the second, and the fourth term can be seen to equal

$$u^2 V_R^2 \left\langle \overline{\rho_s^N}(\cdot,\delta_N R)\otimes\overline{\rho_s^N}(\cdot,\delta_N R), \frac{\overline{n_s^N}^2}{\overline{n_s^N}+1}\phi\otimes\phi \right\rangle.$$

We conclude

$$\begin{aligned}
&\left| \left\langle \mathfrak{M}_t^N(\phi) \right\rangle - u^2 V_R^2 \int_0^t \left[\left\langle \overline{\rho_s^N}(\cdot, \delta_N R), \frac{\overline{n_s^N}^2}{\overline{n_s^N}+1} \overline{\phi}(\cdot, \delta_N R)^2 \right\rangle \right. \right. \\
&\qquad \left. \left. - \left\langle \overline{\rho_s^N}(\cdot, \delta_N R) \otimes \overline{\rho_s^N}(\cdot, \delta_N R), \frac{\overline{n_s^N}^2}{\overline{n_s^N}+1} \phi \otimes \phi \right\rangle \right] ds \right| \\
&\leq \delta_N 4 t u^2 n_{\max}^2 R V_R^2 \max_{|\kappa|=1} \|\partial_\kappa \phi\|_\infty \|\phi\|_1.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\left| \frac{\overline{n_s^N}^2}{\overline{n_s^N}+1} - \frac{\overline{m_s^N}^2}{\overline{m_s^N}+1} \right| \\
&\leq \left| \overline{n_s^N}^2 (\overline{m_s^N}+1) - \overline{m_s^N}^2 (\overline{n_s^N}+1) \pm \overline{n_s^N}^2 (\overline{n_s^N}+1) \right| \\
&= \left| \overline{n_s^N}^2 (\overline{m_s^N} - \overline{n_s^N}) + (\overline{n_s^N} - \overline{m_s^N})(\overline{n_s^N} + \overline{m_s^N})(\overline{n_s^N}+1) \right| \\
&\leq n_{\max}^2 \left| \overline{n_s^N} - \overline{m_s^N} \right| + \left| \overline{n_s^N} - \overline{m_s^N} \right| 2 n_{\max}(n_{\max}+1) \\
&\leq 3 n_{\max}(n_{\max}+1) \left| \overline{n_s^N} - \overline{m_s^N} \right|.
\end{aligned}$$

Setting $C = 3n_{\max}(n_{\max}+1)$, we obtain the bound

$$\begin{aligned}
&\mathbb{E}\left| \left\langle \overline{\rho_s^N}(\cdot, \delta_N R), \left(\frac{\overline{n_s^N}^2}{\overline{n_s^N}+1} - \frac{\overline{m_s^N}^2}{\overline{m_s^N}+1} \right) \overline{\phi}(\cdot, \delta_N R)^2 \right\rangle \right| \\
&\leq \frac{C}{\sqrt{N \delta_N^{2-d}}} \mathbb{E}\left[\int_{\mathbb{T}^d} \overline{F_s^N}(x, \delta_N R)^2 dx \right]^{1/2} \\
&\qquad\qquad \mathbb{E}\left[\int_{\mathbb{T}^d} \left(\int_{[0,1]} \left| \overline{\phi}(x, \delta_N R) \right|^2 \overline{\rho_s^N}(x, dk, \delta_N R) \right)^2 dx \right]^{1/2} \\
&\leq \left(N \delta_N^2 \right)^{-\frac{1}{2}} C \sqrt{\frac{C'}{V_R}} \|\phi\|_4^2,
\end{aligned}$$

which converges to zero by assumption (4.15). The same bound also holds for

$$\mathbb{E}\left| \left\langle \overline{\rho_s^N}(\cdot, \delta_N R) \otimes \overline{\rho_s^N}(\cdot, \delta_N R), \left(\frac{\overline{n_s^N}^2}{\overline{n_s^N}+1} - \frac{\overline{m_s^N}^2}{\overline{m_s^N}+1} \right) \phi \otimes \phi \right\rangle \right|.$$

As $N \to \infty$, the convergence $\rho_s^N \to n_s \lambda$ in probability is implied by the central limit theorem. The same convergence follows for the averages $\overline{\rho_s^N}(\cdot, \delta_N R)$ by Proposition 4.3.2 for the averaged population size $\overline{n_s^N}(\cdot, \delta_N R)$. As we are evaluating

against deterministic test functions, the complete expression

$$u^2 V_R^2 \int_0^t \left[\left\langle \overline{\rho_s^N}(\cdot, \delta_N R), \frac{\overline{m_s^N}^2}{\overline{m_s^N} + 1} \overline{\phi}(\cdot, \delta_N R)^2 \right\rangle - \left\langle \overline{\rho_s^N}(\cdot, \delta_N R) \otimes \overline{\rho_s^N}(\cdot, \delta_N R), \frac{\overline{m_s^N}^2}{\overline{m_s^N} + 1} \phi \otimes \phi \right\rangle \right] ds$$

and thus $\langle \mathfrak{M}^N(\phi) \rangle_t$ converges to (4.37). We can now proceed as in [FP17, Section 4.5]. By Theorem 4.6.6, the sequence $(\mathfrak{M}_t^N(\phi))_{t \geq 0}$ converges in distribution to $(\mathfrak{M}_t(\phi))_{t \geq 0}$ in $\mathbb{D}(\mathbb{R}_+, \mathcal{D}'(\mathbb{T}^d \times [0,1]))$ and is in particular tight. Additionally, Theorem 4.6.6 yields the convergence of finite-dimensional distributions of $\mathfrak{M}^N$, as $\langle \mathfrak{M}^N(\phi_i), \mathfrak{M}^N(\phi_j) \rangle_t$ is uniquely described by $\langle \mathfrak{M}^N(\phi_i + \phi_j) \rangle_t$ and $\langle \mathfrak{M}^N(\phi_i - \phi_j) \rangle_t$ using the polarization identity. An application of Theorem 4.6.8 finishes this proof. □

4.6.4 Conclusion of the proof of the central limit theorem

In the previous subsections we have collected all the ingredients for a proof of Theorem 4.3.3. Tightness of the sequences $((\langle Z_t^N, \phi \rangle)_{t \geq 0}, N \geq 1)$ has been proven in Section 4.6.2. The convergence of finite-dimensional distributions follows from an application of Theorem A.3.3. Its prerequisites are the convergence properties of the integrands g^N (Lemma 4.6.1, Lemma 4.6.2) and the convergence of the martingale measures Lemma 4.4.1. Theorem 4.3.3 now follows from an application of the following theorem.

Theorem 4.6.8 ([Wal86, Theorem 6.15])**.** *Suppose for a sequence $((X_t^N)_{t \geq 0}, N \geq 0)$ of distribution-valued stochastic processes in $D([0,T], \mathcal{D}'(\mathbb{T}^d))$*

1. *for every test function $\phi \in \mathcal{D}(\mathbb{T}^d)$ the sequence $((X_t^N(\phi))_{t \geq 0}, N \geq 0)$ is tight,*
2. *for all test functions $\phi_1, \ldots, \phi_k \in \mathcal{D}(\mathbb{T}^d)$ and times $t_1, \ldots, t_k \in [0,T]$, the vector $(X_{t_1}^N(\phi_1), \ldots, X_{t_k}^N(\phi_k))$ converges weakly in distribution.*

Then there exists limit $X = (X_t)_{t \geq 0}$ in distribution of X^N in $\mathbb{D}([0,T], \mathcal{D}'(\mathbb{T}^d))$.

4.7 Proof of the Wright-Malécot formula

We can describe the probability of identity by descent for two probability densities functions $\psi_1, \psi_2 \in \mathcal{D}(\mathbb{T}^d)$ by

$$N\eta_N P_t^N(\phi,\psi) = \mathbb{E}\left[\frac{\langle Z_t^N \otimes Z_t^N, \mathbb{1}_{\diagdown}\psi_1\otimes\psi_2\rangle}{\langle n_t^N,\psi_1\rangle\langle n_t^N,\psi_2\rangle}\right].$$

A potential exchange of the limit $N\to\infty$ and the expectation was discussed in Remark 4.3.8. Here, we will compute the limiting Wright-Malécot formula

$$\mathbb{E}\left[\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown}\psi_1\otimes\psi_2\rangle}{\langle n_t,\psi_1\rangle\langle n_t,\psi_2\rangle}\right] \tag{4.41}$$

given in Theorem 4.3.6.

Proof of Theorem 4.3.6. For $\phi\in\mathcal{D}(\mathbb{T}^d\times[0,1])$, the mild solution to the SPDE (4.17) can be written as

$$\langle Z_t,\phi\rangle = \int_{[0,t]\times\mathbb{T}^d\times[0,1]} g_{s,t}(z,k)\mathfrak{M}(ds\,dz\,dk),$$

where g is the solution to

$$\begin{cases} -\partial_s g_{s,t} = uV_R\Big(\dfrac{R^2}{d+2}\Delta + r(n_s)\Big)g_{s,t} - \mu g_{s,t} \\ \qquad\qquad + \mu\mathcal{H}g_{s,t} + uV_R n_s r'(n_s)\mathcal{H}g_{s,t}, \\ g_{t,t} = \phi. \end{cases}$$

As $\mathfrak{M}$ is Gaussian, $\langle Z_t,\phi\rangle$ is also Gaussian with variance

$$\mathbb{E}\Big[\Big\langle Z_t\otimes Z_t, \phi_1\otimes\phi_2\Big\rangle\Big] = \int_0^t\int_{(\mathbb{T}^d)^2}\int_{[0,1]^2} f_{s,t}(x_1,k_1,x_2,k_2)\mathcal{Q}_s(dx_1dk_1dx_2dk_2)ds, \tag{4.42}$$

for the covariation measure $\mathcal{Q}_s$ given in (4.18). Here,

$$(g_{s,t}\otimes g_{s,t})\Big(\phi_1\otimes\phi_2\Big)(x_1,k_1,x_2,k_2) := f_{s,t}(x_1,k_1,x_2,k_2)$$

is defined as the solution $f_{s,t}$ to

$$\begin{cases} -\partial_s f_{s,t} = uV_R\Big(\dfrac{R^2}{d+2}(\Delta\otimes Id + Id\otimes\Delta) + r_{x_1}(n_s(x_1)) + r_{x_2}(n_s(x_2))\Big)f_{s,t} \\ \qquad + \mu\Big((\mathcal{H}-Id)\otimes Id\Big)f_{s,t} + \mu\Big(Id\otimes(\mathcal{H}-Id)\Big)f_{s,t} \\ \qquad + uV_R n_s(x_1)r'_{x_1}(n_s(x_1))\Big(\mathcal{H}\otimes Id\Big)f_{s,t} \\ \qquad + uV_R n_s(x_2)r'_{x_2}(n_s(x_2))\Big(Id\otimes\mathcal{H}\Big)f_{s,t}, \\ f_{t,t} = \phi_1\otimes\phi_2. \end{cases}$$

The operators $\Delta \otimes Id$ and $\mathcal{H} \otimes Id$ act on the first space and type parameter - correspondingly $Id \otimes \Delta$ and $Id \otimes \mathcal{H}$ act on the second. Using linearity and approximating sequences, we can generalise the expression (4.42) to functions $\Phi \in \mathcal{D}\Big((\mathbb{T}^d \times [0,1])^2\Big)$. We obtain

$$\mathbb{E}\Big[\Big\langle Z_t \otimes Z_t, \Phi \Big\rangle\Big] = \int_0^t \int_{(\mathbb{T}^d)^2} \int_{[0,1]^2} f_{s,t}(x_1, k_1, x_2, k_2) \mathcal{Q}_s(dx_1 dk_1 dx_2 dk_2) ds,$$

where $\mathcal{Q}$ is now defined as

$$\begin{aligned} &\Big\langle \mathcal{Q}_s, \Phi \Big\rangle \\ &= u^2 V_R^2 \int_{\mathbb{T}^d} \frac{n_s(z)^2}{n_s(z)+1} \left(\int_{[0,1]} \Phi(z,k,z,k) dk - \int_{[0,1]^2} \Phi(z,k_1,z,k_2) dk_1 dk_2 \right) dz. \end{aligned} \tag{4.43}$$

Let us now return to the quantity (4.41). Again we define $\vartheta_i = \frac{n_t \psi_i}{\langle n_t, \psi_i \rangle}$, where ψ_i only depends on space, and set $\Phi = \mathbb{1}_{\diagdown} \psi_1 \otimes \psi_2$. We get

$$\begin{aligned} &\mathbb{E}\left[\frac{\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown} \psi_1 \otimes \psi_2 \rangle}{\langle n_t, \psi_1 \rangle \langle n_t, \psi_2 \rangle} \right] \\ &= \mathbb{E}\left[\left\langle Z_t \otimes Z_t, \mathbb{1}_{\diagdown} \frac{\vartheta_1}{n_t} \otimes \frac{\vartheta_2}{n_t} \right\rangle \right] \\ &= \int_0^t \int_{(\mathbb{T}^d)^2} \int_{[0,1]^2} (g_{s,t} \otimes g_{s,t}) \left(\mathbb{1}_{\diagdown} \frac{\vartheta_1}{n_t} \otimes \frac{\vartheta_2}{n_t} \right) \mathcal{Q}_s(dx_1 dk_1 dx_2 dk_2) ds. \end{aligned}$$

We can show with a calculation, which we provide in Lemma 4.7.1 below, that this is equal to

$$\int_0^t \int_{(\mathbb{T}^d)^2} \int_{[0,1]^2} \mathbb{1}_{\diagdown}(k_1, k_2) P_{s,t}\left(\frac{\vartheta_1}{n_t}\right)(x_1) P_{s,t}\left(\frac{\vartheta_2}{n_t}\right)(x_2) Q_s(dx_1 dk_1 dx_2 dk_2) ds, \tag{4.44}$$

where $P_{s,t}\vartheta$ solves, for $\vartheta \in \mathcal{D}(\mathbb{T}^d)$,

$$\begin{cases} -\dfrac{d}{ds} P_{s,t}\vartheta = \dfrac{u V_R R^2}{d+2} \Delta P_{s,t}\vartheta + u V_R r(n_s) P_{s,t}\vartheta - \mu P_{s,t}\vartheta, \\ \qquad P_{t,t}\vartheta = \vartheta. \end{cases} \tag{4.45}$$

The indicator $\mathbb{1}_{\diagdown}$ was defined as $\mathbb{1}_{k_1 = k_2}$ and hence

$$\int_{[0,1]} \mathbb{1}_{\diagdown}(k,k) dk = 1, \qquad \int_{[0,1]^2} \mathbb{1}_{\diagdown}(k_1, k_2) dk_1 dk_2 = 0.$$

Therefore, under $\Phi = \mathbb{1}_{\diagdown} \psi_1 \otimes \psi_2$ the negative term of (4.43) vanishes and (4.44) is equal to

$$u^2 V_R^2 \int_0^t \int_{\mathbb{T}^d} \frac{n_s(z)^2}{n_s(z)+1} P_{s,t}\left(\frac{\vartheta_1}{n_t}\right)(z) P_{s,t}\left(\frac{\vartheta_2}{n_t}\right)(z) dz ds. \tag{4.46}$$

Similar to Doob's h-transform we define

$$\tilde{P}_{s,t}\vartheta := n_s P_{s,t}\left(\frac{\vartheta}{n_t}\right).$$

We will see in Lemma 4.7.2 below that $\tilde{P}_{s,t}\vartheta$ solves the equation

$$\begin{cases} -\dfrac{d}{ds}\tilde{P}_{s,t}\vartheta = A_s^*\tilde{P}_{s,t}\vartheta - \mu\tilde{P}_{s,t}\vartheta, \\ \qquad \tilde{P}_{t,t} = \vartheta, \end{cases} \tag{4.47}$$

where the operator A_s is defined as

$$A_s\vartheta = \frac{uV_R R^2}{d+2}\Big(\Delta\vartheta + 2\frac{\nabla n_s}{n_s}\cdot\nabla\vartheta\Big).$$

Finally, the solution $\tilde{P}_{s,t}\vartheta$ to (4.47) can be written as $e^{-\mu(t-s)}G_{s,t}\vartheta$, where $G_{s,t}\vartheta$ solves

$$\begin{cases} -\dfrac{d}{ds}G_{s,t}\vartheta = A_s^* G_{s,t}\vartheta, \\ \qquad G_{t,t}\vartheta = \vartheta. \end{cases}$$

Substituting $e^{-\mu(t-s)}G_{s,t}\vartheta$ in (4.46) results in Theorem 4.3.6. □

We end this section with the two lemmas we used in the proof of Theorem 4.3.6.

Lemma 4.7.1. *For $\vartheta_1, \vartheta_2 \in \mathcal{D}(\mathbb{T}^d)$, the following equality holds*

$$(g_{s,t}\otimes g_{s,t})\Big(\mathbb{1}_{\diagdown}\vartheta_1\otimes\vartheta_2\Big) = \mathbb{1}_{\diagdown}P_{s,t}\Big(\vartheta_1\Big)P_{s,t}\Big(\vartheta_2\Big),$$

where $P_{s,t}\vartheta$ is defined as the solution to

$$\begin{cases} -\dfrac{d}{ds}P_{s,t}\vartheta = \dfrac{uV_R R^2}{d+2}\Delta P_{s,t}\vartheta + uV_R r(n_s)P_{s,t}\vartheta - \mu P_{s,t}\phi, \\ \qquad P_{t,t}\vartheta = \vartheta. \end{cases} \tag{4.48}$$

Proof. Recall that

$$(g_{s,t}\otimes g_{s,t})\Big(\mathbb{1}_{\diagdown}\vartheta_1\otimes\vartheta_2\Big) := f_{s,t}(x_1,k_1,x_2,k_2),$$

where $f_{s,t}$ satisfies the differential equation

$$\begin{cases} -\dfrac{d}{ds}f_{s,t} = uV_R\Big(\dfrac{R^2}{d+2}(\Delta\otimes Id + Id\otimes\Delta) + r_{x_1}(n_s(x_1)) + r_{x_2}(n_s(x_2))\Big)f_{s,t} \\ \qquad\qquad + \mu\Big((\mathcal{H}-Id)\otimes Id\Big)f_{s,t} + \mu\Big(Id\otimes(\mathcal{H}-Id)\Big)f_{s,t} \\ \qquad\qquad + uV_R n_s(x_1)r'_{x_1}(n_s(x_1))\Big(\mathcal{H}\otimes Id\Big)f_{s,t} \\ \qquad\qquad + uV_R n_s(x_2)r'_{x_2}(n_s(x_2))\Big(Id\otimes\mathcal{H}\Big)f_{s,t}, \\ \qquad f_{t,t} = \mathbb{1}_{\diagdown}\vartheta_1\otimes\vartheta_2. \end{cases} \tag{4.49}$$

We aim to show

$$(\mathcal{H} \otimes Id) f_{s,t} = (Id \otimes \mathcal{H}) f_{s,t} = 0.$$

Abbreviating $(\mathcal{H} \otimes Id) f_{s,t}$ by $\tilde{f}_{s,t}$, its derivative and initial condition are

$$\begin{cases} -\frac{d}{ds}\tilde{f}_{s,t} = uV_R\Big(\frac{R^2}{d+2}(\Delta \otimes Id + Id \otimes \Delta) + r_{x_1}(n_s(x_1)) + r_{x_2}(n_s(x_2))\Big)\tilde{f}(s,t) \\ \qquad\qquad + \mu\Big(Id \otimes (\mathcal{H} - Id)\Big)\tilde{f}_{s,t} \\ \qquad\qquad + uV_R n_s(x_1) r'_{x_1}(n_s(x_1))\tilde{f}_{s,t} + uV_R n_s(x_2) r'_{x_2}(n_s(x_2))\Big(Id \otimes \mathcal{H}\Big)\tilde{f}_{s,t}, \\ \qquad \tilde{f}_{t,t} = (\mathcal{H} \otimes Id)\Big(\mathbb{1}_{\diagdown}\vartheta_1 \otimes \vartheta_2\Big) = 0. \end{cases}$$

For the initial condition we used $\mathcal{H}(\mathbb{1}_{\diagdown}) = \int_0^1 \mathbb{1}(k_1, k_2)dk_1 = 0$. Indeed, as the initial condition is zero, we obtain that $\tilde{f}_{s,t} = 0$. Returning to (4.49), this shows

$$\begin{cases} -\frac{d}{ds} f_{s,t} = uV_R\Big(\frac{R^2}{d+2}(\Delta \otimes Id + Id \otimes \Delta) + r_{x_1}(n_s(x_1)) + r_{x_2}(n_s(x_2))\Big) f_{s,t} \\ \qquad\qquad - 2\mu f_{s,t}, \\ \qquad f_{t,t} = \mathbb{1}_{\diagdown}\vartheta_1 \otimes \vartheta_2. \end{cases}$$

As the operator only depends on the spatial components, we can now write

$$\begin{aligned} (g_{s,t} \otimes g_{s,t})\Big(\mathbb{1}_{\diagdown}\vartheta_1 \otimes \vartheta_2\Big) &= (P_{s,t} \otimes P_{s,t})\Big(\mathbb{1}_{\diagdown}\vartheta_1 \otimes \vartheta_2\Big) \\ &= \mathbb{1}_{\diagdown}(P_{s,t} \otimes P_{s,t})\Big(\vartheta_1 \otimes \vartheta_2\Big) \\ &= \mathbb{1}_{\diagdown} P_{s,t}\Big(\vartheta_1\Big) P_{s,t}\Big(\vartheta_2\Big). \end{aligned}$$

□

Lemma 4.7.2. *For any $\vartheta \in \mathcal{D}(\mathbb{T}^d)$, define*

$$\tilde{P}_{s,t}\vartheta := n_s P_{s,t}\left(\frac{\vartheta}{n_t}\right),$$

where $(n_t)_{t\geq 0}$ solves the differential equation Proposition 4.3.2. Then $\tilde{P}_{s,t}\vartheta$ solves the differential equation

$$\begin{cases} -\frac{d}{ds}\tilde{P}_{s,t}\vartheta = A_s^*\tilde{P}_{s,t}\vartheta - \mu\tilde{P}_{s,t}\vartheta, \\ \qquad \tilde{P}_{t,t} = \vartheta, \end{cases}$$

where the operator A_s is defined as

$$A_s\phi = \frac{uV_R R^2}{d+2}\Big(\Delta\vartheta + 2\frac{\nabla n_s}{n_s} \cdot \nabla\vartheta\Big).$$

Proof. Integrating by parts,

$$\begin{aligned}\int \frac{\partial_x n}{n}\partial_x\vartheta\cdot\phi &= -\int \vartheta\cdot\partial_x\Big(\frac{\partial_x n}{n}\cdot\phi\Big)\\ &= -\int\vartheta\Big(\frac{\partial_{xx}n}{n}-\frac{(\partial_x n)^2}{n^2}\Big)\phi-\int\vartheta\frac{\partial_x n}{n}\partial_x\phi\end{aligned}$$

the adjoint of A will be of form

$$A^*\vartheta=\frac{uV_RR^2}{d+2}\Big(\Delta\vartheta-2\frac{\Delta n}{n}\vartheta+2\frac{\|\nabla n\|^2}{n^2}\vartheta-2\frac{\nabla n}{n}\cdot\nabla\vartheta\Big).$$

Recall that $\partial_t n_t=\frac{uV_RR^2}{d+2}\Delta n_t+uV_Rr(n_t)n_t$ and

$$\frac{d}{ds}P_{s,t}\vartheta=-\Big(\frac{uV_RR^2}{d+2}\Delta P_{s,t}\vartheta+uV_Rr(n_s)P_{s,t}\vartheta-\mu P_{s,t}\vartheta\Big).$$

Then we can calculate

$$\begin{aligned}\frac{d}{ds}\tilde P_{s,t}\vartheta &= \Big(\frac{d}{ds}n_s\Big)P_{s,t}\left(\frac{\vartheta}{n_t}\right)+n_s\frac{d}{ds}P_{s,t}\left(\frac{\vartheta}{n_t}\right)\\ &= \Big(\frac{uV_RR^2}{d+2}\Delta n_s+uV_Rr(n_s)n_s\Big)\frac{1}{n_s}\tilde P_{s,t}\vartheta\\ &\quad -n_s\Big(\frac{uV_RR^2}{d+2}\Delta+uV_Rr(n_s)-\mu\Big)\left(\frac{1}{n_s}\tilde P_{s,t}\vartheta\right)\\ &= \frac{uV_RR^2}{d+2}\left[\frac{\Delta n_s}{n_s}\cdot\tilde P_{s,t}\vartheta-n_s\Delta\left(\frac{1}{n_s}\tilde P_{s,t}\vartheta\right)\right]+\mu\tilde P_{s,t}\vartheta.\end{aligned}$$

To calculate $\Delta\left(\frac{1}{n_s}\tilde P_{s,t}\vartheta\right)$ note that

$$\begin{aligned}\partial_x\left(\frac{\phi}{n_s}\right) &= \frac{\partial_x\phi}{n_s}-\frac{\partial_x n_s}{n_s^2}\cdot\phi,\\ \partial_{xx}\left(\frac{\phi}{n_s}\right) &= \frac{\partial_{xx}\phi}{n_s}-\frac{\partial_x n_s}{n_s^2}\partial_x\phi-\frac{\partial_{xx}n_s}{n_s^2}\phi-\frac{\partial_x n_s}{n_s^2}\partial_x\phi+2\frac{(\partial_x n_s)^2}{n_s^3}\phi\\ &= \frac{\partial_{xx}\phi}{n_s}-2\frac{\partial_x n_s}{n_s^2}\partial_x\phi-\frac{\partial_{xx}n_s}{n_s^2}\phi+2\frac{(\partial_x n_s)^2}{n_s^3}\phi.\end{aligned}$$

This implies first

$$n_s\partial_{xx}\Big(\frac{1}{n_s}\tilde P_{s,t}\vartheta\Big)=\partial_{xx}\tilde P_{s,t}\vartheta-2\frac{\partial_x n_s}{n_s}\partial_x\tilde P_{s,t}\vartheta-\frac{\partial_{xx}n_s}{n_s}\tilde P_{s,t}\vartheta+2\frac{(\partial_x n_s)^2}{n_s^2}\tilde P_{s,t}\vartheta$$

and in turn

$$n_s\Delta\Big(\frac{1}{n_s}\tilde P_{s,t}\vartheta\Big)=\Delta\tilde P_{s,t}\vartheta-2\frac{\nabla n_s}{n_s}\nabla\tilde P_{s,t}\vartheta-\frac{\Delta n_s}{n_s}\tilde P_{s,t}\vartheta+2\frac{\|\nabla n\|^2}{n_s^2}\tilde P_{s,t}\vartheta.$$

In total, we can conclude

$$
\begin{aligned}
&\frac{d}{ds}\tilde{P}_{s,t}\vartheta \\
&= \frac{uV_R R^2}{d+2}\left[\frac{\Delta n_s}{n_s}\cdot\tilde{P}_{s,t}\vartheta - n_s\Delta\left(\frac{1}{n_s}\tilde{P}_{s,t}\vartheta\right)\vartheta\right] + \mu\tilde{P}_{s,t}\vartheta \\
&= \frac{uV_R R^2}{d+2}\left[\frac{\Delta n_s}{n_s}\cdot\tilde{P}_{s,t}\vartheta - \Delta\tilde{P}_{s,t}\vartheta + 2\frac{\nabla n_s}{n_s}\nabla\tilde{P}_{s,t}\vartheta + \frac{\Delta n_s}{n_s}\tilde{P}_{s,t}\vartheta - 2\frac{\|\nabla n\|^2}{n_s^2}\tilde{P}_{s,t}\vartheta\right] \\
&\qquad + \mu\tilde{P}_{s,t}\vartheta \\
&= \frac{uV_R R^2}{d+2}\left[-\Delta\tilde{P}_{s,t}\vartheta + 2\frac{\nabla n_s}{n_s}\nabla\tilde{P}_{s,t}\vartheta + 2\frac{\Delta n_s}{n_s}\tilde{P}_{s,t}\vartheta - 2\frac{\|\nabla n\|^2}{n_s^2}\tilde{P}_{s,t}\vartheta\right] + \mu\tilde{P}_{s,t}\vartheta \\
&= -(A^* - \mu)\tilde{P}_{s,t}\vartheta. \qquad \square
\end{aligned}
$$

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Central limit theorems describing isolation by distance under varying population size
Publication Status	□Published □ Accepted for Publication □Submitted for Publication ■Unpublished and unsubmitted work written in a manuscript style
Publication Details	arxiv.org/abs/2407.10211

Student Confirmation

Student Name:	Bastian Wiederhold		
Contribution to the Paper	I wrote most of the manuscript. Raphaël Forien contributed crucial elements of the proof strategy.		
Signature	Bastian Wiederhold	Date	16th July 2024

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Alison Etheridge		
Supervisor comments I am confident that Bastian Wiederhold made a very substantial contribution to this publication.		
Signature	Date	16th July 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

5
Conclusion and outlook

In this thesis, we have seen how forwards in time approaches to describing genealogies can allow insights into complex models placed in the realm of Λ-Fleming-Viot processes. In Chapter 2, lookdown constructions were used to deal with the implications of a non-reversible population size, which conveys information about past events - and thus renders a standard sampling duality intractable. In Chapter 3 and Chapter 4, we have demonstrated that profound connections between central limit theorems and the pairwise asymptotic behaviour of ancestral lineages exist.

The application to real-world data and the incorporation of novel biological effects are potential research continuations. In this discussion, I wish to focus on two other topics. The first one is how the techniques in this document relate to one another and to the wider literature on genealogies of population models. The second are research ideas related to generalised Wright-Malécot formulae.

5.1 Potential connections between genealogical approaches

We have seen two approaches to recovering backwards in time information through a forwards in time construction. The notable parallel is the use of 'neutral tracers'. In the lookdown construction, one might argue that the exchangeable labels have an effect on reproduction events, while on the other hand their uniform assignment and non-existing influence on the average process justify viewing them as neutral. The genetic types of Chapter 3 and Chapter 4 can clearly be seen as tracers of families in

the population and do not influence reproduction events. Is there a way to view the central limit theorem technique from a lookdown perspective? Or is the similarity merely a product of requiring tracers of relationships in forwards in time approaches?

The mutation mechanism of Chapter 3 and Chapter 4 is reminiscent of [DK99a, Section 6], which constructs markers on the space $[0,1]^\infty$ with each of the components assuming new uniform values from the interval $[0,1]$ at unit rate. While this construction is used to trace ancestors in a lookdown model, the precise construction is introduced to allow uniform sampling from all possible paths in the model and seems more of a technical nature.

As mentioned in Section 4.3.4, [Eth+24] derives a very similar motion of ancestral lineages using the lookdown framework and the DPhil thesis of Terence Tsui Ho Lung [Tsu23, Conclusion] indicates an extension to two lineages. Indeed, an underlying idea is to quantify the number of individuals at the present location of interest, which have originated from a position in the past. This tracer idea has also been in used in, for example, in [BEK21] and [HN08].

Tracer dynamics are arguably the more natural approach to calculating probabilities of identity than the central limit theorems of Chapter 3 and Chapter 4. The central limit theorems are a particularly nice solution to dealing with the limiting effects of ancestral lines not meeting - but are they necessary? Could the same be achieved by applying the correct scaling to the probability of identity and using the tracers directly?

Apart from the genealogies based on lookdown constructions, as mentioned in Chapter 2, marked metric measure spaces [DGP11] can be used to describe genealogies - and have even been fused with the lookdown framework in [Bla+23]. Can the central limit theorem approach be interpreted from the perspective of genealogies as metric spaces? In the context of marked metric measures spaces, it has been proven fruitful to consider typical particles, so-called spines of the population ([FS23]). Does the Wright-Malécot formula relate to spinal particles?

5.2 Wright-Malécot formulae continuations

While the Wright-Malécot formula has an interpretation as the limiting behaviour of two ancestral lineages, at least two interesting questions remain to be studied: what is the influence of sampling at a 'smaller' or 'larger' separation? how does the Wright-Malécot formula relate to the complete genealogy of the model?

5.2.1 Sampling at different scales

In this subsection, we will outline a strategy for sampling individuals from a smaller distance than in Chapter 3 and Chapter 4.

Let us first introduce a notation for the limiting behaviour of the probability of identity $P_t^N(\phi, \psi)$. Recall that in the fixed radius case of Chapter 3, we saw the existence of the limit

$$\lim_{N\to\infty} N\delta_N^{2-d} P_t^N(\phi_N, \psi_N) = \lim_{N\to\infty} \mathbb{E}\Big[\big\langle Z_t^N \otimes Z_t^N, \mathbb{1}_\Delta \phi_N \otimes \psi_N \big\rangle\Big].$$

Informally, we will denote this relationship by

$$P_t^N(\phi_N, \psi_N) \sim \frac{1}{N\delta_N^{2-d}},$$

indicating that both quanties are of the same order as $N \to \infty$.

As $N \to \infty$, in the setting of Chapter 2 and Chapter 3, the densities ϕ, ψ according to which the positions are sampled for the probability of identity remain the same - while the process and thus space is rescaled. This effectively corresponds to sampling individuals from further and further apart in space or as we scale space by $1/\delta_N$ on the order $O(1/\delta_N)$. By scaling the densities ϕ_N, ψ_N simultaneously as we scale space and time, one can generate different behaviours for how the separation between samples evolves as the model approaches its limit. We could introduce the scaling

$$\phi_N(x) = \frac{1}{\delta_N^{(1-\beta)d}} \phi(x/\delta_N^{1-\beta})$$

with ψ_N defined correspondingly. This would allow us to interpolate between our setting of Chapter 3 and Chapter 4, corresponding to $\beta = 1$, in which the

distance between samples is $O(1/\delta_N)$, and $\beta = 0$, in which the distance remains of $O(1)$ as $N \to \infty$. What is the order of $P_t^N(\phi_N, \psi_N)$ for different values of $\beta \in [0,1]$ as $N \to \infty$?

Again note that the fixed radius, $r_1 = r_2$ case of Chapter 2 (corresponding to $\beta = 0$) indicates

$$P_t^N(\phi^N, \psi^N) \sim \frac{1}{N\delta_N^{2-d}} \sim \begin{cases} \dfrac{1}{N\delta_N} & d = 1, \\ \dfrac{1}{N} & d = 2, \\ \dfrac{1}{N\delta_N^{-1}} & d \geq 3. \end{cases} \tag{5.1}$$

Heuristic calculations with Raphaël Forien for the fixed radius case with general $\beta \in [0,1]$ suggest

$$P_t^N(\phi^N, \psi^N) \sim \begin{cases} \dfrac{1}{N\delta_N} & d = 1, \\ \dfrac{1 + \left|\log\left(\delta_N^{2(1-\beta)}\right)\right|}{N} & d = 2, \\ \dfrac{1}{N\delta_N^{2-d}}\delta_N^{(1-\beta)(2-d)} & d \geq 3. \end{cases} \tag{5.2}$$

Reassuringly, the behaviour of (5.2) generalises (5.1). In one dimension, $\frac{1}{N\delta_N}$ is independent of β and thus the separation between samples does not change the order of the probability of identity. This is likely explained by the recurrent behaviour of Brownian motions in one dimension. In contrast, (5.2) suggests that in dimension two and three at a separation of distance $O(1/\delta_N)$ the probability of identity is much lower than for samples chosen at distance $O(1)$. Again, this might be linked to the neighborhood-recurrent or transient behaviour of the Brownian motions. The asymptotics of (5.2) might also allow a closer alignment to small separation regimes of the physics paper [SW23], mentioned in the introduction of Chapter 3 and Chapter 4. In any case, these observations indicate that future works might be able to unravel more precise behaviours.

5.2.2 Wright-Malécot and genealogies

Stemming from the idea of probability of identity by descent, it is natural to interpret the Wright-Malécot formula in terms of the behaviour of ancestral lineages.

However, there are situations such as in [Eth+24], in which the same profile of the population might give rise to different genealogies. What would be the Wright-Malécot formula in such a scenario? More evidence is necessary to deduce that the Wright-Malécot formula is equivalent to pairwise behaviour of lineages within the genealogy of a model.

Another exciting question is whether the central limit theorem approach of Chapter 3 and Chapter 4 can be used to calculate the probability of finding the same type in a higher number $n \geq 3$ of initial samples. This would amount to determining limits of form

$$\lim_{N\to\infty} \gamma^N \mathbb{E}\left[\left\langle \bigotimes_{i=1}^{n} \rho_t^N, \mathbb{1}_\Delta \bigotimes_{i=1}^{n} \phi_i \right\rangle\right], \tag{5.3}$$

where $\mathbb{1}_\Delta = \prod_{i=1}^{n-1} \mathbb{1}_{k_i=k_{i+1}}$ and $\phi_i : \mathbb{R}^d \to \mathbb{R}^+$ are probability densities. The scaling sequence γ^N is yet to be identified, though in light of Chapter 3 and Chapter 4 the n-th power of $\sqrt{N\delta_N^{2-d}}$ is a candidate. The resulting formulae would reflect the asymptotic genealogies of the process.

In view of the difficulties of Chapter 4, it might be challenging to prove sufficient bounds to ensure the convergence of prelimiting probabilities to their limiting quantity. On the other hand, credibility to this aim has been added by a recent work of Peter Koepernick, which introduces a generalisation of sampling-consistency to spatial coalescents ([Koe24]). The resulting canonical object, the spatial Brownian coalescent, connects, conditionally on the coalescence locations and times, the samples by Brownian bridges. It is described by the terms one would expect for a Wright-Malécot formula for a higher number of samples (see [Koe24, Example 1.9]).

The coalescent tree of two samples is described by the coalescence location and time. The Wright-Malécot formula contains an exponential for the probability of not observing a mutation until the coalescence and a Gaussian transition kernel for the displacement, which the ancestral lineages need to cover to the coalescence location. The exponential stems from the mutation, but effectively provides a time-frame for coalescence. For a higher number of samples, one requires the locations and times of all coalescence events. The density of the Brownian spatial coalescent

then incorporates exponential terms for each of the waiting times until the next coalescence and again a Gaussian term for each of the spatial distances lineages need to cover in between coalescence events (see e.g. [Koe24, Figure 1]).

It would be fascinating to understand how one could use the probability of identity to uncover complete genealogies underlying a spatial population model.

5.3 Closing remarks

In view of the complexity of biology mentioned in the introduction, mathematical genetics will continue to amaze generations of scientists. Even a rather simple heuristic idea such as the original Wright-Malécot formula can unveil a complex underlying dynamic. Surely, we have only witnessed the beginning of the synthesis of models from ecology and genetics. The development of more realistic models will constantly pose analytical challenges, while the availability of empirical evidence will increase with each scientific generation. It has been an honour to contribute a small part to our understanding of our wonderful world.

Bibliography

[Ald78] David Aldous. “Stopping Times and Tightness”. In: *The Annals of Probability* 6.2 (1978), pp. 335–340.

[BDE02] Nicholas Barton, Frantz Depaulis, and Alison Etheridge. “Neutral evolution in spatially continuous populations”. In: *Theoretical Population Biology* 61.1 (2002), pp. 31–48.

[BES04] Nicholas Barton, Alison Etheridge, and Anja Sturm. “Coalescence in a random background”. In: *The Annals of Applied Probability* 14.2 (2004), pp. 754–785.

[BEV10] Nicholas Barton, Alison Etheridge, and Amandine Véber. “A new model for evolution in a spatial continuum”. In: *Electronic Journal of Probability* 15 (2010), pp. 162–216.

[BEV13a] Nicholas Barton, Alison Etheridge, and Amandine Véber. “Modelling evolution in a spatial continuum”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2013.1 (Jan. 2013).

[BKE10] Nicholas Barton, Jerome Kelleher, and Alison Etheridge. “A new model for extinction and recolonization in two dimensions: Quantifying phylogeography”. In: *Evolution* 64.9 (2010), pp. 2701–2715.

[BEV13b] Nathanaël Berestycki, Alison Etheridge, and Amandine Véber. “Large scale behaviour of the spatial Λ-Fleming–Viot process”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 49.2 (2013), pp. 374–401.

[BK93] Abhay Bhatt and Rajeeva Karandikar. “Weak convergence to a Markov process: The martingale approach”. In: *Probability Theory and Related Fields* 96 (1993), pp. 335–351.

[BCD16] Matthias Birkner, Jiri Cerny, and Andrej Depperschmidt. “Random walks in dynamic random environments and ancestry under local population regulation”. In: *Electronic Journal of Probability* 21 (2016), pp. 1–43.

[BDS24] Matthias Birkner, Andrej Depperschmidt, and Timo Schlüter. *Quenched CLT for ancestral lineages of logistic branching random walks.* 2024. arXiv: `2403.08567 [math.PR]`.

[Bir+05] Matthias Birkner et al. “Alpha-stable branching and beta-coalescents”. In: *Electronic Journal of Probability* 10 (2005), pp. 303–325.

[BEK21] Niloy Biswas, Alison Etheridge, and Aleksander Klimek. “The spatial Λ-Fleming-Viot process with fluctuating selection”. In: *Electronic Journal of Probability* 26 (2021), pp. 1–51.

[Bla+23] Airam Blancas et al. “Evolving genealogies for branching populations under selection and competition”. In: *The Annals of Applied Probability* 33.6A (2023), pp. 4528–4569.

[CMS00] Michael Cain, Brook Milligan, and Allan Strand. “Long-distance seed dispersal in plant populations”. In: *American Journal of Botany* 87.9 (2000), pp. 1217–1227.

[CS20] Adrian Gonzalez Casanova and Charline Smadi. "On Λ-Fleming–Viot processes with general frequency-dependent selection". In: *Journal of Applied Probability* 57.4 (2020), pp. 1162–1197.

[CPS22] Adrián Gonzalez Casanova, Verónica Pina, and Arno Siri-Jégousse. "The symmetric coalescent and Wright–Fisher models with bottlenecks". In: *The Annals of Applied Probability* 32.1 (2022), pp. 235–268.

[Cho95] Nhansook Cho. "Weak convergence of stochastic integrals driven by martingale measure". In: *Stochastic Processes and their Applications* 59.1 (1995), pp. 55–79.

[Cor+22] Fernando Cordero et al. "Lambda-coalescents arising in a population with dormancy". In: *Electronic Journal of Probability* 27 (2022), pp. 1–34.

[CD02] Theodore Cox and Richard Durrett. "The stepping stone model: new formulas expose old myths". In: *The Annals of Applied Probability* 12.4 (2002), pp. 1348–1377.

[DGP11] Andrej Depperschmidt, Andreas Greven, and Peter Pfaffelhuber. "Marked metric measure spaces". In: *Electronic Communications in Probability* 16 (2011), pp. 174–188.

[DK96] Peter Donnelly and Thomas Kurtz. "A countable representation of the Fleming-Viot measure-valued diffusion". In: *The Annals of Probability* 24.2 (1996), pp. 698–742.

[DK99a] Peter Donnelly and Thomas Kurtz. "Genealogical processes for Fleming-Viot models with selection and recombination". In: *The Annals of Applied Probability* 9.4 (1999), pp. 1091–1148.

[DK99b] Peter Donnelly and Thomas Kurtz. "Particle Representations for Measure-Valued Population Models". In: *The Annals of Probability* 27.1 (1999), pp. 166–205.

[Eth00] Alison Etheridge. *An Introduction to Superprocesses*. Vol. 20. American Mathematical Society Providence, RI, 2000.

[Eth08] Alison Etheridge. "Drift, draft and structure: some mathematical moodels of evolution". In: *Banach Center Publications* 80 (2008), pp. 121–144.

[EFP17] Alison Etheridge, Nic Freeman, and Sarah Penington. "Branching Brownian motion, mean curvature flow and the motion of hybrid zones". In: *Electronic Journal of Probability* 22 (2017), pp. 1–40.

[EK19] Alison Etheridge and Thomas Kurtz. "Genealogical constructions of population models". In: *The Annals of Probability* 47.4 (2019), pp. 1827–1910.

[EVY20] Alison Etheridge, Amandine Véber, and Feng Yu. "Rescaling Limits of the Spatial Lambda-Fleming-Viot Process with Selection". In: *Electronic Journal of Probability* 25 (2020), pp. 1–89.

[Eth+24] Alison Etheridge et al. "Looking forwards and backwards: Dynamics and genealogies of locally regulated populations". In: *Electronic Journal of Probability* 29 (2024), pp. 1–85.

[EK86] Stewart Ethier and Thomas Kurtz. *Markov Processes: Characterization and Convergence*. Wiley series in probability and mathematical statistics. Wiley, 1986.

[Fel75] Joseph Felsenstein. "A Pain in the Torus: Some Difficulties with Models of Isolation by Distance". In: *The American Naturalist* 109.967 (1975), pp. 359–368.

[FV79] Wendell H. Fleming and Michel Viot. "Some Measure-Valued Markov Processes in Population Genetics Theory". In: *Indiana University Mathematics Journal* 28.5 (1979), pp. 817–843.

[For22] Raphaël Forien. "Stochastic partial differential equations describing neutral genetic diversity under short range and long range dispersal". In: *Electronic Journal of Probability* 27 (2022), pp. 1–41.

[FP17] Raphaël Forien and Sarah Penington. "A central limit theorem for the spatial Lambda Fleming-Viot process with selection". In: *Electronic Journal of Probability* 22.5 (2017), pp. 1–68.

[FS23] Félix Foutel-Rodier and Emmanuel Schertzer. "Convergence of genealogies through spinal decomposition with an application to population genetics". In: *Probability Theory and Related Fields* 187 (2023), pp. 697–751.

[FK07] Zdenek Fric and Martin Konvicka. "Dispersal kernels of butterflies: Power-law functions are invariant to marking frequency". In: *Basic and Applied Ecology* 8.4 (2007), pp. 377–386.

[Guf18] Stephan Gufler. "Pathwise construction of tree-valued Fleming-Viot processes". In: *Electronic Journal of Probability* 23 (2018), pp. 1–58.

[HN08] Oskar Hallatschek and David R. Nelson. "Gene surfing in expanding populations". In: *Theoretical Population Biology* 73.1 (2008), pp. 158–170.

[JS03] Jean Jacod and Albert Shiryaev. *Limit Theorems for Stochastic Processes*. Vol. 288. Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 2003.

[JK14] Sabine Jansen and Noemi Kurt. "On the notion(s) of duality for Markov processes". In: *Probability Surveys* 11 (2014), pp. 59–120.

[KK03] Ingemar Kaj and Stephen Krone. "The coalescent process in a population with stochastically varying size". In: *Journal of Applied Probability* 40.1 (2003), pp. 33–48.

[Kat91] Garry Katzenberger. "Solutions of a stochastic differential equation forced onto a manifold by a large drift". In: *The Annals of Probability* 19.4 (1991), pp. 1587–1628.

[Kel+14] Ruth Kelly et al. "Historical data reveal power-law dispersal patterns of invasive aquatic species". In: *Ecography* 37.6 (2014), pp. 581–590.

[KW64] Motoo Kimura and George H Weiss. "The stepping stone model of population structure and the decrease of genetic correlation". In: *Genetics* 49.4 (Apr. 1964), pp. 561–576.

[Kin82] J.F.C. Kingman. "The coalescent". In: *Stochastic Processes and their Applications* 13.3 (1982), pp. 235–248.

[Koe24] Peter Koepernik. *The Brownian spatial coalescent*. 2024. arXiv: 2401.08557 [math.PR].

[Küh18] Franziska Kühn. "Solutions of Lévy-driven SDEs with unbounded coefficients as Feller processes". In: *Proceedings of the American Mathematical Society* 146.8 (Feb. 2018), pp. 3591–3604.

[Kul09] Alexey M. Kulik. "Exponential ergodicity of the solutions to SDE's with a jump noise". In: *Stochastic Processes and their Applications* 119.2 (2009), pp. 602–632.

[Kur98] Thomas Kurtz. "Martingale problems for conditional distributions of Markov processes". In: *Electronic Journal of Probability* 3 (1998), pp. 1–29.

[Kur11] Thomas Kurtz. "Equivalence of stochastic equations and martingale problems". In: *Stochastic Analysis 2010*. Ed. by Dan Crisan. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 113–130.

[KN11] Thomas Kurtz and Giovanna Nappo. "The filtered martingale problem". In: *The Oxford Handbook of nonlinear filtering*. Ed. by Dan Crisan and Boris Rozovskii. Jan. 2011, pp. 129–165.

[Kwa17] Mateusz Kwaśnicki. "Ten equivalent definitions of the fractional Laplace operator". In: *Fractional Calculus and Applied Analysis* 20.1 (2017), pp. 7–51.

[Lau+23] Sally Lau et al. "Genomic evidence for west Antarctic ice sheet collapse during the last interglacial". In: *Science* 382.6677 (2023), pp. 1384–1389.

[LB03] Jean-François Le Gall and Jean Bertoin. "Stochastic Flows Associated to Coalescent Processes". In: *Probability Theory and Related Fields* 126.2 (2003), pp. 261–288.

[Mal48] Gustave Malécot. *Les Mathématiques de l'Hérédité*. Masson et Cie., Paris, 1948.

[Pit99] Jim Pitman. "Coalescents with multiple collisions". In: *The Annals of Probability* 27.4 (1999), pp. 1870–1902.

[Rin+18] Harald Ringbauer et al. "Estimating barriers to gene flow from distorted isolation-by-distance patterns". In: *Genetics* 208.3 (Mar. 2018), pp. 1231–1245.

[Sag99] Serik Sagitov. "The general coalescent with asynchronous mergers of ancestral lines". In: *Journal of Applied Probability* 36.4 (1999), pp. 1116–1125.

[Sch03] Jason Schweinsberg. "Coalescent processes obtained from supercritical Galton–Watson processes". In: *Stochastic Processes and their Applications* 106.1 (2003), pp. 107–139.

[Sga+22] Gabriele Maria Sgarlata et al. "The effect of habitat loss and fragmentation on isolation-by-distance and time". In: *bioRxiv* (2022). eprint: https://www.biorxiv.org/content/early/2022/10/27/2022.10.26.513874.full.pdf.

[SW23] Tyler Smith and Daniel Weissman. "Isolation by distance in populations with power-law dispersal". In: *G3 Genes|Genomes|Genetics* 13.4 (Jan. 2023), jkad023.

[Tre+20] Jelle Treep et al. “Seed dispersal as a search strategy: dynamic and fragmented landscapes select for multi-scale movement strategies in plants”. In: *Movement Ecology* 9.4 (2020).

[Tsu23] Ho Lung Tsui. “Branching and Fragmenting Particle Systems in Biology”. PhD thesis. University of Oxford, 2023.

[Val+17] Valentin Vallaeys et al. “A Lévy-flight diffusion model to predict transgenic pollen dispersal”. In: *Journal of The Royal Society Interface* 14 (2017), p. 20160889.

[VW15] Amandine Véber and Anton Wakolbinger. “The spatial Λ-Fleming–Viot process: An event-based construction and a lookdown representation”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 51.2 (2015), pp. 570–598.

[Wal86] John B. Walsh. “An introduction to stochastic partial differential equations”. In: *Ecole d’été de Probabilités de Saint-Flour, XIV—1984*. Vol. 1180. Lecture Notes in Math. Springer, Berlin, 1986, pp. 265–439.

[Wri43] Sewall Wright. “Isolation by distance”. In: *Genetics* 28.2 (1943), p. 114.

A
Appendix

A.1 Existence and uniqueness of the SLFV with mutations

The aim of this section is to prove Theorem 3.2.2. We adapt the arguments used to prove Theorem 1.2 in [EVY20] to the setting of the SLFV with different dispersal and parent search radius. We start by proving uniqueness of the solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$, using a duality argument.

A.1.1 The dual of the SLFV with mutations

We define an ancestral process $(\mathcal{A}_t, t \geq 0)$ taking values in

$$\Bigg\{\Big((\xi^1, S^1), (\xi^2, S^2), \ldots, (\xi^l, S^l)\Big) \in \bigcup_{n=1}^{N_0} \Big(\mathbb{R}^d \times \{\partial\}\Big)^n \times \mathcal{P}(\{1, \ldots, N_0\}) \\ : \bigcup_{i=0}^{l} S_i = \{0, 1, 2, \ldots, N_0\} \text{ and } \forall 0 \leq i \neq j \leq l \leq N_0 : S_i \cap S_j = \emptyset \Bigg\}.$$

At each time t, $\mathcal{A}_t$ consists of a finite collection of lineages in $\mathbb{R}^d \times \{\partial\}$. Their locations are denoted by $(\xi_t^j, 1 \leq j \leq N_t)$, where N_t is the number of lineages in $\mathcal{A}_t$ at time t. By $\{\partial\}$ we denote a cemetery state for the lineages. Each of the lineages ξ^j comes along with an element S^j of $\mathcal{P}(\{1, \ldots, N_0\})$ indicating the labels of initial samples descended from the lineage.

Definition A.1.1. Let Π be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^d \times (0,1] \times (0,\infty)^2$ with intensity measure $dt \otimes dx \otimes \nu(du, dr_1, dr_2)$. The initial condition is given by

$$\mathcal{A}_0 = \Big((\xi^1, \{1\}), (\xi^2, \{2\}), \dots, (\xi^{N_0}, \{N_0\})\Big),$$

where $\xi^j \in \mathbb{R}^d$ and $N_0 \in \mathbb{N}$. At each event $(t, x, u, r_1, r_2) \in \Pi$,

1. pick a location y uniformly at random from $B(x, r_1)$, the parent search area,
2. for each lineage which finds itself in the replacement area just before the event, i.e. such that $\xi^j_{t^-} \in B(x, r_2)$, mark this lineage independently of the others with probability u, and let $J \subset \{1, \dots, N_0\}$ denote the set of marked lineages,
3. all marked lineages coalesce at time t, so the entries $\{(\xi^j_{t-}, S^j_{t-}), j \in J\}$ are replaced by the single entry $(y, \cup_{j\in J} S^j_{t-})$.

Moreover, each lineage is killed at rate μ independently of each other, and killed lineages instantly jump to ∂, where they are no longer affected by reproduction events.

Notice that each lineage jumps when it finds itself in the replacement area of an event and when it is marked, which happens at rate

$$\int_{(0,1]\times(0,\infty)^2} u\,|B(0, r_2)|\,\nu(du, dr_1, dr_2).$$

By (3.9), this rate is finite, and since we consider only finitely many lineages, the jump rate of $(\mathcal{A}_t, t \geq 0)$ is almost surely bounded by a deterministic constant (the number of lineages cannot increase). Hence the above definition gives rise to a well-defined process. The following proposition states the duality relation between the SLFV with mutations and the ancestral process defined above. Its proof is omitted as it is very similar to the proof of the corresponding result in [EVY20] (Proposition 2.2), see also [VW15, Theorem 1].

Proposition A.1.2. *Any Ξ-valued Markov process $(\rho_t, t \geq 0)$ solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$ is dual to the ancestral process $(\mathcal{A}_t, t \geq 0)$*

in the sense that, for any $N_0 \in \mathbb{N}$*, for any* $\psi : \mathbb{R}^d \to \mathbb{R}$ *continuous and compactly supported and any* $\chi : [0,1] \to \mathbb{R}$ *continuous,*

$$\begin{aligned}&\mathbb{E}_{\rho_0}\Big[\langle \rho_t, \psi \otimes \chi\rangle^{N_0}\Big] \\ &= \int_{(\mathbb{R}^d)^{N_0}} \mathbb{E}_{(x_1,\dots,x_{N_0})}\left[\prod_{i=1}^{N_t}\int_{[0,1]}(\chi(k))^{|S_t^i|}\rho_0(\xi_t^i, dk)\right]\prod_{j=1}^{N_0}\psi(x_j)dx_1\dots dx_{N_0},\end{aligned}$$

with the convention that $\rho_0(\partial, dk) = \lambda(dk)$*. Here,* $\mathbb{E}_{\rho_0}[\cdot]$ *denotes the expectation with respect to the law of* $(\rho_t, t \geq 0)$ *with initial condition* ρ_0*. The cardinality of the set* S *is denoted by* $|S|$ *and* $\mathbb{E}_{(x_1,\dots,x_{N_0})}[\cdot]$ *is the expectation with respect to the law of the ancestral process started from* $\mathcal{A}_0 = \Big((x_1, \{1\}), \dots, (x_{N_0}, \{N_0\})\Big)$*.*

Since the ancestral process is well defined and the expectation of the set of functions of the form $\rho \mapsto \langle \rho, \psi \otimes \chi\rangle^{N_0}$ characterises the distribution of ρ (it is dense in the set of continuous real-valued functions on Ξ, see Lemma 2.1 in [VW15]), Proposition A.1.2 ensures that the solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$ is unique (if it exists).

A.1.2 Proof of Theorem 3.2.2

To prove the existence of a solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$, we follow the argument of [EVY20, Theorem 1.2], i.e. we construct a sequence of processes $(\rho^{(n)}, n \geq 0)$ which converges in distribution to a solution of the martingale problem. The processes $(\rho^{(n)}, n \geq 0)$ will be constructed as solutions to a sequence of approximating martingale problems, $(\mathcal{G}^{(n)}, D(\mathcal{G}^{(n)}))$, such that the overall jump rate of $(\rho_t^{(n)}, t \geq 0)$ is bounded almost surely for each $n \geq 0$.

To achieve this, consider an increasing sequence of compact subsets of $\mathbb{R}^d$, $(E_n, n \geq 0)$, such that $\cup_{n\geq 0} E_n = \mathbb{R}^d$, and an increasing sequence $(\nu_n, n \geq 0)$ of finite measures on $(0,1] \times (0,\infty)^2$ which converges to ν as $n \to \infty$. We then set,

for any $F = F_{f,\phi} \in D(\mathcal{G}^{(n)}) := D(\mathcal{G})$,

$$\begin{aligned}
&\mathcal{G}^{(n)}F_{f,\phi}(\rho)\\
&:= \mu f'(\langle\rho,\phi\rangle)\langle\lambda-\rho,\phi\rangle\\
&\quad + \int_{E_n}\int_{(0,1]\times(0,\infty)^2}\frac{1}{|B(x,r_1)|}\int_{B(x,r_1)}\int_{[0,1]}\Big(F_{f,\phi}(\Theta_{x,r_2,u,k}(\rho)) - F_{f,\phi}(\rho)\Big)\\
&\qquad\qquad \rho(y,dk)dy\,\nu_n(du,dr_1,dr_2)\,dx.
\end{aligned}$$

We then define $(\rho_t^{(n)}, t\geq 0)$ as the unique Ξ-valued Markov process which solves the martingale problem associated to $(\mathcal{G}^{(n)}, D(\mathcal{G}^{(n)}))$. This martingale problem is well posed because the overall jump rate of $(\rho_t^{(n)}, t\geq 0)$ is almost surely bounded (note that the definition of $(\rho_t^{(n)}, t\geq 0)$ amounts to replacing Π by a Poisson point process with intensity $\mathbb{1}_{x\in E_n}dt\otimes dx\otimes\nu_n(du,dr_1,dr_2)$ - which is a finite measure - in Definition 3.2.1).

To prove Theorem 3.2.2, it remains to show that $\rho^{(n)}$ converges in distribution as $n\to\infty$ to a solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$.

Lemma A.1.3. *For any $T>0$, the sequence of Ξ-valued Markov processes $(\rho^{(n)}, n\geq 0)$ is tight in $D([0,T],\Xi)$.*

Proof. First recall that, by Lemma 2.1 in [VW15], Ξ is a compact space. Therefore, by the Aldous-Rebolledo criterion (see e.g. [Eth00, Theorem 1.17]), it is enough to show that both the finite variation parts and the quadratic variations of the real-valued processes $(F_{f,\phi}(\rho_t^{(n)}), t\geq 0)$ are tight, for any $f\in C^1(\mathbb{R})$ and $\phi\in C_c(\mathbb{R}^d)$.

By the definition of $(\rho_t^{(n)}, t\geq 0)$, the finite variation part of $(F_{f,\phi}(\rho_t^{(n)}), t\geq 0)$ is given by

$$\int_0^t \mathcal{G}^{(n)}F_{f,\phi}(\rho_s^{(n)})ds. \tag{A.1}$$

First note that, for any $\rho\in\Xi$,

$$|f'(\langle\rho,\phi\rangle)\langle\lambda-\rho,\phi\rangle| \leq 2\|\phi\|_1 \sup_{|x|\leq\|\phi\|_1}|f'(x)|. \tag{A.2}$$

In addition,

$$|F_{f,\phi}(\Theta_{x,r,u,k}(\rho)) - F_{f,\phi}(\rho)| \leq u \sup_{|x|\leq\|\phi\|_1} |f'(x)| \left|\langle \mathbb{1}_{B(x,r)}(\delta_k - \rho), \phi\rangle\right|$$
$$\leq 2u \sup_{|x|\leq\|\phi\|_1} |f'(x)| \int_{B(x,r)} \sup_{k\in[0,1]} |\phi(y,k)|dy.$$

Since ϕ is continuous and compactly supported, there exists a constant $C > 0$ (depending only on f and ϕ) such that, for any $\rho \in \Xi$,

$$|F_{f,\phi}(\Theta_{x,r,u,k}(\rho)) - F_{f,\phi}(\rho)| \leq Cu(r^d \wedge 1)\mathbb{1}_{B(x,r)\cap supp(\phi)\neq\emptyset}. \tag{A.3}$$

Using (A.2), (A.3) and the fact that there exists a constant $C > 0$ such that

$$\int_{\mathbb{R}^d} \mathbb{1}_{B(x,r)\cap supp(\phi)\neq\emptyset} dx \leq C(r^d \vee 1), \tag{A.4}$$

we see that

$$\left|\mathcal{G}^{(n)} F_{f,\phi}(\rho)\right| \leq C \int_{(0,1]\times(0,\infty)^2} u(r_2)^d \nu_n(du, dr_1, dr_2) + C',$$

for some constants $C, C' > 0$. As a result, under condition (3.9), the modulus of continuity of (A.1) is almost surely uniformly bounded over all $n \in \mathbb{N}$, hence the finite-variation parts of $(F_{f,\phi}(\rho_t^{(n)}), t \geq 0)$ are tight.

The quadratic variation of $(F_{f,\phi}(\rho_t^{(n)}), t \geq 0)$ is given by

$$\int_0^t \Gamma^{(n)} F_{f,\phi}(\rho_s^{(n)}) ds,$$

where

$$\Gamma^{(n)} F_{f,\phi}(\rho) := \int_{E_n} \int_{(0,1]\times(0,\infty)^2} \frac{1}{|B(x,r_1)|} \int_{B(x,r_1)} \int_{[0,1]} \Big(F_{f,\phi}(\Theta_{x,r_2,u,k}(\rho)) - F_{f,\phi}(\rho)\Big)^2$$
$$\rho(y, dk)dy\, \nu_n(du, dr_1, dr_2)\, dx.$$

But, using (A.3) and the fact that $u \in (0, 1]$, there exists a constant $C > 0$ such that, for all $\rho \in \Xi$,

$$|F_{f,\phi}(\Theta_{x,r,u,k}(\rho)) - F_{f,\phi}(\rho)|^2 \leq Cu(r^d \wedge 1)\mathbb{1}_{B(x,r)\cap supp(\phi)\neq\emptyset}.$$

As a result there exists a constant $C > 0$ such that, for all $\rho \in \Xi$,

$$\left|\Gamma^{(n)} F_{f,\phi}(\rho)\right| \leq C,$$

and this shows that the quadratic variation processes of $(F_{f,\phi}(\rho_t^{(n)}), t \geq 0)$ are tight.

By the Aldous-Rebolledo criterion, this concludes the proof of the lemma. □

We can now conclude the proof of Theorem 3.2.2.

Proof of Theorem 3.2.2. Uniqueness of the solution to the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$ follows by the duality relation of Proposition A.1.2. To prove existence, by Lemma A.1.3, there exists a converging subsequence of the sequence $(\rho^{(n)}, n \geq 0)$, and it is enough to show that the limit of any such subsequence must solve the martingale problem associated to $(\mathcal{G}, D(\mathcal{G}))$ (we will also obtain the convergence in distribution of the whole sequence $(\rho^{(n)}, n \geq 0)$). By Theorem 4.8.10 in [EK86], it suffices to prove that, for any $F_{f,\phi} \in D(\mathcal{G})$,

$$\mathcal{G}^{(n)} F_{f,\phi}(\rho) \to \mathcal{G} F_{f,\phi}(\rho) \text{ as } n \to \infty, \text{ uniformly in } \rho \in \Xi.$$

But, using (A.3) and the fact that $\nu_n \leq \nu$, for any $\rho \in \Xi$,

$$\begin{aligned}
&\left|\mathcal{G}^{(n)} F_{f,\phi}(\rho) - \mathcal{G} F_{f,\phi}(\rho)\right| \\
&\leq C \int_{E_n^c} \int_{(0,1]\times(0,\infty)^2} u(r_2^d \wedge 1) \mathbb{1}_{B(x,r_2)\cap supp(\phi)\neq\emptyset} \nu(du, dr_1, dr_2) dx \\
&\quad + C \int_{\mathbb{R}^d} \int_{(0,1]\times(0,\infty)^2} u(r_2^d \wedge 1) \mathbb{1}_{B(x,r_2)\cap supp(\phi)\neq\emptyset} (\nu - \nu_n)(du, dr_1, dr_2).
\end{aligned}$$

The first term on the right hand side tends to zero as $n \to \infty$ by Lebesgue's dominated convergence theorem, using (3.9), and so does the second term, using (A.4). Since the right hand side does not depend on ρ, the convergence is uniform on Ξ, and the result is proved. □

A.2 Definition of worthy of martingale measures

Definition A.2.1 ([Wal86, Chapter 2], [Cho95, Definition 1.1])**.** Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space and $\mathcal{B}(\mathbb{R}^{\mathrm{d}})$ be the Borel-σ-algebra on $\mathbb{R}^{\mathrm{d}}$. A random set function $M : \mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$ is a martingale measure if the following three properties are satisfied:

1. For disjoint $A, B \in \mathcal{B}(\mathbb{R}^{\mathrm{d}}), A \cap B = \emptyset$ the set function M is additive, i.e. $M(t, A \cup B) = M(t, A) + M(t, B)$ almost surely.

2. For $A \in \mathcal{B}(\mathbb{R}^{\mathrm{d}})$, $\Big(M(t, A), t \geq 0\Big)$ is a martingale.

3. There exists an increasing sequence $(E_n, n \geq 1)$ where $E_n \in \mathcal{B}(\mathbb{R}^{\mathrm{d}})$ whose union is $\mathbb{R}^{\mathrm{d}}$ such that

   (i) for every $t > 0$, if $\mathcal{E}_n = \mathcal{B}(E_n)$ then $\sup_{A \in \mathcal{E}_n} \mathbb{E}\Big(M(t, A)^2\Big) < \infty$;

   (ii) and for every $t > 0$, if $(A_k, k \geq 1)$ is a sequence of $\mathcal{E}_n$ decreasing to $\emptyset$, then

$$\mathbb{E}\Big(M(t, A_k)^2\Big) \to 0.$$

Definition A.2.2 ([Wal86, p.291])**.** A martingale measure M is worthy if there exists a random σ-finite measure D on $\mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times \mathbb{R}^{\mathrm{d}})$ with the following properties:

1. D is symmetric with respect to the two spatial marginals and for fixed $A, B \in \mathcal{B}(\mathbb{R}^{\mathrm{d}})$ the process $\Big(D([0,t] \times A \times B), t \geq 0\Big)$ is predictable.

2. For any $(s,t] \times A \times B \subset \mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times \mathbb{R}^{\mathrm{d}}$ the covariance functional is bounded by D, i.e.

$$\Big| \Big\langle M(A), M(B) \Big\rangle_t - \Big\langle M(A), M(B) \Big\rangle_s \Big| \leq D\Big((s,t] \times A \times B\Big).$$

A.3 Convergence Theorems

Theorem A.3.1 ([For22, Theorem 3.7])**.** *Let $(M^N, N \geq 1)$ be a sequence of worthy martingale measures on $\mathbb{R}^{\mathrm{d}} \times [0,1]$ with corresponding dominating measures $(D^N, N \geq 1)$. Let*

$$\psi^N : \Big\{(s,t) : 0 \leq s \leq t\Big\} \times \mathbb{R}^{\mathrm{d}} \times [0,1] \to \mathbb{R} \tag{A.5}$$

be a sequence of functions and U_t^N be the sequence of stochastic integrals

$$U_t^N := \int_{[0,t] \times \mathbb{R}^{\mathrm{d}} \times [0,1]} \psi_{s,t}^N(x,k) M^N(ds, dx, dk).$$

If $(\psi^N, N \geq 1)$ and $(D^N, N \geq 1)$ satisfy the conditions below with constants $C_1, C_2, C_3 > 0$ and $k \in \mathbb{N}$, the sequence U_t^N is tight in $D(\mathbb{R}_+, \mathbb{R})$ and for all $T > 0$ there exists a constant $C(T)$ such that for all $N \geq 1$

$$\mathbb{E}\left(\sup_{t \in [0,T]} |U_t^N|^2 \right) \leq C(T) C_1 k (C_2^2 + C_3^2).$$

The conditions are:

1. *There exists* $C_1 > 0$ *and* $k \in \mathbb{N}$ *such that for any* $0 \leq s \leq t$*, for all* $N \geq 1$ *and any Schwartz function* $\phi \in \mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$

$$\int_{[s,t]\times(\mathbb{R}^{\mathrm{d}}\times[0,1])^2} \phi(x_1,k_1)\phi(x_2,k_2)D^N(ds,dx_1,dk_1,dx_2,dk_2) \leq C_1|t-s|\sum_{p=1}^{k}\|\phi\|_p^2. \tag{A.6}$$

2. *For any* $0 \leq s \leq t$ *and for all* $N \geq 1$ *the function* $\psi^N_{s,t}$ *is in* $\mathcal{S}(\mathbb{R}^{\mathrm{d}} \times [0,1])$ *and the functions* $t \mapsto \psi^N_{s,t}$ *and* $s \mapsto \psi^N_{s,t}$ *are both continuous.*

3. *There exist* $C_2, C_3, \mu > 0$ *such that for all* $N \geq 1, 0 \leq s \leq t$ *and* $p \in [1,k]$

$$\left\|\psi^N_{s,t}\right\|_p \leq C_2 e^{-\mu(t-s)},$$

and for all $N \geq 1, 0 \leq s' \leq s \leq t \leq t', p \in [1,k]$

$$\left\|\psi^N_{s,t'} - \psi^N_{s,t}\right\|_p \leq C_3|t'-t|e^{-\mu(t-s)}, \qquad \left\|\psi^N_{s',t} - \psi^N_{s,t}\right\|_p \leq C_3|s'-s|e^{-\mu(t-s)}.$$

Theorem A.3.2 ([JS03, Theorem VIII 3.11])**.** *Let* $((X^N_t)_{t\geq 0}, N \geq 1)$ *be a sequence of càdlàg, locally square-integrable, d-dimensional martingales and let* $(X_t, t \geq 0)$ *be a continuous, d-dimensional Gaussian martingale. We write* $X^N_t = \left(X^{N,1}_t, ..., X^{N,d}_t\right)$ *and* $X_t = \left(X^1_t, ..., X^d_t\right)$ *for the corresponding vectors. Suppose that*

1. $\sup_{t\geq 0}\left|X^N_t - X^N_{t^-}\right|$ *is uniformly bounded for* $N \geq 1$ *and converges to zero in probability,*

2. *for each* $t \in \mathcal{Q}$ *and* $1 \leq i, j \leq d$*, the pairwise predictable quadratic variations*

$$\left\langle X^{N,i}, X^{N,j}\right\rangle_t \xrightarrow[N\to\infty]{} \left\langle X^i, X^j\right\rangle_t$$

converge in probability.

Then X^N *converges to* X *in distribution in* $C(\mathbb{R}_+, \mathbb{R}^{\mathrm{d}})$*.*

Theorem A.3.3 ([For22, Proposition E.1])**.** *Consider a sequence of worthy martingale measures* $(M^N, N \geq 1)$ *on* $\mathbb{R}^d \times [0,1]$ *with corresponding dominating measures* $(D^N, N \geq 1)$*. Suppose that* $(M^N, N \geq 1)$ *converges in distribution to a martingale measure* M *in* $D(\mathbb{R}_+, \mathcal{S}'(\mathbb{R}^d \times [0,1]))$ *and that the dominating measures are bounded in the sense of* (A.6) *for a* $k \geq 1$*. If a family of deterministic real-valued functions* $(f^N_i, N \geq 1, i \leq n)$ *on* $[0,T] \times \mathbb{R}^d \times [0,1]$ *satisfies*

1. *for all* $N \geq 1, 1 \leq i \leq n$ *and* $s \in [0,T]$ *the function* $f_i^N(s,\cdot)$ *is contained in* $\mathcal{S}(\mathbb{R}^d \times [0,1])$,

2. *the functions* f_i^N *are uniformly bounded in* L^p *for any* $p \in [1,k]$ *and* $1 \leq i \leq n$,

$$\sup_{N\geq 1} \sup_{s\in[0,T]} \left\| f_i^N(s,\cdot) \right\|_p < +\infty,$$

3. *there exists limiting functions* $f_i : [0,T] \times \mathbb{R}^d \times [0,1] \to \mathbb{R}$, $1 \leq i \leq n$ *such that the convergence is uniform in* L^p *for* $p \in [1,k]$,

$$\lim_{N\to\infty} \sup_{s\in[0,T]} \left\| f_i^N(s,\cdot) - f_i(s,\cdot) \right\|_p = 0,$$

then the finite-dimensional distributions

$$\Big(M_t^N(f_1^N), ..., M_t^N(f_n^N) \Big)_{t\in[0,T]} \xrightarrow[N\to\infty]{} \Big(M_t(f_1), ..., M_t(f_n) \Big)_{t\in[0,T]}$$

converge in distribution in $D([0,T], \mathbb{R}^n)$.

B
Appendix

B.1 Details on the simulation

This section details the principles underlying Figure 4.2.

B.1.1 Simulation of the process

We consider a discretized version of the process defined by Definition 4.2.2 on $\{0, 1, 2, ..., 99, 100\}$. We represent the uniform proportion of the population by one type, while allowing for up to 2000 other types in the population. If the uniform proportion type is chosen as the parental type in an event, we choose instead the next still empty type or replace the type with the smallest total mass in the population. We start the process with a uniform mass of three everywhere and let it evolve for a time of 125 under the growth function

$$r_x(n) = \max\left\{\min\left\{14\frac{|x-50|}{50}, 7\right\} + 1 - n, -1\right\}.$$

Additionally, we choose

$$\mu = 0.0001, \qquad u = 0.04, \qquad R = 4.$$

B.1.2 Analytical prediction

For the analytical part we consider ancestral lineages backwards in time. Again, we consider the discretized space $\{0, ..., 100\}$. We consider the population at

stationarity solving

$$\frac{uV_R R^2}{3}\Big(n(x+1)+n(x-1)-2n(x)\Big)+r_x(n(x))n(x)=0 \ \ \forall x\in\{1,...,99\},$$

with boundary condition $n(0)=8=n(100)$ and precisely the same growth function r as for the process simulation. To describe the movement of ancestral lineages, we consider the continuous time random walk with transition rates

$$Q(i,j)=\begin{cases}\dfrac{n(j)}{n(i)} & j=i\pm 1,\\ -\dfrac{n(i+1)}{n(i)}-\dfrac{n(i-1)}{n(i)} & j=i,\\ 0 & \text{else.}\end{cases}$$

This corresponds to a discretization of the lineages evolving according to the operator (4.21). We can calculate the transition probabilities of the ancestral lineages as the matrix exponential

$$\mathfrak{P}=e^{Q}.$$

We can then consider the sequence $K_0=diag\{1,...,1\}$, the 100×100 matrix with only ones on the diagonal, and $K_t=K_{t-1}\mathfrak{P}$. The (i,j) entry of K_t will encapsulate the probability to transition from i to j over the period t. We can therefore describe the probability of identity by descent for two positions l_1,l_2 by

$$\Theta(l_1,l_2)=\sum_{t\geq 1}^{t_{\max}}\sum_{0\leq z\leq 100}e^{-2\mu i}\frac{1}{n(z)+1}K_i(l_1,z)K_i(l_2,z).$$

Defining the diagonal matrix $D(n)=\{(n(0)+1)^{-1},...,(n(100)+1)^{-1}\}$, we can calculate the probability of identity between all possible combinations of positions as the matrix

$$\Theta=\sum_{t\geq 1}^{t_{\max}}e^{-2\mu t}K_t D(n)(K_t)^T,$$

where $(K_t)^T$ is the transposition of K_t. For Figure 4.2 we chose $t_{\max}=28$ as due to the finite range the ancestral lineages start to equilibriate, leading to a much shallower decay as on would expect on a large domain.

B.1.3 Alignment of simulation and prediction

As mentioned in the introduction, The analytical result A corresponds to the limiting Wright-Malecot formula (in particular not necessarily being a probability), whereas we obtain probabilities of identity $\mathcal{P}$ from the simulation of the process. The outcome of simulation $\mathcal{P}^N(x)$ and analytical prediction $\Theta(\mu, u, x)$ for locations at distance x can be connected through

$$N\delta_N \mathcal{P}^N(x) \simeq \Theta\Big(\frac{N}{\delta_N^2}\mu_N, Nu_N, \delta_N x\Big),$$

corresponding to our scaling of the probability of identity, space and the parameters. As we do not consider a sequence of processes, in our case we can set $\mu_N = \mu = 0.0001$ and $u_N = u = 0.04$, the same parameters as we used for the simulation of the process. To avoid complications about the spatial scaling, we set simply set $\delta_N = 1$. Choosing $N = 2.8$ led to the observed alignment of simulation and prediction.

B.2 Properties of g^N

B.2.1 Boundedness

Proof of Lemma 4.6.1. Similar to [FP17, Lemma 4.3], we can represent the solution as

$$g^N_{s,t}(x,k) = G^N_{t-s} * \phi(x,k) + \int_s^t G^N_{u-s} * \Big(\mathcal{R}_{m^N_u}(g^N_{u,t}) - \mu g^N_{u,t} + \mu \mathcal{H} g^N_{u,t} + uV_R \overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t}}}\Big)(x,k)du,$$

again with G^N being the semigroup associated to $\mathcal{L}^N$. The inequality $\|G^N_{t-s} * \phi\|_p \leq \|\phi\|_p$ for functions $\phi \in \mathcal{D}(\mathbb{T}^d)$ extends to our functions $\phi \in \mathcal{D}(\mathbb{T}^d \times [0,1])$ with a

type component and the extended p-norms of (4.2)

$$\begin{aligned}\|G^N_{t-s} * \phi\|_p &= \left(\int_{\mathbb{T}^d} \sup_{k\in[0,1]} \left| \int_{\mathbb{T}^d} G^N_{t-s}(z-y)\phi(y,k)dy \right|^p dz \right)^{1/p} \\ &\leq \left(\int_{\mathbb{T}^d} \sup_{k\in[0,1]} \int_{\mathbb{T}^d} G^N_{t-s}(z-y) \left|\phi(y,k)\right|^p dydz \right)^{1/p} \\ &\leq \left(\int_{\mathbb{T}^d} \int_{\mathbb{T}^d} G^N_{t-s}(z-y) \sup_{k\in[0,1]} \left|\phi(y,k)\right|^p dydz \right)^{1/p} \\ &= \left(\int_{\mathbb{T}^d} \sup_{k\in[0,1]} \left|\phi(y,k)\right|^p dy \right)^{1/p} = \|\phi\|_p.\end{aligned}$$

Applying $(a+b)^2 \leq 2a^2 + 2b^2$ and Jensen's inequality again

$$\begin{aligned}\left\|g^N_{s,t}\right\|^p_p &\leq (2(t-s))^{p-1} \int_s^t \Big\| uV_R \overline{r(\overline{m}^N_t)\overline{g^N_{u,t}}} - \mu g^N_{u,t} \\ &\qquad + \mu \mathcal{H} g^N_{u,t} + uV_R \overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t}}} \Big\|^p_p du + 2^{p-1} \left\|\phi\right\|^p_p \\ &\leq (2(t-s))^{p-1} \int_s^t \left(uV_R r_{\max} + 2\mu + uV_R n_{\max} \sup_{z\in\mathbb{T}^d} \|r'_z\|_\infty \right) \left\|g^N_{s,t}\right\|^p_p du \\ &\qquad + 2^{p-1} \left\|\phi\right\|^p_p.\end{aligned}$$

We also used that $\|\overline{\phi}\|_p \leq \|\phi\|_p$ and that, as $[0,1]$ is finite,

$$\left|\mathcal{H} g^N_{u,t}\right| \leq \sup_{k\in[0,1]} \left|g^N_{u,t}\right|.$$

Grownwall's inequality allows us to conclude the statement and an inductive argument shows the generalization to the derivatives. □

B.2.2 Convergence

Lemma B.2.1. *Let m^N be the sequence of deterministic approximate population sizes of* (4.16) *converging to the solution n of the reaction-diffusion equation* (4.13) *by Lemma 4.4.1. Let $\phi : \mathbb{T}^d \to \mathbb{R}$ be a four times continuously-differentiable function and assume $\max_{|\kappa|\leq 4} \|\partial_\kappa \phi\|_p < \infty$ for $1 \leq p \leq \infty$. Then there exists constants $C_1, C_2 > 0$ such that*

$$\left\| \left(\mathcal{L}^N + \mathcal{R}^N_{m^N_t} \right) \phi \right\|_p \leq C_1 \max_{|\kappa|\leq 2} \|\partial_\kappa \phi\|_p \tag{B.1}$$

and

$$\left\|\left(\mathcal{L}^N+\mathcal{R}^N_{m_t^N}\right)\phi-uV_R\Big(\frac{R^2}{d+2}\Delta+r(n_s)\Big)\phi\right\|_p\leq\delta_N C_2. \tag{B.2}$$

Proof. From ([FP17, Proposition A.1]) we know

$$\left\|\mathcal{L}^N\phi-uV_R\frac{R^2}{d+2}\Delta\phi\right\|_p\leq uV_R R^4\frac{d^3}{3}(\delta_N)^2\max_{|\kappa|\leq 4}\|\partial_\kappa\phi\|_p$$

and

$$\left\|\overline{\phi}(\cdot,\delta_N R)-\phi\right\|_p\leq\frac{d}{2}(\delta_N R)^2\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p. \tag{B.3}$$

The growth operator $\mathcal{R}^N_{m_t^N}$ was defined as

$$\begin{aligned}\mathcal{R}^N_{m_t^N}\phi(z)&=uV_R\frac{1}{V_{\delta_N R}}\int_{B(z,\delta_N R)}\frac{1}{V_{\delta_N R}}\int_{B(x,\delta_N R)}r_x(\overline{m_s^N}(x,\delta_N R))\phi(y)dydx\\&=uV_R\overline{r(\overline{m_t^N})\overline{\phi}}(\cdot,\delta_N R).\end{aligned}$$

Adding and substracting terms, we obtain

$$\begin{aligned}&\left\|\overline{r(\overline{m_t^N})\overline{\phi}}(\cdot,\delta_N R)-r(n_t)\phi\right\|_p\\&\leq\left\|\overline{r(\overline{m_t^N})\overline{\phi}}(\cdot,\delta_N R)-\overline{r(\overline{n_t})\overline{\phi}}(\cdot,\delta_N R)\right\|_p+\left\|\overline{r(\overline{n_t})\overline{\phi}}(\cdot,\delta_N R)-r(n_t)\phi\right\|_p\\&\leq\left\|\Big[r(\overline{m_t^N})-r(\overline{n_t})\Big]\overline{\phi}(\cdot,\delta_N R)\right\|_p+\left\|\overline{r(\overline{n_t})\overline{\phi}}(\cdot,\delta_N R)-r(\overline{n_t})\overline{\phi}\right\|_p\\&\qquad+\left\|r(\overline{n_t})\overline{\phi}-r(\overline{n_t})\phi\right\|_p+\left\|r(\overline{n_t})\phi-r(n_t)\phi\right\|_p\\&\leq\sup_{z\in\mathbb{T}^d}\|r'_z\|_\infty\left\||\overline{m_t^N}-\overline{n_t^N}|\overline{\phi}\right\|_p+\frac{d}{2}(\delta_N R)^2\max_{|\kappa|=2}\left\|\partial_\kappa\Big(r(\overline{n_t})\overline{\phi}\Big)\right\|_p\\&\qquad+r_{\max}\frac{d}{2}(\delta_N R)^2\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p+\sup_{z\in\mathbb{T}^d}\|r'_z\|_\infty\left\|\left|\overline{n_t}-n_t\right|\phi\right\|_p\\&\leq\sup_{z\in\mathbb{T}^d}\|r'_z\|_\infty\delta_N C\|\phi\|_p+\frac{d}{2}(\delta_N R)^2\max_{|\kappa|=2}\left\|\partial_\kappa\Big(r(\overline{n_t})\overline{\phi}\Big)\right\|_p\\&\qquad+r_{\max}\frac{d}{2}(\delta_N R)^2\max_{|\kappa|=2}\|\partial_\kappa\phi\|_p+\sup_{z\in\mathbb{T}^d}\|r'_z\|_\infty\,\delta_N R\max_{|\kappa|=1}\|\partial_\kappa n_t\|_\infty\|\phi\|_p,\end{aligned}$$

where we used Lemma 4.4.1. □

Lemma B.2.2. *For any twice-continuously differentiable function $\phi:\mathbb{T}^d\to\mathbb{R}$ there exists a constant $C>0$ such that*

$$\left\|uV_R\overline{\overline{m_u^N}r'(\overline{m_u^N})\overline{\phi}}-uV_R n_u r'(n_u)\phi\right\|_p\leq\delta_N C.$$

Proof. We can write

$$\begin{aligned}
&\Big\|\overline{\overline{m}_u^N r'(\overline{m}_u^N)\overline{\phi}} - \overline{n_u r'(\overline{m}_u^N)\overline{\phi}}\Big\|_p + \Big\|\overline{n_u r'(\overline{m}_u^N)\overline{\phi}} - \overline{n_u r'(n_u)\overline{\phi}}\Big\|_p \\
&\qquad + \Big\|\overline{n_u r'(n_u)\overline{\phi}} - \overline{n_u r'(n_u)\phi}\Big\|_p + \Big\|\overline{n_u r'(n_u)\phi} - n_u r'(n_u)\phi\Big\|_p \\
&\leq \Big\|\overline{(\overline{m}_u^N - n_u) r'(\overline{m}_u^N)\overline{\phi}}\Big\|_p + \Big\|\overline{n_u(r'(\overline{m}_u^N) - r'(n_u))\overline{\phi}}\Big\|_p \\
&\qquad + \Big\|\overline{n_u r'(n_u)(\overline{\phi} - \phi)}\Big\|_p + \Big\|\overline{n_u r'(n_u)\phi} - n_u r'(n_u)\phi\Big\|_p \\
&\leq \delta_N C \sup_{z\in\mathbb{T}^d} \|r_z'\|_\infty \|\phi\|_p + n_{\max} \sup_{z\in\mathbb{T}^d} \|r_z''\|_\infty \delta_N C \|\phi\|_p \\
&\qquad + n_{\max} \sup_{z\in\mathbb{T}^d} \|r_z'\|_\infty \frac{d}{2}(\delta_N R)^2 \max_{|\kappa|=2} \|\partial_k \phi\|_p + \frac{d}{2}(\delta_N R)^2 \max_{|\kappa|=2} \Big\|\partial_\kappa (n_u r'(n_u)\phi)\Big\|_p.
\end{aligned}$$

We used Lemma 4.4.1 for the first and second term and the inequality on $\|\overline{\phi} - \phi\|_p$ from the previous lemma for the third and fourth term. □

Proof of Lemma 4.6.2. To show the convergence, we use the representation

$$\begin{aligned}
g_{s,t} &= G_{t-s} * \phi(x) + \int_s^t G_{u-s} * \Big(uV_R r(n_u)(g_{u,t}) - \mu g_{u,t} \\
&\qquad + \mu \mathcal{H} g_{u,t} + uV_R n_u r'(n_u) \mathcal{H} g_{u,t} \Big) du, \\
g_{s,t}^N &= G_{t-s} * \phi(x) + \int_s^t G_{u-s} * \Big(\mathcal{L}^N(g_{u,t}^N) - uV_R \frac{R^2}{d+2} \Delta g_{u,t}^N + \mathcal{R}_{m_s^N}(g_{u,t}^N) \\
&\qquad - \mu g_{u,t}^N + \mu \mathcal{H} g_{u,t}^N + uV_R \overline{\overline{m}_u^N r'(\overline{m}_u^N) \overline{\mathcal{H} g_{u,t}^N}} \Big) du.
\end{aligned}$$

Here, G_t denotes the Gaussian kernel associated to the Laplacian $uV_R \frac{R^2}{d+2}\Delta$. With the help of Lemma B.2.1 and Lemma B.2.2, we obtain

$$\begin{aligned}
&\Big\| g_{s,t}^N - g_{s,t} \Big\|_p \\
&\leq \int_s^t \Big\| \mathcal{L}^N(g_{u,t}^N) - uV_R \frac{R^2}{d+2} \Delta g_{u,t}^N + \mathcal{R}_{m_s^N}(g_{u,t}^N) - uV_R r(n_u)(g_{u,t}) \\
&\qquad - \mu\Big(g_{u,t}^N - g_{u,t}\Big) + \mu\Big(\mathcal{H} g_{u,t}^N - \mathcal{H} g_{u,t}\Big) \\
&\qquad + uV_R \overline{\overline{m}_u^N r'(\overline{m}_u^N) \overline{\mathcal{H} g_{u,t}^N}} - uV_R n_u r'(n_u) \mathcal{H} g_{u,t} \Big\|_p du \\
&\leq (t-s)\delta_N C + \int_s^t \Big\| uV_R r(n_u)(g_{u,t}^N) - uV_R r(n_u)(g_{u,t}) - \mu\Big(g_{u,t}^N - g_{u,t}\Big) \\
&\qquad + \mu\Big(\mathcal{H} g_{u,t}^N - \mathcal{H} g_{u,t}\Big) + uV_R n_u r'(n_u) \mathcal{H} g_{u,t}^N - uV_R n_u r'(n_u) \mathcal{H} g_{u,t} \Big\|_p du \\
&\leq (t-s)\delta_N C + \Big(uV_R (\sup_{z\in\mathbb{T}^d} \|r\|_\infty + n_{max} \sup_{z\in\mathbb{T}^d} \|r_z'\|_\infty) + 2\mu \Big) \int_s^t \Big\| g_{u,t}^N - g_{u,t} \Big\|_p du.
\end{aligned}$$

Again Grownwall's inequality yields the desired result. □

B.2.3 Continuity in the second time component

Proof of Lemma 4.6.3. This proof is similar to [FP17, Lemma 4.9]. As in Lemma 4.6.5, we extend g^N to $s,t \in [0,T]$ by setting

$$g^N_{s,t}(x,k) = g^N_{s\wedge t,t}(x,k),$$

and assume $t' > t \geq s$. We can then write

$$\begin{aligned}
&g^N_{s,t}(x,k)\\
&= G^N_{t-s} * \phi(x,k)\\
&\quad + \int_s^T G^N_{u-s} * \Big(\mathcal{R}^N_{m^N_u}(g^N_{u,t}) - \mu g^N_{u,t} + \mu\mathcal{H} g^N_{u,t} + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t}}}\Big)(x,k)du\\
&\quad - \int_t^T G^N_{u-s} * \Big(\mathcal{R}^N_{m^N_u}(\phi) - \mu\phi + \mu\mathcal{H}\phi + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H}\phi}}\Big)(x,k)du.
\end{aligned}$$

This transforms the difference into

$$\begin{aligned}
&\big(g^N_{s,t'} - g^N_{s,t}\big)(x,k)\\
&= G^N_{t'-s} * \phi(x,k) - G^N_{t-s} * \phi(x,k)\\
&\quad + \int_s^T G^N_{u-s} * \Bigg[\mathcal{R}_{m^N_u}(g^N_{u,t'}) - \mu g^N_{u,t'} + \mu\mathcal{H} g^N_{u,t'} + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t'}}}\\
&\qquad\quad - \mathcal{R}_{m^N_u}(g^N_{u,t}) - \mu g^N_{u,t} + \mu\mathcal{H} g^N_{u,t} + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t}}}\Bigg](x,k)du\\
&\quad + \int_t^{t'} G^N_{u-s} * \Big(\mathcal{R}_{m^N_u}(\phi) - \mu\phi + \mu\mathcal{H}\phi + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H}\phi}}\Big)(x,k)du.
\end{aligned}$$

We continue similar as in the proof of Lemma 4.6.1

$$\begin{aligned}
&\Big\|g^N_{s,t'} - g^N_{s,t}\Big\|^p_p\\
&\leq 3\Big\|G^N_{t'-s} * \phi - G^N_{t-s} * \phi\Big\|^p_p\\
&\quad + 3^{p-1}(T-s)^{p-1}\int_s^T \Big\|\mathcal{R}_{m^N_u}(g^N_{u,t'}) - \mu g^N_{u,t'} + \mu\mathcal{H} g^N_{u,t'} + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t'}}}\\
&\qquad\quad - \mathcal{R}_{m^N_u}(g^N_{u,t}) - \mu g^N_{u,t} + \mu\mathcal{H} g^N_{u,t} + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H} g^N_{u,t}}}\Big\|^p_p du\\
&\quad + 3^{p-1}(t'-t)^{p-1}\int_t^{t'} \Big\|\mathcal{R}_{m^N_u}(\phi) - \mu\phi + \mu\mathcal{H}\phi + uV_R\overline{\overline{m}^N_u r'(\overline{m}^N_u)\overline{\mathcal{H}\phi}}\Big\|^p_p du.
\end{aligned}$$

The bound on $\|\overline{\phi} - \phi\|_\infty$ from (B.3), yields for the first term

$$\begin{aligned}\left\|G^N_{t'-s} * \phi - G^N_{t-s} * \phi\right\|_p^p &= \left\|\int_t^{t'} G^N_{u-s} * \mathcal{L}^N \phi du\right\|_p^p \\ &= (t'-t)^{p-1} \int_t^{t'} \int_{\mathbb{T}^d} \sup_{k\in[0,1]} \left|\mathcal{L}^N \phi(x,k)\right|^p dx du \\ &\leq C(t'-t)^p Vol(\mathbb{T}^d)\Big(\sup_{k\in[0,1]} \max_{|\beta|=2} \|\partial_\beta \phi(\cdot,k)\|_\infty\Big)^p.\end{aligned}$$

Both the second and the third summand can be bounded similar to the previous lemma resulting in

$$\begin{aligned}&\left\|g^N_{s,t'} - g^N_{s,t}\right\|_p^p \\ &\leq 3^{p-1}C(t'-t)^p + 3^{p-1}(t'-t)^p(uV_R(r_{\max} + n_{max}\|r'\|_\infty) + 2\mu)\left\|\phi\right\|_p^p \\ &\quad + 3^{p-1}(T-s)^{p-1}\Big(uV_R(\sup_{z\in\mathbb{T}^d}\|r\|_\infty + n_{max}\sup_{z\in\mathbb{T}^d}\|r'_z\|_\infty) + 2\mu\Big)\int_s^T \left\|g_{u,t'} - g_{u,t}\right\|_p^p du.\end{aligned}$$

A final application of Grönwall's inequality completes this section. □