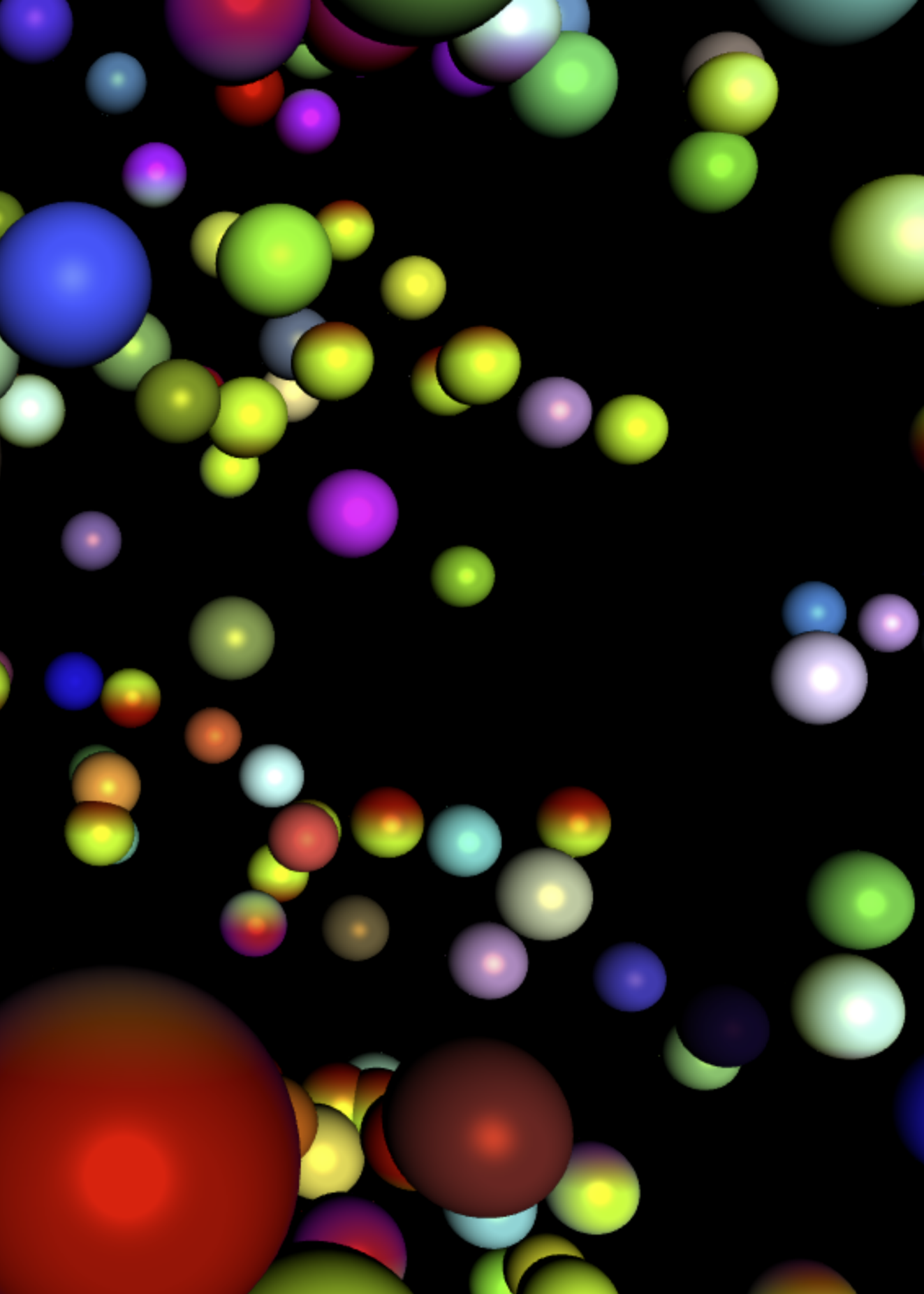
Dormancy in stochastic interacting systems

Nandan, S.



CHAPTER 1
Introduction

The present thesis consists of two parts. Part I focusses on the study of a particular class of *interacting particle systems* that describe genetic evolution of spatially structured populations with *seed-banks*. Part II focusses on the study of the *hydrodynamic scaling limit* of three interacting particle systems that incorporate *dormancy* and on the analysis of their non-equilibrium behaviour in the presence of boundary reservoirs.

§1.1 Introduction to Part I

Probability theory is the area of mathematics that aims at understanding the intrinsic stochastic nature of real-world phenomena by means of the abstract language of mathematics. Within this area, population genetics takes a special place because it brings together mathematics and biology. The primary goal of mathematical population genetics is to understand via tailored mathematical models how evolutionary forces, demographic factors, etc., affect the genealogy and frequency distribution of genotypes in biological populations.

We give a brief overview of the basic concepts that are central to understanding of the genetic evolution of a population in Section 1.1.1. We borrow from [82, 45].

§1.1.1 Bits and pieces of genetics

Among the numerous factors that contribute to the evolution of a population *resampling, mutation, natural selection, recombination* and *migration* play a central role.

Resampling (or reproduction, in which individuals transfer their gene type to future generations) is the most basic biological activity of almost any living organism. A biologist would prefer to use the word "random genetic drift" to describe the evolutionary effect of resampling in a *panmictic* population, where every individual is equally likely to be the parent of an offspring. Many populations, such as humans, birds, etc., do not seem to exhibit panmixia when mating is categorised on the basis of certain phenotypical characteristics only, but they often appear to do so when the traits under investigations are genotypes [82]. Therefore the assumption of a population being panmictic (or homogeneously mixing), which we adopt throughout this thesis, is reasonable in many circumstances. Resampling (or random mating) in a population is a source of stochasticity that pervades the gene pool of subsequent generations. It induces random fluctuations of various genotype frequencies in a natural way and drives the population towards forming a *homozygous gene pool*, i.e., a gene pool containing only a single genotype.

Mutation introduces novel gene types into a population. It is the molecular equivalent of errors that typically occur when humans carry out complex activities. In the process of replication of genetic material during resampling, spontaneous local changes may occur in the allelic composition of genes. These errors in the reproduction of genetic material give rise to different genotypes. Mutations can also occur during the reparation of damaged cells. Both beneficial and deleterious mutations are rare, but usually have significant evolutionary effects on the population.

The concept of natural selection in evolutionary theory was introduced by Charles Darwin in the mid 19th century. Selection is a force of nature that acts as a further propellant in creating a homozygous gene pool, containing only the *advantageous* genotypes of a population. Under the influence of selection, fitter types in a population have certain advantages while competing for inheritance, and cause the population to adapt more efficiently to environmental changes over time.

Recombination is a phenomena observed in populations consisting of *diploid* individuals. Diploid individuals carry two copies of genetic material in their cells instead

of one (the latter occurs in populations consisting of *haploid* individuals). Offspring of diploid populations, such as humans, have two parents. During reproduction, instead of inheriting a single identical copy of the genetic material from each parent, they inherit a recombined version in which the two copies of the parental genes undergo molecular changes via exchange of material between them. Therefore, even though a recombination event affects the genotype frequencies in an offspring, it does not alter the overall frequency of the alleles that constitute a specific genotype. In this thesis, we will only be concerned with the evolutionary behaviour of haploid populations where recombination is of no relevance.

The demography of populations is in general structured, in the sense that they admit a *carrying capacity* imposed by the surrounding habitat. Even biological cells always arrange themselves in a certain spatial order and this affects the transfer of genetic material. In population genetics the term migration, or more precisely, *migration of genetic material* is therefore construed in the broadest possible sense. A major goal of population genetics is to gain a better understanding of the effect of population structure on evolutionary quantities, such as the heterozygosity in a population, the fixation probability, (i.e., the likelihood of a specific genotype overtaking an entire population) etc. In this thesis, we analyse the role of migration in structured populations with varying capacities. For this purpose, we only consider a *conservative* migration that preserves the local population sizes. This particular choice of migration may not seem the most sensible from a pragmatic point of view. However, as we will see later, the assumption of conservative migration allows for a considerable simplification of the underlying mathematics.

In recent years, researchers in population genetics have started to analyse populations with a *seed-bank* in which individuals temporarily become dormant. Dormancy refers to the ability of an organism to enter into a reversible state of reduced metabolic activity in response to adverse environmental conditions. In the dormant state, organisms refrain from reproduction, and other phenotypic development, until they become active again. While dormancy is a trait found mostly in microbial populations, the natural analogue of dormancy in plant populations is the suspension of seed germination in difficult ecological circumstances. Several experiments suggest that populations exhibiting dormancy have better heterogeneity, survival fitness and resilience [149, 157]. Dormancy appears to be ubiquitous to many forms of life, and to be an important evolutionary trait [109, 142]. The direct effect of this trait is not easily detected when viewed on the evolutionary time scale. Various attempts have been made to better understand dormancy from a mathematical perspective (see e.g. [108, 18] for a broad overview).

§1.1.2 Mathematics of evolution

Now that the reader knows what basic evolutionary biology is all about and what it consists of, we shift our focus towards the mathematical aspects. In this thesis, we only deal with stochastic models of genetic evolution that incorporate resampling, migration and dormancy in spatially structured populations. For models that include other evolutionary forces, such as mutation, selection, recombination, etc., we refer

the reader to [153, 124, 87, 57, 5].

Fisher-Wright model. The mathematics of population genetics starts with the pioneering works of Fisher [61], Wright [153, 154] and Moran [123, 124]. Fisher and Wright introduced a classical model – later called the Fisher-Wright model – that describes the evolution of a panmictic population of constant and finite size under the sole influence of resampling. In this model, the offspring of the population follows a multinomial sampling distribution, reflecting the panmictic nature of the population, and the offspring replaces the entire parent generation at discrete instants of time. Under the model dynamics, each offspring inherits the genetic type of an arbitrarily chosen parent and the total number of offspring produced in each generation is the same as the total size of the parent population. Therefore, in this model the initial population size is conserved over time. This type of modelling is suitable for seasonally breeding small populations, such as plants, animals, etc., with a fixed average life span, in which the successive generations are non-overlapping. The Fisher-Wright model is computationally intensive, but it encompasses evolutionary behaviour of haploid populations as well as diploid populations. The application of the model to diploid populations is valid only if the population is panmictic and monoecious (such as plant populations where self-fertilisation can occur) with size $2N$, where $N \in \mathbb{N}$ is interpreted as the true size of the population.

Moran model. In many biological populations, such as microbes, humans, etc., the assumption of non-overlapping generations breaks down and evolution takes place in continuous time. In such scenarios, discrete-time mathematical models do not approximate the evolutionary behaviour of the population well enough and a need for continuous-time models arises. In 1958 Moran introduced a mathematical model [123] – later called the *Moran model* – that is a continuous-time birth and death process with finite state space, and describes the genetic evolution of a panmictic haploid population with finite size. This model, although less popular among biologists, retains all the basic qualitative features of the Fisher-Wright model. Moreover, one advantage in working with this model is that it is analytically more tractable. In this thesis, the Moran model will serve as the primary building block for the modelling of resampling, migration and dormancy in spatially structured populations. Therefore it is useful to take a closer look at its ingredients.

In the Moran model, one considers a finite population of $N \in \mathbb{N}$ reproductively (via resampling) active haploid individuals. Each individual initially carries a genotype that comes from the gene pool or type space (the collection of all potential genotypes) of the population. For simplicity, we assume that the type space contains only two genotypes, say ♡ and ♠. Models that deal with populations having infinitely many genotypes are known as Fleming-Viot processes (see e.g. [42]) and will not be considered here. According to the Moran dynamics, the population evolves over time via resampling as follows (see Fig. 1.1):

- Each individual in the population carries a *resampling clock* that rings after a random time with exponential distribution of mean 2. When the clock rings, the

individual chooses a parent from the N individuals (possibly itself) uniformly at random and *adopts* its type.

An equivalent and perhaps more natural description is:

- Each (unordered) *pair* of individuals in the population carries a clock that rings after a random time with exponential distribution of mean 2. When a clock rings, one of the two individuals gives birth to an offspring and the other individual dies.

We will stick to the former description of the model because it is mathematically more convenient. Note that the individuals are assumed to have an equal birth and death rate, which, similarly as in the Fisher-Wright model, forces the total population size to remain constant over time. Also observe that the rate of resampling is chosen to be $\frac{1}{2}$. This choice is made only to make the Moran model run at the same time scale as the Fisher-Wright model and has no other reasoning behind it. In population genetics, one is usually interested in the collective behaviour of an evolving population in which the genotypic information on a specific individual hardly matters. Because of this, we may choose not to label each individual of the population and instead to focus on the genetic configuration of the population as a whole. Since the individuals carry one

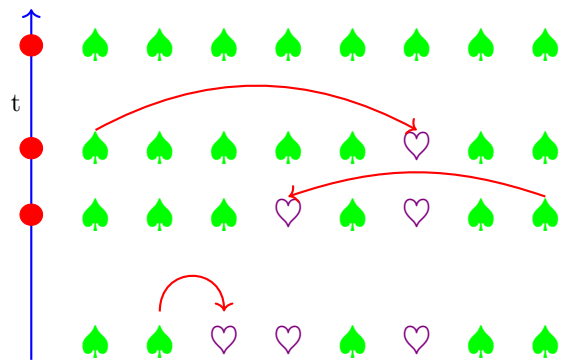


Figure 1.1: A schematic representation of a haploid population evolving under Moran dynamics. Individuals carry one of two types: ♡ and ♠. Red dots in the continuous time line stand for a resampling event. The arrows indicate simultaneous birth and death event involved in a pair of individuals.

of two genotypes and the population size is a conserved quantity, we need only one variable in order to fully specify the overall genetic evolution of the finite population, namely, the *number* of individuals having a particular type. Let us denote by $X(t)$ the number of type-♡ individuals in the population at time t. Since the time lapses between successive resampling events are assumed to be exponentially distributed and the population is panmictic, we see that

$$z := (X(t))_{t\geq 0} \tag{1.1}$$

becomes a continuous-time Markov process with state space $[N] := \{0, 1, \ldots, N\}$. As is the case with any Markov process, the time evolution of z is characterised by

its behaviour in an infinitesimal time interval. Mathematically, this information is contained in the so-called *infinitesimal generator* of the process z. As the concept of the generator of a Markov process will be frequently exploited in various parts of this thesis, we briefly elaborate on the connection between an infinitesimal generator and a Markov process. In order to avoid technicalities, we will skip the subtleties behind the estimation.

Markov generator. The generator of a time-homogeneous Feller Markov process $Z := (Z_t)_{t\geq 0}$ is a linear operator G defined on a suitable dense subspace $dom(G)$ (referred to as the domain of G) of a Banach space V (a normed and complete linear space) containing functions (often assumed to be continuous and bounded) on the state space $\mathcal{X}$ of the process Z, which often is an uncountable set. Due to the Markovian nature of Z, the canonical law of Z is determined by the family of one-dimensional distributions $(\mu_t)_{t\geq 0}$, where μ_t is the distribution of Z_t. These one-dimensional distributions, in turn, can be fully characterised by a one-parameter family of linear contraction operators – the so-called *semigroup* $(S_t)_{t\geq 0}$ associated with Z – that are defined on V. The relation between the contractions $(S_t)_{t\geq 0}$ and the distributions $(\mu_t)_{t\geq 0}$ comes from a *topological duality* (cf. [112, Theorem 1.5, Chapter I]) and is given by

$$\int_{\mathcal{X}} S_t f \, \mathrm{d}\mu_0 = \int_{\mathcal{X}} f \, \mathrm{d}\mu_t, \quad f \in V,\, t \geq 0. \tag{1.2}$$

In particular, taking μ_0 to be the Dirac distribution concentrated at $z \in \mathcal{X}$, we see that

$$(S_t f)(z) = \mathbb{E}_z[f(Z_t)], \quad f \in V,\, t \geq 0, \tag{1.3}$$

where $\mathbb{E}_z$ denotes the expectation taken w.r.t. the law of Z started at z. Therefore, constructing the canonical law of Z is equivalent to specifying the semigroup $(S_t)_{t\geq 0}$. This is where the densely defined linear operator G becomes relevant.

In order to construct the semigroup $(S_t)_{t\geq 0}$ of the Markov process Z, one can appeal to the Hille-Yosida theorem (cf. [58, Theorem 2.6]), which provides a necessary and sufficient criterion on G to *generate* the semigroup. Alternatively, one can obtain the associated semigroup by formulating a *well-posed Martingale Problem* for the generator G (cf. [112, Section 5, Chapter I]). In this thesis we will adopt the latter approach in order to extend the Moran model to the context of spatially structured populations. The generator and the semigroup are related by

$$Gf = \lim_{t\downarrow 0} \frac{S_t f - f}{t}, \quad f \in dom(G) \subseteq V, \tag{1.4}$$

where the above convergence is in the chosen Banach space V. In general, it is not easy to specify the full domain $dom(G)$ of the generator G explicitly. However, if the state space $\mathcal{X}$ of the process Z is a countable set equipped with the discrete topology, then in most situations both $dom(G)$ and V can be taken as $\mathcal{F}_b(\mathcal{X})$, the space of all bounded functions on $\mathcal{X}$ endowed with the sup norm $\|\cdot\|_\infty$, which is defined as

$$\|f\|_\infty := \sup_{z\in\mathcal{X}} |f(z)|, \quad f \in \mathcal{F}_b(\mathcal{X}). \tag{1.5}$$

In case of the Moran model, the generator G_{Mor} of the Markov process z defined in (1.1) is given by

$$G_{\mathrm{Mor}}f(x) = \tfrac{(N-x)x}{2N}[f(x+1) - f(x)] + \tfrac{x(N-x)}{2N}[f(x-1) - f(x)], \tag{1.6}$$

where $x \in [N]$ and $f \in \mathcal{F}_b([N])$. To see how this expression comes about, we can use (1.3)–(1.4) to write

$$G_{\mathrm{Mor}}f(x) = \lim_{t\to 0} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t} = \lim_{t\to 0} \mathbb{E}_x\left[\frac{f(X_t) - f(X_0)}{t}\right]. \tag{1.7}$$

In other words, $G_{\mathrm{Mor}}f(x)$ is the average infinitesimal rate of change of the observable f when the population evolving via the Moran dynamics initially contains x individuals of type ♡. In an infinitesimal time interval, the number x of type-♡ individuals increases by 1 if a type-♠ individual in the population chooses a type-♡ individual as its parent, and decreases by 1 if a type-♡ individual chooses a type-♠ individual as its parent. The former event reflects the change $[f(x+1) - f(x)]$ of the observable f and happens at a cumulative average rate $\frac{(N-x)x}{2N}$, the latter event reflects the change $[f(x-1) - f(x)]$ of the observable f and happens at the same rate $\frac{x(N-x)}{2N}$. The reason for the rates being equal can be explained as follows. Each individual in the population resamples at rate $\frac{1}{2}$. For the former event to occur, an individual of type ♠ must resample, which happens at total rate given by $\frac{(N-x)}{2}$, the number of type-♠ individuals during resampling multiplied by the resampling rate, and while resampling this individual must pick uniformly at random a type-♡ individual, which has probability $\frac{x}{N}$. Therefore, the cumulative rate of occurrence of the former event is $\frac{x(N-x)}{2N}$. A similar argument applies to the latter event.

Genealogy in the Moran model. The Moran model is particularly popular in population genetics because it is equivalent to a *birth* and *death* process, which is well understood in the Markov process literature. Many quantities of biological interests, such as the probability of two randomly chosen individuals being identical by descent, the amount of heterozygosity in the population, etc., can be explicitly computed.

Another advantage in the Moran model is that the *genealogy* (i.e., the process that tracks the ancestral lineages of individuals backwards in time) of finitely many individuals sampled from the panmictic Moran population is *exactly* governed by the so-called *Kingman coalescent process* (cf. [96]). This is in contrast to the Fisher-Wright model, where the individual ancestral lines inherit the Kingman coalescent structure only in the *large-population-size* limit, and when viewed on a time scale proportional to the size of the population. The method of analysing the evolution of a population by tracing individual genealogies all the way back to their ancestors was initiated by Kingman [96], who introduced the aforementioned coalescent process. The genealogical approach to studying evolutionary stochastic processes is now a widespread technique in population genetics. The pioneering work in [96] has in fact inspired the current development of coalescent theory that encompasses not only the Kingman coalescent, but also other coalescent processes [137, 7], such as the β-coalescent, the Λ-coalescent, etc.

The Kingman coalescent process

$$\mathcal{C} := (\mathcal{C}_t)_{t\geq 0} \tag{1.8}$$

is a continuous-time Markov process and takes values in the set of all partitions of the natural numbers. A state of the partition-valued process is a mathematical representation of the genealogical relation between individuals of a population that reproduce by resampling. In particular, a block in the time-t state $\mathcal{C}_t$ stands for an ancestor of the individuals that are alive in the evolved Moran population at time t. The individuals that descend from the ancestor specified by a block are marked by the natural numbers within the block. The coalescent process $\mathcal{C}$ evolves *backwards* in time, while the population in the Moran process evolves *forwards* in time. In the time evolution of the process, each pair of blocks in a partition coalesces at rate 1 to form a new partition containing one block less than before. This process appropriately describes the genealogy of individuals as long as we assume that the individuals reproduce independently at rate $\frac{1}{2}$ and measure time in units of length N, where N is the size of the constituent population.

In [55] (see also [138, 77]) the connection between the genealogy in a Moran population and the Kingman coalescent was established in a mathematical framework, where a "particle model" representation of an *infinite* population model is obtained via the so-called "look-down" construction. The look-down construction demonstrates that a population of any finite size evolving according to the Moran dynamics can be consistently embedded into the infinite population model (cf. [55, Lemma 2.1]). In this formulation of the Moran model, one obtains a strong (pathwise) form of *stochastic duality* between the Moran process and the Kingman coalescent process. The strong duality, in turn, implies what is known as a *weak* stochastic duality between the Moran process z and the *block-counting process*

$$\bar{\mathcal{C}} := (|\mathcal{C}_t|)_{t\geq 0} \tag{1.9}$$

associated to the Kingman coalescent process. Here, for a partition $\mathcal{R}$ of natural numbers, $|\mathcal{R}|$ denotes the number of blocks in $\mathcal{R}$. Let us elaborate a bit on the notion of weak stochastic duality and on the process $\bar{\mathcal{C}}$, as these will be central to the theme of this thesis.

Stochastic duality. The concept of weak stochastic duality relates two Markov processes in an *intertwined state*. More precisely, we say that two Markov processes $(K_t)_{t\geq 0}$ and $(L_t)_{t\geq 0}$, taking values in their respective state spaces, say Ω and $\widehat{\Omega}$, are dual to each other w.r.t. a (bounded and measurable) duality function $D(\cdot\,,\cdot) : \Omega \times \hat{\Omega} \to \mathbb{R}$ if the following intertwining relation is satisfied:

$$\mathbb{E}_k[D(K_t, l)] = \mathbb{E}^l[D(k, L_t)], \quad \forall t \geq 0,\ (k,l) \in \Omega \times \hat{\Omega}, \tag{1.10}$$

where the expectation in the left-hand (resp. right-hand) side is taken w.r.t. the law of the process $(K_t)_{t\geq 0}$ (resp. $(L_t)_{t\geq 0}$) started at $k \in \Omega$ (resp. $l \in \hat{\Omega}$). When the duality function is nice enough, the above relation can be characterised in terms of the infinitesimal generators of the two processes. In particular, one can see from [91,

Proposition 1.2] that if, for all $(k,l) \in \Omega \times \hat{\Omega}$, the two functions $D(\cdot\,,l) : \Omega \to \mathbb{R}$ and $D(k,\cdot) : \hat{\Omega} \to \mathbb{R}$ are in the domain of the infinitesimal generators $\hat{K}$ and $\hat{L}$ of the two respective Markov processes, then the relation in (1.10) holds if and only if

$$(\hat{K}D(\cdot\,,l))(k) = (\hat{L}D(k,\cdot))(l), \quad \forall (k,l) \in \Omega \times \hat{\Omega}. \tag{1.11}$$

In the case where the generators $\hat{K}$ and $\hat{L}$ are equal and belong to the same Markov process, the latter is said to be *self-dual* w.r.t. the duality function $D(\cdot\,,\cdot)$.

The notion of weak stochastic duality discussed above is very general and has developed into a powerful technique for analysing Markov processes. A sample of references for an overview on this topic is [71, 25, 24, 91]. In population genetics, weak stochastic duality between two Markov processes often originates from a strong pathwise duality, where the dual process is *graphically* constructed by looking at the original process backwards in time. This is because, in certain special situations, the original Markov process models the evolution of a biological population in such a way that the underlying genealogical process also is Markovian. These dualities are referred to as *sampling duality* relations in the literature on population genetics, because the associated duality function can be seen as a *formula* for sampling individuals from the population. The Moran model is no exception in this respect. Indeed, the process z in (1.1) is in a sampling duality relation with the block-counting process $\bar{\mathcal{C}}$ in (1.9). The duality relation, as is demonstrated in [55], comes from a strong pathwise duality between the Moran process and the Kingman coalescent process. In practice, stochastic duality is relevant in the context of Markov processes only. Let us therefore point out that $\bar{\mathcal{C}}$ is in fact a pure death Markov process with values in the set $\mathbb{N}$ of all natural numbers. The process $\bar{\mathcal{C}}$ has transition rates

$$n \mapsto n-1 \quad \text{at rate } \binom{n}{2} \mathbb{1}_{\{n \geq 2\}}, \quad n \in \mathbb{N}. \tag{1.12}$$

The sampling duality relation between the process z started at a state $x \in [N]$, $N \in \mathbb{N}$, and the block-counting process $\bar{\mathcal{C}}$ started at $n \in \mathbb{N}$ is given by

$$\mathbb{E}_x \left[\frac{\binom{X_t}{n}}{\binom{N}{n}} \mathbb{1}_{\{n \leq X_t\}} \right] = \mathbb{E}^n \left[\frac{\binom{x}{|\mathcal{C}_t|}}{\binom{N}{|\mathcal{C}_t|}} \mathbb{1}_{\{|\mathcal{C}_t| \leq x\}} \right], \quad t \geq 0, \tag{1.13}$$

where the expectation on the left-hand side is taken w.r.t. the law of the Moran process z and the expectation on the right-hand side is taken w.r.t. the law of the block-counting process $\bar{\mathcal{C}}$. In words, the above relation says: the probability that n individuals sampled from the time-t Moran population of size N have type $\heartsuit$ is the same as the probability that all ancestors identified by tracing the n sampled lineages backwards in time from time t to time 0 have type $\heartsuit$.

The weak form of the above duality conceals the embedded coalescent structure in the backwards time evolution of the lineages. Therefore it gives little insight into the dual process. However, this form of duality is more pronounced in the literature, because it allows for the possibility of *constructing* multiple duality functions and dual processes for a single Markov process. This is usually achieved by studying the so-called *Lie-algebraic* structure of the associated infinitesimal generator (see e.g.,

[71]). The general principle in the algebraic framework of duality is to express the Markov generator of the original process in terms of elementary algebraic operators that constitute some well-known Lie algebra, and perform an Ansatz via a well-chosen duality function for the construction of a dual Markov generator.

The duality relation in (1.13) is extremely useful for obtaining analytic expressions of many quantities related to the Moran process. To demonstrate just how useful the relation in (1.13) is, let us consider the problem of computing the probability that a two-type (♡ and ♠) panmictic population of size $N \in \mathbb{N}$ evolving via the Moran dynamics eventually ends up with a homozygous gene pool containing only the gene type ♡. First observe that, with probability 1, the Moran population eventually fixates to a single gene type. The reason is that the reproduction via resampling is a dissipative mechanism that causes loss of individual genetic information in the Moran population. As the total population is of finite size, only one of the two gene types survives in the long term and the entire population fixates to a single gene type. To compute the fixation probability, we first observe that the process $z = (X_t)_{t\geq 0}$ in (1.1) is a bounded Martingale that converges a.s. to one of the two absorbing states N and 0. Here, we recall that X_t is the number of type-♡ individuals in the population of fixed size N at time t. In particular, from (1.13) we see that the fixation probability (in law) to the type ♡ is given by

$$\lim_{t\to\infty} \mathbb{E}_x \left[\frac{\binom{X_t}{n}}{\binom{N}{n}} \mathbb{1}_{\{n\leq X_t\}} \right],$$

which by the duality relation in (1.13) is equal to

$$\lim_{t\to\infty} \mathbb{E}^n \left[\frac{\binom{x}{|\mathcal{C}_t|}}{\binom{N}{|\mathcal{C}_t|}} \mathbb{1}_{\{|\mathcal{C}_t|\leq x\}} \right],$$

with the expectation taken w.r.t. the block-counting process $(|\mathcal{C}_t|)_{t\geq 0}$ corresponding to the Kingman coalescent with initial state n. Since the block-counting process starting from any natural number n eventually fixates at the value 1 in the limit $t \to \infty$, the above expression equals $\frac{x}{N}$ which gives the desired fixation probability.

In [25] the Lie-algebraic method of duality is applied to the context of mathematical models in population genetics. In particular, the duality in (1.13) is retrieved from an algebraic representation of the infinitesimal generator associated to the Moran process given in (1.6) (see e.g., [25, Section 4]). In this thesis, instead of following the standard route of genealogy-tracing, we exploit the Lie-algebraic framework of duality in order to obtain two dual processes corresponding to, respectively, the single and the multi-colony Moran process with seed-banks. In the next section we extend the standard Moran model to include a *seed-bank* component that models the presence of dormancy in the population. The single-colony Moran model with seed-bank serves as the building block for the construction of the multi-colony (spatial) Moran model with seed-banks.

§1.1.3 The Moran model with seed-bank

In a stochastic individual-based model, dormancy is mathematically incorporated by turning off *resampling* for a random and possibly extended period of time. This way of modelling dormancy introduces memory, and thereby gives rise to a rich behaviour of the underlying stochastic system. The first mathematical model dealing with the effect of dormancy goes back to [34]. Since then several other ways to model seed-banks have emerged [92, 16, 14]. For example, in the model proposed in [92], the Fisher-Wright model [153] was extended to include a *weak* seed-bank, where individuals reproduce offspring several generations ahead in time, with the skipped generations being interpreted as a dormant period for the offspring. In this model the resulting genealogy of the population becomes stretched over time and retains the same coalescent structure described by the Kingman coalescent process $\mathcal{C} = (\mathcal{C}_t)_{t\geq 0}$. In [13, 12], a different qualitative behaviour was observed by including a *strong* seed-bank component, which enables the dormant individuals to have wake-up times with fat tails. A trade-off in these models was the loss of the Markov property in the time evolution of the system. This issue was partially tackled in [14], which introduced the *seed-bank coalescent*, a new class of coalescent structures that, broadly speaking, describe the genealogy of a population exhibiting extreme dormancy.

While the works mentioned above deal with seed-bank models only in the *diffusive regime*, obtained after taking the *large-colony-size-limit* of individual-based models, it is biologically more reasonable to consider seed-bank models with populations that have *finite* sizes. A natural candidate for models dealing with finite populations is the Moran model introduced earlier. In this section we extend the Moran model to include a seed-bank component that captures the effect of dormancy in the Moran population.

Single-colony Moran model with seed-bank. The seed-bank modelling in the Moran process is achieved by subdividing the constituent population of total size, say $(N + M) \in \mathbb{N}$, into two subpopulations, namely, an *active* population of size $N \in \mathbb{N}$ and a *dormant* population of size $M \in \mathbb{N}$, and turning reproduction via resampling off in the dormant population. In order to preserve the flow of gene information between the two subpopulations, we further introduce an *exchange* mechanism. More precisely, during the exchange events individuals of the active population swap places with the individuals in the dormant population. While doing so both the dormant and the active individuals keep their gene type. In this way, individuals can be either in an active state or a dormant state depending on the subpopulation they reside in. However, as the dormant individuals do not resample (i.e., do not reproduce), they cause an overall slow-down of the random genetic drift that arises from random resampling. Because of this, we refer to the dormant population as the seed-bank of the active population. A schematic description of the single-colony Moran process with seed-bank is given in Fig. 1.2. Likewise, in the Moran process without seed-bank the total sizes of the two subpopulations remain constant in time. Therefore, as long as the quantities of interest are the gene frequencies, we may describe the biological system with just two variables, namely, the number $X(t)$ of type-♡ active individuals and the number $Y(t)$ of type-♡ dormant individuals at time $t \geq 0$. In terms of mathematics, the individuals

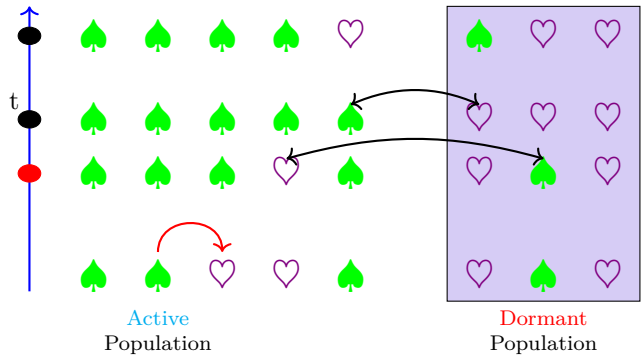


Figure 1.2: A schematic representation of a haploid population with seed-bank evolving under Moran dynamics. The active population has size $N = 5$ and the seed-bank is of size $M = 3$. Individuals carry one of two types: ♡ and ♠. Red (black) dots in the continuous timeline stand for resampling (exchange) event. The red (black) arrows indicate simultaneous birth-death (exchange) event involved in a pair of individuals.

update their gene types according to the following rules:

(1) At rate $\frac{1}{2}$ each individual in the active population resamples type.

(2) At rate $\lambda > 0$ each individual in the active population exchanges *type* with an individual chosen uniformly at random from the seed-bank.

As the reproduction and the exchange events happen at uniform rates, the process

$$z = (z(t))_{t\geq 0}, \qquad z(t) = (X(t), Y(t)), \tag{1.14}$$

forms a bivariate Markov process in continuous time. The process z lives in the state space $[N] \times [M]$ and makes the transitions

$$(x,y) \mapsto \begin{cases} (x-1,y) & \text{at rate } \frac{x(N-x)}{2N}, \\ (x+1,y) & \text{at rate } \frac{x(N-x)}{2N}, \\ (x-1,y+1) & \text{at rate } \lambda x \frac{M-y}{M}, \\ (x+1,y-1) & \text{at rate } \lambda K \frac{N-n}{N}, \end{cases} \tag{1.15}$$

where $x \in [N]$, $y \in [M]$, and $K^{-1} := \frac{M}{N}$ is the relative strength of the seed-bank. This model is in fact a continuous-time version of the two-island model introduced in [126] and allows for different sizes of the two subpopulations. While the model in [126] was analysed in the large population size limit, we keep the population size finite. By using the algebraic framework of stochastic duality we characterise the equilibrium behaviour of the model in Chapter 2. We also do the same for the multi-colony model with seed-banks, which we introduce in the next section.

Genealogy in the single-colony Moran model with seed-bank. In the single-colony Moran model with seed-bank, the *genealogy* of finitely many individuals sampled

from the two subpopulations of size N and M can be explained in terms of a partition-valued coalescent process similarly as in case of the Moran process. However, because of the addition of a seed-bank component, blocks of a partition can be in one of two states: A (active state) and D (dormant state). Recall that the genealogy in the Moran model is described by the Kingman coalescent process, where a pair of blocks of a partition coalesce at rate 1 independently of each other. This is in contrast to the genealogy in our Moran model with seed-bank. Because of the restriction to *finite size* of the active and the dormant population, and also due to the exchange mechanism involved in the two subpopulations, the independence formerly present in the coalescent structure of the genealogy is partially lost. In particular, active and dormant blocks of a partition *interact* with each other. The interaction between the blocks, or more precisely "the ancestors", appears because of the subdivision of the Moran population into two subpopulations of finite sizes, which in some way destroys the exchangeable labelling proposed in [55] of the individuals in the population. For this reason, we name the associated partition valued genealogical process an *interacting seed-bank coalescent*. To remain consistent with the previous terminologies, we will use the word lineage to refer to a block in a partition.

Let $\mathcal{P}_k$ be the set of partitions of $\{1, 2, \ldots, k\}$. For $\xi \in \mathcal{P}_k$, denote the number of lineages in ξ by $|\xi|$. Furthermore, for $j, k, l \in \mathbb{N}$, define

$$\mathcal{M}_{j,k,l} = \left\{ \vec{u} \in \{A, D\}^j : \begin{array}{l} \text{the numbers of } A \text{ and } D \text{ in } \vec{u} \\ \text{are at most } k \text{ and } l, \text{ respectively} \end{array} \right\}. \tag{1.16}$$

The state space of the genealogical process is $\mathcal{P}_{N,M} = \{(\xi, \vec{u}) : \xi \in \mathcal{P}_{N+M}, \vec{u} \in \mathcal{M}_{|\xi|,N,M}\}$. Note that $\mathcal{P}_{N,M}$ contains only those marked partitions of $\{1, 2, \ldots, N+M\}$ that have at most N active lineages and M dormant lineages. This is because we can only sample at most N active and M dormant individuals from the population.

Before we give the formal definition, let us introduce some basic notations. For $\pi, \pi' \in \mathcal{P}_{N,M}$, we say that $\pi \succ \pi'$ if π' can be obtained from π by merging two active lineages. Similarly, we say that $\pi \bowtie \pi'$ if π' can be obtained from π by altering the state of a single lineage ($A \to D$ or $D \to A$). We write $|\pi|_A$ and $|\pi|_D$ to denote the number of active and dormant lineages present in π, respectively.

Definition 1.1.1 (Interacting seed-bank coalescent). The *interacting seed-bank coalescent* is the continuous-time Markov chain with state space $\mathcal{P}_{M,N}$ characterised by the following transition rates:

$$\pi \mapsto \pi' \text{ at rate } \begin{cases} \frac{1}{N} & \text{if } \pi \succ \pi', \\ \lambda\big(1 - \frac{|\pi|_D}{M}\big) & \text{if } \pi \bowtie \pi' \text{ by a change of state of one lineage in } \pi \text{ from } A \text{ to } D, \\ \lambda K\big(1 - \frac{|\pi|_A}{N}\big) & \text{if } \pi \bowtie \pi' \text{ by a change of state of one lineage in } \pi \text{ from } D \text{ to } A, \end{cases} \tag{1.17}$$

■

where $\pi, \pi' \in \mathcal{P}_{N,M}$ and $K = \frac{N}{M}$. The factor $1 - \frac{|\pi|_D}{M}$ in the transition rate of a single active lineage when π becomes dormant reflects the fact that, as the seed-bank gets

full, it becomes more difficult for an active lineage to enter the seed-bank. Similarly, as the number of active lineages decreases due to the coalescence, it becomes easier for a dormant lineage to leave the seed-bank and become active. This also tells us that there is a *repulsive interaction* between the lineages of the same state (A or D). As the sizes N and M of the two subpopulations get large, the interaction becomes weak. In particular, as $N, M \to \infty$, after proper space-time scaling, the interacting seed-bank coalescent converges weakly to a limiting coalescent process known as the *seed-bank coalescent* [14], where interaction between the lineages is no longer present.

Single-colony block-counting process and duality. In order to obtain a sampling duality relation between the Moran model with seed-bank and the interacting seed-bank coalescent, we consider the block-counting process associated with the coalescent. If

$$\mathcal{C}_{\text{in}} := (\mathcal{C}_{\text{in}}(t))_{t\geq 0} \tag{1.18}$$

denotes the interacting seed-bank coalescent process in Definition 1.1.1, then we define by n_t (resp. m_t) the number of *active* (resp. *dormant*) lineages in the time-t state $\mathcal{C}_{\text{in}}(t)$ of the partition-valued process. We see from the definition of $\mathcal{C}_{\text{in}}$ that the process

$$z_* = (z_*(t))_{t\geq 0}, \qquad z_*(t) = (n_t, m_t), \tag{1.19}$$

also forms a continuous-time Markov process with values in $[N] \times [M] \backslash \{(0,0)\}$. The transition rates of the block-counting process are given by

$$(n,m) \mapsto \begin{cases} (n-1,m) & \text{at rate } \frac{1}{N}\binom{n}{2}\mathbb{1}_{\{n\geq 2\}}, \\ (n-1,m+1) & \text{at rate } \lambda n \frac{M-m}{m}, \\ (n+1,m-1) & \text{at rate } \lambda K m \frac{N-n}{N}, \end{cases} \tag{1.20}$$

where $(n,m) \in [N] \times [M]$ is such that $(n,m) \neq (0,0)$ and $K = \frac{N}{M}$. The first transition in (1.20) corresponds to the coalescence of two active lineages in the coalescent process $\mathcal{C}_{\text{in}}$, while the last two transitions reflect the transition of a lineage from an active (resp. dormant) state to a dormant (resp. active) state. In the sense of the earlier discussed weak stochastic duality, the block-counting process z_* is dual to the Moran process z with seed-bank given in (1.14). In particular, they satisfy the sampling duality relation given by

$$\mathbb{E}_{(x,y)}\left[\frac{\binom{X(t)}{n}}{\binom{N}{n}}\frac{\binom{Y(t)}{m}}{\binom{M}{m}}\mathbb{1}_{\{n\leq X(t),\, m\leq Y(t)\}}\right] = \mathbb{E}^{(n,m)}\left[\frac{\binom{x}{n_t}}{\binom{N}{n_t}}\frac{\binom{y}{m_t}}{\binom{M}{m_t}}\mathbb{1}_{\{n_t\leq x,\, m_t\leq y\}}\right], \quad t \geq 0, \tag{1.21}$$

where the expectation on the left-hand side is taken w.r.t. the law of the process z started at $(x,y) \in [N] \times [M]$ and the expectation on the right-hand side is taken w.r.t. the law of the process z_* started at $(n,m) \in [N] \times [M]$. The duality relation in (1.21) contains all the essential information on the process z that is needed in order to carry out an analysis of its long-time behaviour. Indeed, with the help of this relation we easily obtain the following characterisation:

Theorem 1.1.2 (Equilibrium, [Corollary 2.3.4, Chapter 2]). *Suppose that z starts from initial state $(X, Y) \in [N] \times [M]$. Then $(X(t), Y(t))$ converges in law as $t \to \infty$ to a random vector (X_∞, Y_∞) with distribution*

$$\mathcal{L}_{(X,Y)}(X_\infty, Y_\infty) = \tfrac{X+Y}{N+M}\,\delta_{(N,M)} + \Big(1 - \tfrac{X+Y}{N+M}\Big)\delta_{(0,0)}, \tag{1.22}$$

where, for $v \in [N] \times [M]$, δ_v denotes the Dirac distribution concentrated at v.

The mathematical details of the above result can be found in Section 2.3 of Chapter 2. In words, this result says that as time progresses the heterozygosity in the Moran population with seed-bank is lost and the entire gene pool fixates (in law) to one of the two types: ♡ and ♠. The probability of fixation to the all type-♡ configuration in the long run is given by $\frac{X+Y}{N+M}$, which is the initial frequency of type ♡ in the *entire* population. Thus, the addition of a seed-bank component has no significant effect on the overall qualitative behaviour of the model. However, we will see later that for the spatial model with seed-banks this is no longer the case, and seed-banks can potentially change the quantitative as well the qualitative behaviour of the model.

§1.1.4 Spatially inhomogeneous Moran model with seed-banks

All models discussed so far study the effect of dormancy in a single-colony population and are mainly concerned with the underlying genealogy in the diffusive regime. Seed-bank models dealing with *geographically structured* populations are rare, and mathematically rigorous results are still under development. Only recently, in [76] (see also [48]), single-colony seed-bank models were extended to the *spatial* setting by incorporating *migration* of individuals between different colonies. These works are concerned with structured populations having *large* sizes, where the evolution of the demographics, such as gene frequencies, etc., is primarily governed by a system of coupled stochastic differential equations. In these works, the challenge of modelling seed-banks with fat-tailed exit times is overcome by adding internal layers to the seed-banks, where *active* individuals before entering into a layer of the seed-bank acquire a *colour* that determines the wake-up time. Three different seed-bank models of increasing generality were introduced. A full description of the different regimes in the long-time behaviour of these models was obtained in [76] for the geographic space $\mathbb{Z}^d$, $d \geq 1$, whereas a multi-scale *renormalisation* analysis on the hierarchical group was carried out in [75]. Moreover, the finite-systems scheme was established [130, 74] as well (i.e., how a truncated version of the system behaves on a properly tuned time scale as the truncation level tends to infinity).

Spatially inhomogeneous Moran model with seed-banks. The novelty in the spatial model introduced in Section 2.4 is that it addresses geographically structured populations with seed-banks having preassigned *finite* sizes. Mathematically, the model is described in terms of an *interacting particle system* (see [112] for an overview) evolving in an *inhomogeneous* state space. The spatial model is the main object of our

study in Chapters 2–4 and captures the interplay of three fundamental evolutionary forces, namely, resampling, dormancy and migration, in structured populations.

Informally, we may describe the model as follows. A schematic description of the model is given in Fig. 1.3. We consider multiple colonies consisting of two subpopulations, namely, an active population and a dormant population. The colonies are labelled by the d-dimensional integer lattice $\mathbb{Z}^d$, which plays the role of a geographic space. The dormant population at colony $i \in \mathbb{Z}^d$ is called the seed-bank of the corresponding active population. As in the single-colony model, each individual in the population carries one of the two gene types: ♡ and ♠. The active and the dormant population at colony $i \in \mathbb{Z}^d$ have finite sizes given by, respectively, $N_i \in \mathbb{N}$ and $M_i \in \mathbb{N}$. With each colony $i \in \mathbb{Z}^d$ we associate the variables $(X_i(t), Y_i(t))$, with $X_i(t)$ and $Y_i(t)$, respectively, the number of type-♡ active and dormant individuals at colony i at time $t \geq 0$. The gene types of the individuals in each colony evolve over time according to the resampling and exchange dynamics described earlier in the context of the single-colony Moran model with seed-bank. To simplify our analysis in the spatial model and to be consistent with the single-colony model, we fix the *intra-colony* resampling and exchange rates at $\frac{1}{2}$ and $\lambda > 0$, respectively. In order to also introduce *interaction* between the subpopulations at different colonies, we incorporate *conservative migration* of *active* individuals. The latter is achieved by letting individuals in the active populations resample gene types not only from the active population in their own colony, but also from active populations in other colonies. In this way, the genetic information can still flow between the subpopulations at different colonies. However, the individuals themselves stay put, which results in conservation over time of the initial local population sizes $(N_i, M_i)_{i \in \mathbb{Z}^d}$.

We specify the *inter-colony* resampling rates for the active individuals by a *migration kernel* denoted by $a(\cdot\,,\cdot)$. The kernel $a(\cdot\,,\cdot)$ is an irreducible matrix of transition

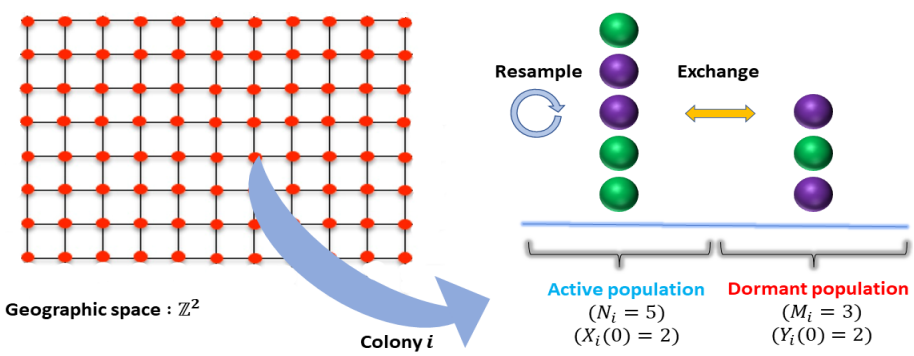


Figure 1.3: A schematic representation of the spatial populations on geographic space $\mathbb{Z}^2$ for the choice of population sizes $\mathfrak{c} := (N_k, M_k)_{k \in \mathbb{Z}^2}$. Purple individuals are of type ♡ and green individuals are of type ♠. The active (resp. dormant) population at colony i has size $N_i = 5$ (resp. $M_i = 3$). The system evolves in time under the influence of resampling and exchange.

rates whose entries are labelled by the elements in $\mathbb{Z}^d \times \mathbb{Z}^d$ and satisfies

$$a(i,j) = a(0,j-i) \quad \forall\, i,j \in \mathbb{Z}^d, \quad \sum_{i \in \mathbb{Z}^d} a(0,i) < \infty. \tag{1.23}$$

Here, $a(i,j)$ is the rate at which active individuals of colony $i \in \mathbb{Z}^d$ resample from the active population at colony $j \in \mathbb{Z}^d$. Note that our previous assumption on the intra-colony resampling rates requires us to put $a(0,0) = \frac{1}{2}$. As indicated before, the process defined by

$$Z := (Z(t))_{t \geq 0}, \quad Z(t) := (X_i(t), Y_i(t))_{i \in \mathbb{Z}^d}, \tag{1.24}$$

forms an interacting particle system taking values in the inhomogeneous configuration space

$$\mathcal{X} := \prod_{i \in \mathbb{Z}^d} [N_i] \times [M_i]. \tag{1.25}$$

The configuration $Z(t)$ specifies the gene types of the individuals in *all* the subpopulations at time t. As is typically the case for interacting particle systems, the time evolution of a single component in the configuration $Z(t)$ is not Markovian in nature. However, the configuration $Z(t)$ itself as a whole evolves in a Markovian manner. The different components of the process Z interact with each other due to the presence of the three evolutionary forces: resampling, dormancy and migration. The population sizes $(N_i, M_i)_{i \in \mathbb{Z}^d}$ and the migration kernel $a(\cdot\,,\cdot)$ are key parameters that dictate the long-run behaviour of $Z(t)$. Whereas, the intra-colony exchange rate λ only affects the time scale on which different components of the configuration $Z(t)$ evolve. Because the rate λ does not vary across colonies, it does not have a significant role in the analysis of the process Z.

Spatially interacting seed-bank coalescent. As we observed before, stochastic duality plays an important role in the analysis of models in population genetics. Duality is a formidable tool that allows one to perform exact computations in many stochastic interacting systems. Because the local population sizes in our spatial model are conserved quantities, the model has the advantage that it admits a dual process like in the single-colony Moran process with seed-bank. The underlying genealogy of the spatial model is described by a *spatially interacting structured seed-bank coalescent*. In the spatial seed-bank coalescent, lineages switch between an active and a dormant state, and perform interacting coalescing random walks on the geographic space $\mathbb{Z}^d$. To avoid technicalities, we refrain from providing a formal description of the genealogical process via partition-valued Markov chain. Our principle aim in this thesis is to characterise equilibrium behaviour of the spatial Moran process Z with the help of the dual process. For this purpose, an analysis of the block-counting process Z_* associated with the spatial seed-bank coalescent process is sufficient. We will introduce the block-counting process Z_* in the next paragraph. For the sake of completeness, we briefly describe the spatial seed-bank coalescent process via an interacting particle system (see Fig. 1.4). At each site $i \in \mathbb{Z}^d$ there are two reservoirs, an *active* reservoir and a *dormant* reservoir, with, respectively, N_i and M_i labelled locations. Each location can accommodate at most one particle. We refer to the particles in an active

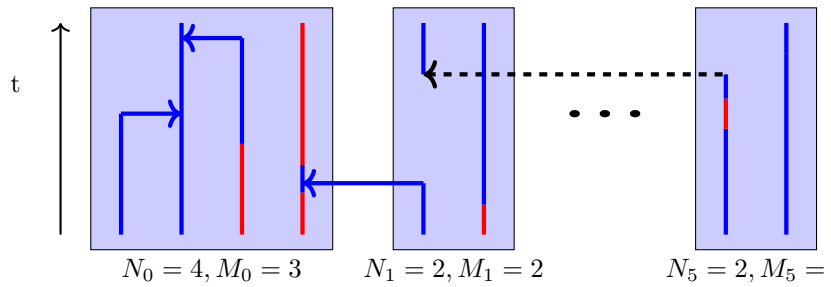


Figure 1.4: Schematic transitions of the particles in the spatially interacting structured seed-bank coalescent in dimension $d = 1$. Each block depicts the reservoirs located at sites of $\mathbb{Z}$. The blue lines represent the evolution of active particles, whereas the red lines represent the evolution of dormant particles.

and a dormant reservoir as *active* particles and *dormant* particles, respectively. The system evolves according to the following rules:

(a) An active particle at site $i \in \mathbb{Z}^d$ becomes dormant at rate λ by moving into a random labelled location (out of M_i many) in the dormant reservoir at site i when the chosen labelled location is empty, otherwise it remains in the active reservoir.

(b) A dormant particle at site $i \in \mathbb{Z}^d$ becomes active at rate λK_i with $K_i = \frac{N_i}{M_i}$ by moving into a random labelled location (out of N_i many) in the active reservoir at site i when the chosen labelled location is empty, otherwise it remains in the dormant reservoir.

(c) An active particle at site i chooses a random labelled location (out of N_j many) from the active reservoir at site j at rate $a(i, j)$ and does the following:

- If the chosen location in the active reservoir at site j is empty, then the particle moves to site j and thereby migrates from the active reservoir at site i to the active reservoir at site j.
- If the chosen location in the active reservoir at site j is occupied by a particle, then it coalesces with that particle.

Observe that an active particle can migrate between different sites in $\mathbb{Z}^d$ and two active particles can coalesce even when residing in different colonies.

Spatial block-counting process and stochastic duality. We obtain the block-counting *dual process*

$$Z_* := (Z_*(t))_{t\geq 0}, \quad Z_*(t) := (n_i(t), m_i(t))_{i\in\mathbb{Z}^d}, \tag{1.26}$$

from the spatial coalescent by counting the number of particles at each site $i \in \mathbb{Z}^d$. More precisely, we define by $n_i(t)$ (resp. $m_i(t)$) the number of active (resp. dormant) particles that are present at site $i \in \mathbb{Z}^d$ at time $t \geq 0$. Like the spatial Moran process

Z, the block-counting dual process Z_* is also an interacting particle system and takes values on the same state space $\mathcal{X}$. Under mild conditions on the *active* population sizes $(N_i)_{i\in\mathbb{Z}^d}$ and the migration kernel $a(\cdot\,,\cdot)$, a sampling duality relation can be established between the two processes Z and Z_*. In particular, if $D(\cdot\,,\cdot) : \mathcal{X}\times\mathcal{X} \to [0,1]$ is defined as

$$D((X_i,Y_i)_{i\in\mathbb{Z}^d},(n_i,m_i)_{i\in\mathbb{Z}^d}) := \prod_{i\in\mathbb{Z}^d} \frac{\binom{X_i}{n_i}}{\binom{N_i}{n_i}} \frac{\binom{Y_i}{m_i}}{\binom{M_i}{m_i}} \mathbb{1}_{\{n_i\le X_i,\, m_i\le Y_i\}} \tag{1.27}$$

with $(X_i,Y_i)_{i\in\mathbb{Z}^d},(n_i,m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}$, then one has

$$\mathbb{E}_\eta[D(Z(t),\xi)] = \mathbb{E}^\xi[D(\eta,Z_*(t))] \quad \forall t\ge 0, \tag{1.28}$$

where the expectation on the left-hand side is taken w.r.t. the law of the process Z started at $\eta\in\mathcal{X}$ and the expectation on the right-hand side is taken w.r.t. the law of the process Z_* started at $\xi\in\mathcal{X}$. The duality relation in (1.28) is very useful for analysing the spatial Moran process Z. In fact, the relation fully characterises all the mixed moments of the process Z in terms of the dual process Z_*. Even though the dual process is tricky to analyse because of the interaction in the dual particles, it is much simpler than the spatial Moran process Z.

§1.1.5 Summary of Part I

Having introduced the basic ingredients of Part I in Sections 1.1.3–1.1.4, we can now summarise the primary goals:

(1) Introduce a stochastic model that addresses genetic evolution in spatially structured populations with seed-banks whose sizes are *finite* and depend on the geographic location of the populations. Prove *existence* and *uniqueness* of the process $Z = (Z(t))_{t\ge 0}$ via well-posedness of an associated martingale problem and duality with a system of interacting coalescing random walks.

(2) The constructed process Z modelling the genetic evolution falls in the class of interacting particle systems that are Markov processes with large number of interacting components. An interesting phenomenon often observed in the long-time behaviour of such systems is the occurrence of a *phase transition*. Loosely speaking, a phase transition corresponds to an abrupt change in equilibrium behaviour as underlying model parameters cross certain critical values. In our model, the parameter controlling the phase transition turns out to be the dimension of the geographic space. In low dimensions, the invariant distributions of the model are *degenerate*, in the sense that they are concentrated on the absorbing configurations of the process. These are nothing but the two extremes of the possible gene type configurations, where either all individuals carry type ♡ or all carry type ♠. Convergence phenomena such as these are called *clustering* because locally mono-type clusters grow in the geographic space as the system approaches equilibrium. In higher dimensions, however, the model admits a *one-parameter family* of invariant distributions labelled by a continuous parameter, namely, the *average density* of a specific gene type in the population. In this

case, the system at equilibrium exhibits *coexistence*, i.e., individuals of different gene types coexist with each other. One goal in Part I is to identify a necessary and sufficient criterion for the occurrence of such a dichotomy in the equilibrium behaviour of our model.

(3) As we indicated before, the population sizes $(N_i, M_i)_{i\in\mathbb{Z}^d}$ and the migration kernel $a(\cdot\,,\cdot)$ are the primary parameters that determine the dichotomy of coexistence versus clustering in the spatial Moran process. Another goal in Part I is to identify the range of these parameter values under which the criterion for clustering versus coexistence is met.

(4) Identify the *domain of attraction* of each equilibrium in the clustering and in the coexistence regime. Here, for an equilibrium state ν of the process Z, the domain of attraction of ν is the set of all probability distributions μ such that the process Z starting from initial distribution μ converges to the equilibrium state ν as time evolves.

(5) In the clustering regime, the equilibrium states of the spatial process concentrate on homozygous gene configurations. A quantity of particular interest in this regime is the *fixation probability*, which quantifies the probability of a specific gene type, say type $\heartsuit$, taking over the entire population. If model parameters such as the population sizes are arbitrary, then standard techniques fail to provide closed-form expressions for this probability. However, as the theory of stochastic homogenisation suggests, macroscopic quantities such as the fixation probability do not feel the irregularities in the microscopic parameters when they are modelled by a *random environment*. A random environment in a stochastic model adds a second source of randomness and is typically used to capture stochastic effects in the irregularities. In most scenarios, a law of large numbers sets in and many macroscopic quantities behave similarly as those evolving in a suitably *homogenised* environment. Another goal in Part I is to see whether homogenisation occurs. The spatial model can be naturally extended to the setting of a random environment by sampling the population sizes $(N_i, M_i)_{i\in\mathbb{Z}^d}$ beforehand at random from a preassigned probability distribution. In this context, the aim is to carry out a clustering analysis for the spatial process and show that the fixation probability homogenises as the result of an appropriate 'averaging effect'.

We address the above 5 goals in the three chapters of Part I, which are based on the material of three papers on the spatial Moran process Z defined in (1.24).

Chapter 2. In this chapter we address the goals outlined in (1) and (2). To this end, we first lay out the mathematical foundations for modelling genetic evolution of *structured* and *finite* populations with *seed-banks* in stochastic settings. In particular, the main objective is to construct the spatial Moran process Z, which is a novel interacting particle system (see [112] for an overview) modelling stochastic evolution of gene types in spatially structured finite populations with seed-banks. Modelling genetic evolutions of finite populations via interacting particle systems is rare in mathematical population genetics. Most research in this area concerns only large populations,

and deal with stochastic differential equations arising from the so-called diffusive limit of individual-based models. Inclusion of seed-banks in such models is relatively new as well.

In [16, 14, 15], the continuum version of the celebrated Fisher-Wright model is modified to include a seed-bank component, and in [76, 75] this model is further extended to incorporate spatialness. However, the results in these works apply to *large* populations only. The starting point of this chapter is therefore the Moran model describing a single population of finite size.

We modify the single-colony Moran model to accommodate a suitable seed-bank component. A brief introduction of this model is given in Section 1.1.3. In order to characterise the equilibrium behaviour of this model, we utilise the recently developed theory of stochastic duality [71, 25, 24]. In particular, we identify the associated dual process by following the Lie-algebraic approach to stochastic duality described in [148]. We derive a finite-dimensional representation for the infinitesimal Markov generator of the process by viewing it as an abstract element of the $\mathfrak{su}(2)$ Lie algebra. Furthermore, by making an Ansatz with respect to a well-chosen intertwiner (i.e., the duality function), we identify the dual representation of the Markov generator, which indeed turns out to be the infinitesimal generator of a dual Markov process. This finding aligns with the general prognosis of the Lie-algebraic approach, namely, that identifying Markov generators in terms of elementary operators from a carefully chosen Lie algebra may lead to constructing new dual processes. We exploit the duality to fully characterise the equilibrium behaviour of the single-colony Moran model with seed-bank. It turns out that, despite the presence of a seed-bank component, the model qualitatively behaves as a single Moran population of finite size without a seed-bank. This is so because both the seed-bank and the reproductively active population have finite capacity.

Subsequently, we extend the single-colony model to the multi-colony Moran process $Z = (Z(t))_{t\geq 0}$ introduced in Section 1.1.4 which describes spatially structured populations of finite sizes each equipped with their own seed-bank. Using the same representations of the $\mathfrak{su}(2)$ Lie algebra, we identify the process $Z_* = (Z_*(t))_{t\geq 0}$ in (1.26) as a dual of Z. We construct the process Z by establishing well-posedness of an appropriate martingale problem, where the uniqueness of the process follows from the duality relation in (1.28). We also characterise the structural properties of the set of all invariant distributions for Z, by establishing a dichotomy between *clustering* and *coexistence*. This kind of dichotomy in the equilibrium behaviour mainly surfaces in spatial models that possess more than one *absorbing* configuration. Examples of such models include the voter model [111], the stepping stone model [140], and the model introduced in [39] addressing populations with spatial structure but no seed-banks. The same dichotomy is found in the more recent models introduced in [76, 75], which include both spatialness and seed-bank effects. Our main result in this chapter confirms that a similar dichotomy holds even when the constituent population sizes are finite and spatially varying. The dichotomy is determined by a necessary and sufficient criterion formulated in terms of the time evolution of the dual process Z_* started from two lineages (particles). In particular, the criterion says that the process Z remains in the clustering regime if and only if two dual lineages in the process Z_*

eventually coalesce with probability 1. The duality relation between Z and Z_* allows us to express the average heterogeneity in the subpopulations at time t in terms of the time-t state of two dual particles. We use this to show that the heterogeneity vanishes everywhere if and only if the two particles coalesce eventually with probability 1.

Chapter 3. In this chapter we address the goal outlined in (3). We focus on the parameter regime for which the spatial process Z exhibits clustering. From the clustering criterion given in Chapter 2, it is clear that this regime is uniquely characterised by the long-time behaviour of the dual process Z_*. In particular, eventual coalescence of two dual lineages is equivalent to the existence of a common ancestor for the spatial populations, and therefore the almost sure occurrence of this event necessarily eliminates the possibility of Z attaining a multi-type equilibrium, where individuals of different gene types can coexist.

The above scenario is common in spatial models (see e.g., [39]) where the stochastic evolution of demographics such as allele frequencies in subdivided populations are diffusively approximated. The recent results in [76] establish similar dichotomies between clustering and coexistence for three diffusively rescaled models describing spatial populations with seed-banks. It is shown that when the sizes of the seed-banks are a constant multiple of the sizes of the active populations, the dichotomy of clustering versus coexistence is solely determined by the underlying migration kernel and, apart from causing a quantitative delay in the loss of heterozygosity of the populations, seed-banks have no significant qualitative effect.

The main result in this chapter asserts that the picture remains the same for spatially structured finite populations with seed-banks of varying capacities, as long as the variations in the relative sizes of the seed-banks are of the same order. In particular, we show that if the relative sizes of the seed-banks, defined as the ratio of the dormant and the active population sizes, are uniformly bounded over the geographic space $\mathbb{Z}^d$, then the process Z clusters if and only if the symmetrised migration kernel defined by

$$\hat{a}(i,j) := \tfrac{1}{2}[a(i,j) + a(j,i)], \quad i,j \in \mathbb{Z}^d, \tag{1.29}$$

is recurrent. The last result is proven under a non-clumping criterion on the active population sizes $(N_i)_{i\in\mathbb{Z}^d}$, and the converse is proven under the stronger assumption of symmetry of the migration kernel $a(\cdot\,,\cdot)$. The non-clumping criterion imposed on the sizes $(N_i)_{i\in\mathbb{Z}^d}$ of the active populations requires that

$$\inf_{i\in\mathbb{Z}^d} \sum_{j:\|j-i\|\le R} \tfrac{1}{N_j} > 0 \tag{1.30}$$

for some $R < \infty$. This essentially says that there exists a threshold $N < \infty$ and a range $R < \infty$ such that within any finite region of the geographic space of radius R, there is at least one active population of size at most N. This criterion ensures that the time scale at which pairs of lineages in different parts of the geographic space coalesce are of the same order. We expect a close connection between the above criterion and the existence of a common ancestor of the spatial populations. We derive an alternative clustering criterion for the clustering versus coexistence dichotomy. This alternative criterion is defined in terms of almost sure absorption of an auxiliary Markov process and turns out to be easier to analyse than the original criterion.

Chapter 4. In this chapter we address the goals outlined in (4) and (5). We study the spatial process Z in the parameter regime where clustering occurs. In particular, we provide a full description of the set of all initial distributions for which the spatial process Z converges to an equilibrium.

A well-established method in the literature for studying stationary states of a Markov process having duality properties is to characterise all functions that are harmonic for an associated dual process. Depending on the complexity of the dual process and the duality function, often a full characterisation is possible. For instance, a generous use of this method is found in [111, 105], where ergodic properties of many well-known interacting particle systems are derived. Relevant examples in the context of diffusion processes arising in population genetics include [140, 39, 76, 75]. The standard technique used in this method involves constructing a successful coupling between two copies of the dual process started from two different initial states, which necessarily forces all bounded harmonic functions of the process to be constant. By leveraging the duality relation, this result is transferred to the original process, and a criterion is established that intertwines the domain attraction of each equilibrium with the set of constant harmonic functions of the dual.

In our context, it turns out that a successful coupling between two copies of the dual started from different initial configuration indeed exists when the original process Z exhibits clustering. This enables us to derive a necessary and sufficient criterion for determining the initial distributions that converge weakly to a mono-type equilibrium under the time evolution of the spatial process. This criterion is formulated in terms of the transition kernel of a single dual particle, and is valid only in the clustering regime of the original process. The criterion also characterises the fixation probability. The fixation probability quantifies the probability of a specific gene type, say type $\heartsuit$, spreading over every subpopulation at the attained equilibrium. This probability depends on how the type-$\heartsuit$ individuals are initially distributed over different subpopulations. As pointed out in goal (5), an explicit characterisation of this probability is not feasible when the model parameters are arbitrary. However, by sampling the population sizes $(N_i, M_i)_{i\in\mathbb{Z}^d}$ from a random field that is stationary and ergodic under translation, we are able to derive an expression for this probability. The formula is given in terms of an annealed average of the type-$\heartsuit$ densities in the active and the dormant population, biased by the ratio of the two population sizes at the target colony. We obtain this result under the assumption that the migration kernel $a(\cdot\,,\cdot)$ is symmetric and recurrent, and the initial frequency distribution of the type-$\heartsuit$ individuals in each colony is consistent with a global profile of the population sizes. Our results in this chapter hold only when the geographic space $\mathbb{Z}^d$ has dimension $d \leq 2$.

For the proof of the results, we make heavy use of the associated single-particle dual process. To be precise, we show that under the symmetry and recurrence assumptions on the migration kernel, the environment seen by a single dual particle in a typical random environment converges in law to an invariant distribution. Finally, by exploiting the intertwining relation between the domain of attraction of the mono-type equilibrium of Z and the transition probability kernel of a single dual particle, we lift this convergence to the spatial process Z.

§1.2 Introduction to Part II

The main motivation for studying systems of interacting particles originally comes from statistical physics. Time evolution of state variables in physical systems, such as gaseous material in a closed container, or flow of water through a pipe, etc., are complex processes governed by many parameters. An obvious attribute that is common to all such systems is the presence of small particles in large numbers. The motion of each individual particle is often subject to a local interaction rule and typically correlates to the characteristics of all nearby particles. As an outcome, the evolution equation for a single particle is no longer closed. Only a large number of coupled equations can describe the particle motion in a satisfactory manner. The disadvantage is that we lose tractability – or so it would seem in hindsight.

In a series of pioneering works [144]-[146] in the late 1960s, Spitzer initiated the study of Markov processes with locally interacting components. In subsequent years, Liggett, along with many other authors, provided a complete description of all the possible invariant measures for several such processes. These works, most of which are summarised in Liggett's monograph [112], gave birth to the novel mathematical framework of interacting particle systems, and have since developed into a prominent field of study. Within this framework, it becomes possible to rigorously describe the spatio-temporal evolution of a microscopic system that, in principle, can consist of infinitely many particles.

As explained earlier, Part I of this thesis uses the tools and methods of the interacting particle system framework. While this formulation is predominant in the context of physical systems, in Part I we draw motivations from mathematical biology instead, and utilise the framework to describe evolutionary consequences of dormancy in spatially structured populations. It is, however, not necessary to view dormancy as a trait inherent to biological systems alone. Indeed, in a physical system, dormancy may be considered as an internal state of the particles that causes hindrance to their microscopic dynamics, such as motion under the influence of a driving force. Even in chemical reactions, variation in activity levels of a reactant may be interpreted as a form of dormancy. In Part II, our principle aim is to investigate the effect of dormancy in the broader framework of interacting particle systems.

In this thesis, three such interacting particle systems are considered: the independent particle system, the exclusion process and the inclusion process. The first two systems were originally introduced by Spitzer [145], and have been treated extensively in the literature. The ergodic behaviour of these two systems is well-understood and their scaling limits are also well-known. The inclusion process, on the other hand, was introduced in [70], and its invariant distributions were fully characterised only recently [105]. Given the rich equilibrium behaviour of many mathematical population genetics models with dormancy, it is appealing to endow these three particle systems with dormant characteristics and see how they behave in the long run. Unfortunately, many of the standard techniques, such as stochastic duality, break down after the inclusion of a dormant state. Consequently, we must be careful in choosing how to incorporate dormancy. In Part II we deal with suitably crafted *multi-layer* versions of the particle systems that capture interesting phenomena under dormancy while pre-

serving the original duality properties in a natural way. Before we set the stage, in the next two sections we briefly shed light on two important aspects of interacting particle systems – hydrodynamic scaling limits (see [120] for an extensive overview) and non-equilibrium steady state behaviour – both of which are central to the analysis presented in Part II.

§1.2.1 Hydrodynamic scaling limit

The primary reason for studying interacting particle systems is to arrive at a mathematically rigorous microscopic description of the evolution of physical systems. But the usefulness lies not only in explaining microscopic properties, but also in predicting the behaviour of macroscopic observables associated with the physical system. In particular, the stochastic nature of interacting particle systems puts sophisticated probabilistic tools, such as the law of large numbers and the central limit theorem, at our disposal to elucidate emergent phenomena of physical systems in a rigorous mathematical framework. Here, physically emergent phenomena include, but are not limited to, universal laws of physics, such as Fourier's law of heat conduction or Fick's law of diffusion. The general idea behind the so-called "hydrodynamic scaling" formalism is to give a mathematically precise meaning to these emergent phenomena by exploiting various probabilistic limiting techniques and space-time scaling arguments. One may view such formalism as the transition from the microscopic world of particles to the macroscopic world of measurable observables.

In many cases, the precise choice of the type of interaction between the physical particles turns out to be irrelevant, because the emergent phenomena are often insensitive to the fine details of the microscopic laws of interaction. For example, it is possible to derive the evolution equation of heat conduction as the hydrodynamic scaling limit of both the independent particle system and the simple symmetric exclusion process. To explain the last statement, let us recall the definitions of the three particle systems, and briefly elaborate on how a suitable space-time scaling of these systems can give rise to the heat equation in the macroscopic limit.

Independent particle system. The independent particle system is a mathematical description of the time evolution of a *collection* of *indistinguishable* particles that do not influence each other in any way and move on a countable phase space $\mathcal{S}$ in a Markovian manner. For simplicity, let us fix the phase space $\mathcal{S}$ to be the integer lattice $\mathbb{Z}$. We assume that each particle performs a continuous-time simple symmetric random walk on $\mathbb{Z}$ at rate 2. Following the terminologies of the interacting particle system framework, we can specify the time-t state of such a system by a configuration

$$\eta_{\text{in}}(t) \in \mathbb{N}_0^{\mathbb{Z}}, \ \eta_{\text{in}}(t) := (\eta_{\text{in}}(i,t))_{i\in\mathbb{Z}},$$

with $\eta(i,t)$ being interpreted as the number of particles at site $i \in \mathbb{Z}$ at time $t \geq 0$. The process η_{in} defined by

$$\eta_{\text{in}} := (\eta_{\text{in}}(t))_{t\geq 0}, \quad \eta_{\text{in}}(t) = (\eta_{\text{in}}(i,t))_{i\in\mathbb{Z}}, \quad t \geq 0, \tag{1.31}$$

is the simplest example of an interacting particle system, with a Markov generator

$$(L_{\text{in}}f)(\eta) := \sum_{x\in\mathbb{Z}} \eta_x \sum_{|x-y|=1} [f(\eta^{x,y}) - f(\eta)] \tag{1.32}$$

acting on a suitable test function f and evaluated at a configuration $\eta = (\eta_x)_{x\in\mathbb{Z}}$. Here, for $x, y \in \mathbb{Z}$ and a configuration $\eta = (\eta_i)_{i\in\mathbb{Z}}$, $\eta^{x,y}$ denotes the configuration obtained from η by removing a particle from an occupied site x and putting it at site y. In other words,

$$\eta^{x,y} := (\eta_i - \mathbb{1}_{\{i=x,\eta_x\geq 1\}} + \mathbb{1}_{\{i=y,\eta_x\geq 1\}})_{i\in\mathbb{Z}}. \tag{1.33}$$

Note that, in order for the generator L_{in} to uniquely specify a Markov process, some regularity restrictions must be imposed on the initial configuration of the process. We refrain from addressing these technical subtleties here.

Simple symmetric exclusion process. While the independent particle system is a natural example, it does a poor job in modelling physical systems in which particles are interacting. Studies of even the simplest form of particle interaction can provide useful insights. The aforementioned simple symmetric exclusion process (SSEP) was introduced by Spitzer [145] as a toy model for lattice gases at infinite temperature, and has been studied extensively in the literature since. This process is obtained from the independent particle system by imposing a local interaction called *exclusion rule*: two particles are not allowed to occupy the same location. Consequently, all jumps of the independent particles leading to a violation of the exclusion rule are suppressed. For the exclusion rule to make sense, one must of course start the system at a configuration where all particles are initially at distinct locations. The resulting Markov process

$$\bar{\eta}_{\text{ex}} := (\bar{\eta}_{\text{ex}}(t))_{t\geq 0}, \quad \bar{\eta}_{\text{ex}}(t) := (\bar{\eta}_{\text{ex}}(i,t))_{i\in\mathbb{Z}}, \quad t \geq 0, \tag{1.34}$$

evolves on the state space $\{0,1\}^{\mathbb{Z}}$ and has the formal generator

$$(L_{\text{ex}}f)(\eta) = \sum_{\substack{x,y\in\mathbb{Z},\\ |x-y|=1}} \eta_x(1-\eta_y)[f(\eta^{x,y}) - f(\eta)], \tag{1.35}$$

where $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ is a cylinder function and $\eta \in \{0,1\}^{\mathbb{Z}}$.

Simple symmetric inclusion process. The simple symmetric inclusion process (SIP) introduced in [70] is the opposite analogue of the exclusion process. In this process the underlying particles interact by "inviting" the neighbouring particles to their own locations rather than driving them away. The additional interaction is superimposed onto the independent motions of the microscopic particles and the interaction strength is assumed to be linearly increasing with the number of particles at a destination site. The resulting process

$$\eta_{\text{inc}} := (\eta_{\text{inc}}(t))_{t\geq 0}, \quad \eta_{\text{inc}}(t) := (\eta_{\text{inc}}(x,t))_{x\in\mathbb{Z}}, \quad t \geq 0, \tag{1.36}$$

obtained from counting the number of particles at each site in $\mathbb{Z}$ evolving over time, therefore lives on the state space $\mathbb{N}^{\mathbb{Z}}$. The process η_{inc} is Markovian and has the formal generator

$$(L_{\text{inc}}f)(\eta) := \sum_{x\in\mathbb{Z}} \eta_x \sum_{|x-y|=1} (1+\eta_y)[f(\eta^{x,y}) - f(\eta)], \tag{1.37}$$

where f is a suitable test function and $\eta = (\eta_x)_{x\in\mathbb{Z}}$ is a configuration.

Hydrodynamics: the heat equation. With the above mathematical definitions at hand, we can now concentrate on hydrodynamic scaling. As we have already explained, hydrodynamic behaviour of a microscopic system refers to a description of how the constituent quantities evolve when viewed from a macroscopic frame of reference. In the macroscopic world, most wildly fluctuating microscopic quantities scale down to trivial states, and only a few of the many degrees of freedom survive in the form of certain conserved thermodynamic quantities such as *energy, temperature, particle density*, etc. These macroscopic quantities behave in a much smoother way than their microscopic counterparts. Depending on how one models the underlying microscopic randomness, these quantities can often be shown to satisfy a *deterministic* partial differential equation. As we will note shortly, in case of the independent particle system and the exclusion process, the partial differential equation associated with the hydrodynamic limit is nothing but the heat equation.

In order to carry out the hydrodynamic scaling, one must first renormalise space and time by suitable scaling parameters that quantify the relationship between the microscopic and the macroscopic world. In this regard, it is standard to assume that spatial distance scales linearly as one zooms out from the microscopic view to the macroscopic view. However, as time is typically measured relative to the external observer, one should also take the average spatial spread of the microscopic particles into consideration while rescaling the time parameter. In the case of particles that evolve according to the two random processes η_{in} and $\bar{\eta}_{\text{ex}}$ defined above, the average spread in time t is of order $\sqrt{t}$. This is well-known in the context of independent particle systems, but is not obvious for the exclusion process. We refer the interested reader to [119], where it is shown that the typical distance covered by a free particle and a particle subject to the exclusion rule are asymptotically of the same order. Thus, in order to visualise a non-trivial motion of the particles from the macroscopic point of view, temporal scaling should be taken quadratically proportional to the spatial scaling.

Having justified the choice for the space-time scaling parameters, we can now consider the following two measure-valued random quantities associated with the processes in (1.31) and (1.34):

$$X^N_{\text{in}}(t)(\cdot) := \frac{1}{N}\sum_{x\in\mathbb{Z}} \eta_{\text{in}}(x, N^2t)\delta_{x/N}(\cdot),\ X^N_{\text{ex}}(t)(\cdot) := \frac{1}{N}\sum_{x\in\mathbb{Z}} \bar{\eta}_{\text{ex}}(x, N^2t)\delta_{x/N}(\cdot). \tag{1.38}$$

Here, $N \in \mathbb{N}$ is the parameter quantifying the amount of dilation performed while zooming out from the microscopic world to the macroscopic world, and will eventually be set to diverge to infinity. Mathematically, X^N_{in} and X^N_{ex} describe what the empirical

distribution of the particle densities in the two microscopic processes η_{in} and $\bar{\eta}_{\text{ex}}$ look like from the macroscopic perspective. Observe that the two processes $t \mapsto X^N_{\text{in}}(t)$ and $t \mapsto X^N_{\text{ex}}(t)$ indeed take values in the space of non-negative Radon measures. In particular, it is easily seen that for any compact $A \subseteq \mathbb{R}$,

$$X^N_{\text{in}}(t)(A) = \frac{1}{N} \sum_{x \in \mathbb{Z}} \eta_{\text{in}}(x, N^2 t) \mathbb{1}_A(\tfrac{x}{N}) < \infty. \tag{1.39}$$

The hydrodynamic scaling procedure tells that the two processes $\{X^N_{\text{in}}(t) : t \geq 0\}$ and $\{X^N_{\text{ex}}(t) : t \geq 0\}$ converge to a non-trivial *deterministic* limiting process, in a probabilistic sense, as we pass from the microscopic viewpoint to the macroscopic viewpoint by letting $N \to \infty$. More precisely, the following holds:

Theorem 1.2.1 (Hydrodynamic scaling, [120, Theorem 2.8.1] and [69]). *Let $\bar{\rho} \in C_b(\mathbb{R})$ be a bounded and continuous macroscopic profile, and let $(\mu_N)_{N \in \mathbb{N}}$ (resp., $(\bar{\mu}_N)_{N \in \mathbb{N}}$) be a sequence of probability measures on $\mathbb{N}^{\mathbb{Z}}$ (resp., $\{0,1\}^{\mathbb{Z}}$) such that, for any $\delta > 0$, $g \in C_c^\infty(\mathbb{R})$,*

$$\begin{aligned} \lim_{N \to \infty} \mu_N \Big(\eta \in \mathbb{N}^{\mathbb{Z}} : \Big| \tfrac{1}{N} \sum_{x \in \mathbb{Z}} g(\tfrac{x}{N}) \eta(x) - \int_{\mathbb{R}} g(x) \bar{\rho}(x) \, \mathrm{d}x \Big| > \delta \Big) = 0, \\ \lim_{N \to \infty} \bar{\mu}_N \Big(\bar{\eta} \in \{0,1\}^{\mathbb{Z}} : \Big| \tfrac{1}{N} \sum_{x \in \mathbb{Z}} g(\tfrac{x}{N}) \bar{\eta}(x) - \int_{\mathbb{R}} g(x) \bar{\rho}(x) \, \mathrm{d}x \Big| > \delta \Big) = 0. \end{aligned} \tag{1.40}$$

Let $\mathbb{P}_{\mu_N}$ (resp., $\bar{\mathbb{P}}_{\bar{\mu}_N}$) be the law of the measure-valued process $t \mapsto X^N_{\text{in}}(t)$ (resp., $t \mapsto X^N_{\text{ex}}(t)$) in (1.38) induced by the initial distribution μ_N (resp., $\bar{\mu}_N$). Then, for any $T > 0$, $\delta > 0$ and $g \in C_c^\infty(\mathbb{R})$,

$$\begin{aligned} \lim_{N \to \infty} \mathbb{P}_{\mu_N} \left(\sup_{t \in [0,T]} \Big| \int_{\mathbb{R}} g(x) \, \mathrm{d}X^N_{\text{in}}(t)\{x\} - \int_{\mathbb{R}} g(x) \rho(x,t) \, \mathrm{d}x \Big| > \delta \right) = 0, \\ \lim_{N \to \infty} \bar{\mathbb{P}}_{\bar{\mu}_N} \left(\sup_{t \in [0,T]} \Big| \int_{\mathbb{R}} g(x) \, \mathrm{d}X^N_{\text{ex}}(t)\{x\} - \int_{\mathbb{R}} g(x) \rho(x,t) \, \mathrm{d}x \Big| > \delta \right) = 0, \end{aligned} \tag{1.41}$$

where $\rho(\cdot\,,\cdot)$ is the unique strong solution of the heat equation

$$\begin{cases} \partial_t \rho = \Delta \rho_0, \\ \rho(x,0) = \bar{\rho}(x). \end{cases} \tag{1.42}$$

The above theorem asserts that both measure-valued processes $t \mapsto X^N_{\text{in}}(t)$ and $t \mapsto X^N_{\text{ex}}(t)$ converge weakly, in probability, to a limiting measure-valued process $t \mapsto X_t$, where X_t is a deterministic measure on $\mathbb{R}$ for each $t \geq 0$. Furthermore, X_t is absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}$ with density $\rho(\cdot\,,t)$, i.e.,

$$\mathrm{d}X_t\{x\} = \rho(x,t) \, \mathrm{d}x, \quad t \geq 0, x \in \mathbb{R}, \tag{1.43}$$

and $(\rho(\cdot\,,t))_{t \geq 0}$ is the unique strong solution of the heat equation in (1.42). To keep matters simple, we skip the technical details of the proof, which essentially exploits

stochastic self-duality properties of the two microscopic systems, along with Donsker's invariance principle for a simple symmetric random walk.

Now that we have seen how one retrieves the well-known heat equation from two seemingly complex microscopic particle systems, it is natural to wonder about the macroscopic effect of introducing dormant characteristics at the microscopic level. The first half of Part II is devoted to studying such effects. In particular, we discuss the hydrodynamic scaling behaviour of the three microscopic systems introduced above, supplemented with "dormancy". Our results will be summarised in Section 1.2.3.

§1.2.2 Non-equilibrium steady state

In Section 1.2.1 we motivated the study of interacting particle systems as a mathematical way of modelling physical systems consisting of a large number of microscopic components. Interacting particle systems are Markov processes on an uncountable state space that typically deal with the evolution of infinitely many variables, such as the location of infinitely many particles, the infection status or gene type of individuals in an infinite population, etc. In reality, however, physical or biological systems consist of a large yet finite number of components. This presents an undesirable discrepancy between theoretical models and real physical systems. The standard way to overcome this discrepancy is by restricting the proposed model to a finite region of interest.

To illustrate the idea, consider the process of heat conduction on a one-dimensional metal rod (see Fig. 1.5). We can assume that the microscopic structure of the rod is discrete and can be represented by the integer lattice $\mathbb{Z}$. As observed in Section 1.2.1, the spatio-temporal macroscopic heat profile in the rod, which is given by the solution of the heat equation in (1.42), can be thought of as a by-product of the process $t \mapsto (\eta_{\text{in}}(i,t))_{i\in\mathbb{Z}}$ in (1.31) where particles perform independent random walks on the microscopic lattice structure of the metal rod. If initially the metal rod has not yet reached a global thermal equilibrium in terms of the macroscopic heat conduction, and we focus on the heat profile of a segment of the rod, called the *bulk*, with a length that is negligible compared to the total length of the rod, then we will observe that the heat profile in the bulk first attains a *local equilibrium*. This local equilibrium typically depends on the initial cumulative amount of heat contained in the complement of the bulk, called *external reservoirs*, at both ends of the rod. If the sizes of the reservoirs are large enough, then the average amount of heat contained within them endure negligible effects from the heat profile of the bulk, and therefore remain almost constant throughout the macroscopic evolution of the bulk profile. From the microscopic perspective, the spatial extent of the bulk is so small compared to the external reservoirs that the interactions between particles in the reservoirs only have an average effect on the random motions of the particles in the bulk.

Independent particle system with reservoirs. In view of the above, a mathematically more accurate understanding for the local equilibrium can be achieved by modelling the collective effects of the two external reservoirs on the bulk variables with two individual *boundary reservoirs*. The modified dynamics is such that microscopic particles can escape from the bulk to enter the boundary reservoirs independently of

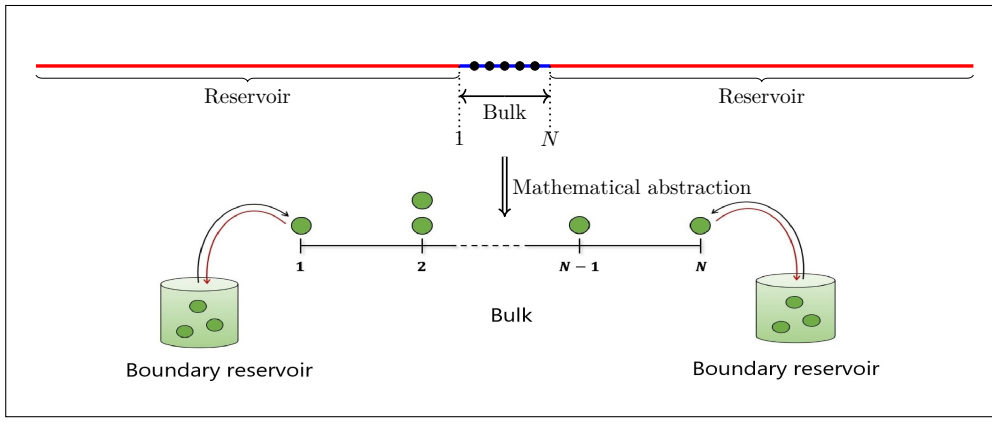


Figure 1.5: A schematic representation of bulk and reservoirs in a one-dimensional metal rod. The length of the bulk is $N \in \mathbb{N}$ in microscopic units, and is small compared to the lengths of the reservoirs.

each other, while each reservoir can put particles at the boundary sites at a *constant rate* determined by the cumulative amount of "heat" contained within them. As an outcome, we obtain a new microscopic process $\widehat{\eta}_{\text{in}}$ defined as

$$\widehat{\eta}_{\text{in}} := (\widehat{\eta}_t)_{t\geq 0}, \quad \widehat{\eta}_t := (\widehat{\eta}(i,t))_{i\in[N]_*}, \quad t \geq 0, \tag{1.44}$$

where $[N]_* := \{1, \dots, N\}$ and $\widehat{\eta}(i,t)$ represents the number of particles at site $i \in [N]_*$ in the bulk at time $t \geq 0$. The process $\widehat{\eta}_{\text{in}}$ is a continuous-time Markov chain with generator

$$\widehat{L}_{\text{in}} := \widehat{L}^{\text{bulk}} + \widehat{L}^{\text{res}}. \tag{1.45}$$

The action of $\widehat{L}^{\text{bulk}}$ and $\widehat{L}^{\text{res}}$ on a test function $f : \mathbb{N}_0^{[N]_*} \to \mathbb{R}$ is as follows:

$$(\widehat{L}^{\text{bulk}} f)(\eta) := \sum_{x\in[N]_*} \eta_x \sum_{\substack{y\in[N]_*,\\ |x-y|=1}} [f(\eta^{x,y}) - f(\eta)], \tag{1.46}$$

and

$$\begin{aligned}(\widehat{L}^{\text{res}} f)(\eta) := \eta_1 \left[f(\eta^{1,-}) - f(\eta)\right] + \eta_N \left[f(\eta^{N,-}) - f(\eta)\right] \\ + \rho_L \left[f(\eta^{1,+}) - f(\eta)\right] + \rho_R \left[f(\eta^{N,+}) - f(\eta)\right],\end{aligned} \tag{1.47}$$

where $\eta := (\eta_x)_{x\in[N]_*} \in \mathbb{N}_0^{[N]_*}$ is a configuration representing the number of particles at each site in the bulk, $\rho_L > 0$ (resp., $\rho_R > 0$) is the rate at which the left (resp., right) reservoir injects new particles at site 1 (resp., N), the configurations $\{\eta^{x,y} : x, y \in [N]_*\}$ are obtained from η by using (1.33), and

$$\eta^{x,-} := (\eta_i - \mathbb{1}_{\{i=x,\eta_i\geq 1\}})_{i\in[N]_*}, \quad \eta^{x,+} := (\eta_i + \mathbb{1}_{\{i=x\}})_{i\in[N]_*}, \quad x \in [N]_*. \tag{1.48}$$

Observe from (1.46)–(1.47) that $\widehat{L}^{\text{bulk}}$ is responsible for the independent motions of the particles in the bulk, while $\widehat{L}^{\text{res}}$ dictates the interactions between the particles and the reservoirs at the two boundary sites 1 and N.

Non-equilibrium steady state. The process $\widehat{\eta}_{\text{in}}$ is well-known in the literature (see e.g., [110, 24]) and admits a unique equilibrium distribution ν_{ρ_L,ρ_R} due to the presence of the reservoirs. It exhibits interesting behaviour in the equilibrium ν_{ρ_L,ρ_R}, whose explicit form is known as well (see [24, Proposition 4.5]). To be specific, when the two boundary reservoirs operate at identical environmental conditions, i.e., when they are in thermal equilibrium w.r.t. each other because $\rho_L = \rho_R = \rho$, the variables $\{\widehat{\eta}_x : x \in [N]_*\}$ denoting the number of particles at different sites in the bulk behave independently of each other with a Poisson distribution of mean ρ. In contrast, if the reservoirs are not at thermal equilibrium (i.e., $\rho_L \neq \rho_R$), then the variables $\{\widehat{\eta}_x : x \in [N]_*\}$ remain independent, but no longer follow an identical marginal distribution under the law ν_{ρ_L,ρ_R}. In this scenario ν_{ρ_L,ρ_R} is referred to as a *non-equilibrium steady state* of the bulk (or, equivalently, of the process $\widehat{\eta}_{\text{in}}$), because it describes the physical phenomenon that, even though the metal rod is not in a global equilibrium (as $\rho_L \neq \rho_R$), it is nonetheless in a microscopic local equilibrium ν_{ρ_L,ρ_R} when viewed only in the bulk.

Fick's law of mass transport. In the presence of the reservoirs, the macroscopic properties of the bulk variables in the metal rod undergo only minor changes. By means of hydrodynamic scaling of the process $\widehat{\eta}_{\text{in}}$ in (1.44), we can easily extract properties of the corresponding *macroscopic local equilibrium*. In fact, the macroscopic heat profile in the bulk still follows the same heat equation (interpreted in the sense of hydrodynamics)

$$\begin{cases} \partial_t \rho = \Delta \rho, \\ \rho(x,0) = \bar{\rho}(x), \end{cases} \qquad x \in [0,1], \tag{1.49}$$

but with additional boundary conditions

$$\begin{cases} \rho(0,t) = \bar{\rho}(0) = \rho_L, \\ \rho(1,t) = \bar{\rho}(1) = \rho_R, \end{cases} \qquad t \geq 0, \tag{1.50}$$

that arise precisely due to the coupling with the two boundary reservoirs.

Heat conduction in a metal rod is but one example where the macroscopic equation in (1.49)–(1.50) is used to model the underlying physical process. Many physical experiments suggest that the transport of a solute between two compartments (or 'reservoirs') separated by a thin layer of membrane (or 'bulk') is governed by the same equation. More generally, the continuity equation for mass transport, which basically is a consequence of the conservation of total mass, states that

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot J, \tag{1.51}$$

where $\rho(x,t)$ is the density of the solute at a macroscopic position $x \in [0,1]$ at time $t \geq 0$, and $J : [0,1] \times \mathbb{R}^+ \to \mathbb{R}$ is the instantaneous diffusion flux measuring the amount of solute passing through a unit area per unit time. Experimental results based on analysis of single-component diffusions in homogeneous media align with the prediction of the so-called *Fick's law*, which postulates that the diffusion flux J is

in a direction opposite to the gradient of the concentration, i.e., for some diffusivity constant $\sigma > 0$,

$$J = -\sigma \nabla \rho. \tag{1.52}$$

Combining (1.51)–(1.52), one recovers the familiar diffusion (or heat) equation given in (1.49).

Uphill diffusion. In situations where Fick's law does not hold, *uphill* diffusion becomes possible. Uphill diffusion is characterised by the flow of a solute from an area with a lower concentration to an area with a higher concentration. In homogeneous media, diffusion of a single component obeys Fick's law and therefore the flow is always downhill. However, in a multi-component mixture, interaction between different components can change their diffusive characteristics in such a way that the overall effect results in uphill diffusion of the total density [102]. In particular, if a single component exhibits 'dormant characteristics' and therefore represents a solute with a mixture of both states (active and dormant), then it is reasonable to expect interesting behaviour at equilibrium, such as the violation of Fick's law, uphill diffusion, etc.

An aim in the second half of Part II is to incorporate boundary reservoirs into the three multi-layer systems with dormancy, which will be briefly introduced in the next section, and study properties of their corresponding microscopic non-equilibrium steady states. Furthermore, in the context of mass transport, by studying the associated macroscopic properties of the systems, we investigate how the interplay between active and dormant states of a single component can give rise to uphill diffusion.

§1.2.3 Summary of Part II

We start by describing how the three earlier defined interacting particle systems are adapted in order to include dormant characteristics.

Three switching interacting particle systems. We modify the three particle systems introduced in Section 1.2.1 by allowing the underlying particles to switch into a "mild" or a "pure" dormant state independently of each other. The mild dormant state of a particle causes a slowdown in its random motion. In particular, the particles move at a slower (or zero) rate in their mild (or pure) dormant state. Formally, for $\sigma \in \{-1, 0, 1\}$ we introduce the modified interacting particle systems on $\mathbb{Z}$ where the particles randomly switch their jump rate between two possible values, 1 and $\epsilon \in [0, 1]$, depending on whether they are in an active or a dormant state. For $\sigma = -1$ the particles evolve as in the simple symmetric exclusion process, for $\sigma = 0$ the particles perform independent random walks, while for $\sigma = 1$ the particles evolve as in the simple symmetric inclusion process. Furthermore, the type of a particle can change at rate $\gamma > 0$ and the total rate of these changes is tuned to the underlying interaction rule. Observe that the dormant particles are still allowed to jump, but at a slower rate ϵ than the active particles. Let

$$\eta_0(x, t) := \text{number of active particles at site } x \text{ at time } t \geq 0,$$
$$\eta_1(x, t) := \text{number of dormant particles at site } x \text{ at time } t \geq 0.$$

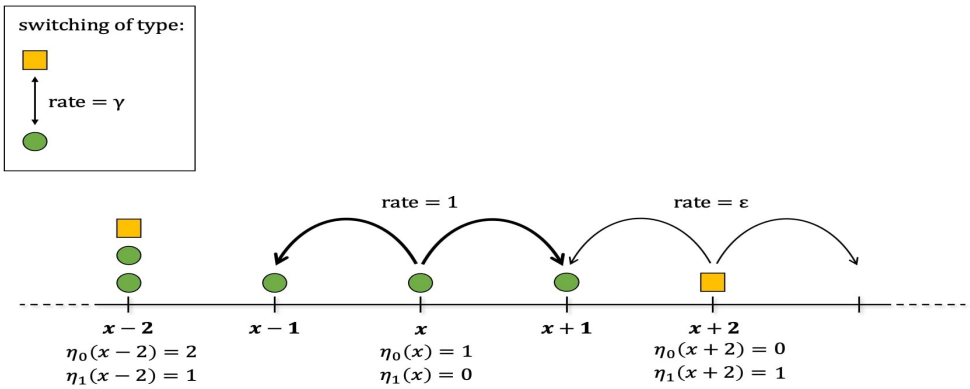


Figure 1.6: Representation of the switching random walks via slow (dormant) *and* fast (active) *particles.*

The process

$$(\eta_0(t), \eta_1(t))_{t\geq 0}, \quad (\eta_0(t), \eta_1(t)) := (\eta_0(x,t), \eta_1(x,t))_{x\in\mathbb{Z}}, \tag{1.53}$$

lives on the state space $\mathcal{X} := \bar{\mathcal{X}} \times \bar{\mathcal{X}}$, where

$$\bar{\mathcal{X}} := \begin{cases} \{0,1\}^{\mathbb{Z}}, & \text{if } \sigma = -1, \\ \mathbb{N}_0^{\mathbb{Z}}, & \text{if } \sigma \in \{0,1\}, \end{cases} \tag{1.54}$$

and forms a Markov process that we refer to as *switching exclusion process* for $\sigma = -1$, *switching random walks* for $\sigma = 0$ (see Fig. 1.6), and *switching inclusion process* for $\sigma = 1$. Before giving the explicit form of the generator, it is convenient to define, for $\sigma \in \{-1, 0, 1\}$,

$$L_\sigma := \begin{cases} L_{\text{ex}}, & \text{if } \sigma = -1, \\ L_{\text{in}}, & \text{if } \sigma = 0, \\ L_{\text{inc}}, & \text{if } \sigma = 1, \end{cases} \tag{1.55}$$

where the three generators $L_{\text{in}}, L_{\text{ex}}$ and L_{inc} are as in (1.32),(1.35) and (1.37), respectively. The generator L_σ encodes the three processes, namely, the independent particle system, the exclusion process and the inclusion process, in a single generator and can be used to describe the generator for the switching process in a simplified form. Indeed, the generator $L_{\epsilon,\gamma}$ of the switching process acts on a suitable test function $f : \mathcal{X} \to \mathbb{R}$ as

$$(L_{\epsilon,\gamma} f)(\eta_0, \eta_1) := (L_\sigma f(\cdot, \eta_1))(\eta_0) + \epsilon (L_\sigma f(\eta_0, \cdot))(\eta_1) + \gamma (L_{0\updownarrow 1} f)(\eta_0, \eta_1), \tag{1.56}$$

where $(\eta_0, \eta_1) := (\eta_0(x), \eta_1(x))_{x\in\mathbb{Z}} \in \mathcal{X}$ and $L_{0\updownarrow 1}$ acts on f as

$$\begin{aligned} (L_{0\updownarrow 1} f)(\eta_0, \eta_1) := \sum_{x\in\mathbb{Z}} \Big\{ & \eta_0(x)(1 + \sigma\eta_1(x)) \left[f(\eta_0^{x,-}, \eta_1^{x,+}) - f(\eta) \right] \\ & + \eta_1(x)(1 + \sigma\eta_0(x)) \left[f(\eta_0^{x,+}, \eta_1^{x,-}) - f(\eta) \right] \Big\} \end{aligned} \tag{1.57}$$

and is part of the generator $L_{\epsilon,\gamma}$ that describes the switching between the two states (active or dormant) of a particle. Here, for a configuration $\eta \in \bar{\mathcal{X}}$ and a site $x \in \mathbb{Z}$ the configurations $\eta^{x,+}, \eta^{x,-}$ are defined as in (1.48).

Observe in (1.56) that the first (resp., the second) term on the right-hand side describes the motions of the active (resp., the dormant) particles according to the interaction rule of the particle system. Also observe that in (1.57) the total rate at which a particle changes its state from active to dormant or vice versa depends on the particular interaction between the particles. Indeed, the switching between the particle types happens independently when $\sigma = 0$. In case $\sigma = -1$, an active particle at a site prohibits another dormant particle at the same site to become active and vice versa. In case $\sigma = 1$, an active particles encourages another dormant particle to become active at the same site and vice versa. We emphasise that the type of interaction between particles of opposite states is intentionally chosen to be the same as the interaction between particles of the same state. This choice is in fact crucial for preserving the self-duality properties of the particle systems without dormancy.

Hydrodynamics: reaction-diffusion equation. In order to study the hydrodynamic scaling limit of the switching process $t \mapsto (\eta_0(t), \eta_1(t))$ introduced in (1.53) we consider the following scaling of space and time. We introduce a coarse-graining parameter $N \in \mathbb{N}$ and scale space by $1/N$, time by N^2, the switching rate γ_N by $1/N^2$, and let $N \to \infty$ to obtain a system of macroscopic equations associated with the switching interacting particle system. Note that while coarse graining, i.e., zooming out of the microscopic world to the macroscopic world, we keep the rates at which particles move constant. This is because we scale time by N^2, which automatically takes care of scaling the rate of the spatial movement of the particles. Similarly as in the three original particle systems, we consider the following (Radon) measure-valued quantities associated with the switching process in order to study hydrodynamic behaviour:

$$\mathsf{X}_0^N(t) := \frac{1}{N}\sum_{x\in\mathbb{Z}} \eta_0(x, tN^2)\,\delta_{x/N}, \qquad \mathsf{X}_1^N(t) := \frac{1}{N}\sum_{x\in\mathbb{Z}} \eta_1(x, tN^2)\,\delta_{x/N}. \tag{1.58}$$

Here, δ_y stands for the Dirac measure at $y \in \mathbb{R}$. The variables $X_0^N(t)$ and $X_1^N(t)$ in (1.58) are the empirical densities of, respectively, the active and the dormant particles at time $t \geq 0$. Note that, because the switching process $t \mapsto (\eta_0(t), \eta_1(t))$ has a càdlàg path, the corresponding path associated with the process $t \mapsto (\mathsf{X}_0^N(t), \mathsf{X}_1^N(t))$ is càdlàg as well. By exploiting the self-duality property along with some mild regularity conditions on the initial distributions of the rescaled switching process, we can show that the weak limit as $N \to \infty$ of $t \mapsto (\mathsf{X}_0^N(t), \mathsf{X}_1^N(t))$ in the Skorokhod topology is the deterministic continuous measure-valued path $t \mapsto (\mathsf{X}_0(t), \mathsf{X}_1(t))$ with

$$\mathrm{d}\mathsf{X}_i(t)\{x\} = \rho_i(x,t)\,\mathrm{d}x, \quad t \geq 0,\, x \in \mathbb{R},\, i \in \{0,1\}, \tag{1.59}$$

where $\rho_0(\cdot\,,\cdot)$ and $\rho_1(\cdot\,,\cdot)$ are the unique bounded strong solutions of the reaction-diffusion equation

$$\begin{cases} \partial_t \rho_0 = \Delta\rho_0 + \Upsilon(\rho_1 - \rho_0), \\ \partial_t \rho_1 = \epsilon\Delta\rho_1 + \Upsilon(\rho_0 - \rho_1), \end{cases} \tag{1.60}$$

with initial conditions

$$\begin{cases} \rho_0(\cdot\,,0) = \bar{\rho}_0(\cdot), \\ \rho_1(\cdot\,,0) = \bar{\rho}_1(\cdot). \end{cases} \tag{1.61}$$

In (1.60) the parameter Υ is the limiting value of the rescaled switching rates $(\gamma_N)_{N\in\mathbb{N}}$ associated with the switching process, i.e., $\lim_{N\to\infty} N^2\gamma_N = \Upsilon$, and intuitively corresponds to the rate of switching (between active and dormant particles) events on the macroscopic scale. In (1.61) the initial macroscopic profiles $\bar{\rho}_0$ and $\bar{\rho}_1$ are assumed to be bounded continuous functions. These regularity conditions on the initial profiles are needed in order to ensure the existence and uniqueness of strong solutions of (1.60) (see e.g., [68, Chapter 5, Section 4, Theorem 4.1]).

The partial differential equations of type (1.60) fall in the class of reaction-diffusion equations, which are used to model time-dependent evolution of concentrations of certain substances in a solution due to diffusion and chemical reaction. Our finding that the hydrodynamic equation of a microscopic process with dormancy is a reaction-diffusion equation suggests that dormancy at a microscopic level can induce non-trivial effects on a macroscopic level and has the potential to change the qualitative behaviour of physical or chemical systems. Indeed, if ρ_0, ρ_1 are smooth enough and satisfy (1.60), then by taking extra derivatives we see that the total density $\rho := \rho_0 + \rho_1$ satisfies the *thermal telegrapher equation*

$$\partial_t \left(\partial_t \rho + 2\Upsilon\rho\right) = -\epsilon\Delta(\Delta\rho) + (1+\epsilon)\Delta\left(\partial_t\rho + \Upsilon\rho\right), \tag{1.62}$$

which is second order in ∂_t and fourth order in ∂_x (see [2, 86] for a derivation). Note from (1.62) that the total density does not satisfy the usual diffusion equation of type (1.49). This fact is investigated in detail in the second half of Part II where we analyse the non-Fick property of ρ.

Non-equilibrium behaviour: uphill diffusion. In the second half of Part II, we look at the non-equilibrium steady state behaviour of the switching process by introducing boundary reservoirs similar to the ones included in the process $\hat{\eta}_{\text{in}}$ in (1.44). In particular, we restrict the switching process to a finite region $[N]_* := \{1, \dots, N\}$ of $\mathbb{Z}$ where $N \geq 2$, and add two boundary reservoirs at each site 1 and N (see Fig. 1.7). The two reservoirs at a boundary site control the injection and absorption of, respectively, active and dormant particles. The rates at which particles are injected or absorbed by the reservoirs are chosen according to the type of interaction rule in the switching process. This is because when the rates associated with the reservoir dynamics are compatible with the dynamics of the particles in the bulk, the switching process admits a dual process. We already mentioned earlier that the switching process without the reservoirs is self-dual, a property it inherits from the three underlying particle systems, namely, the independent particle system, the exclusion process and the inclusion process. In the presence of the reservoirs, the bulk dynamics in the switching process preserves the self-duality property as well, but the reservoirs in the dual process become absorbing. Therefore the corresponding dual also consists of a system of active and dormant particles, where particles perform the same dynamics as before in the bulk, but are eventually absorbed at the boundary sites by the two

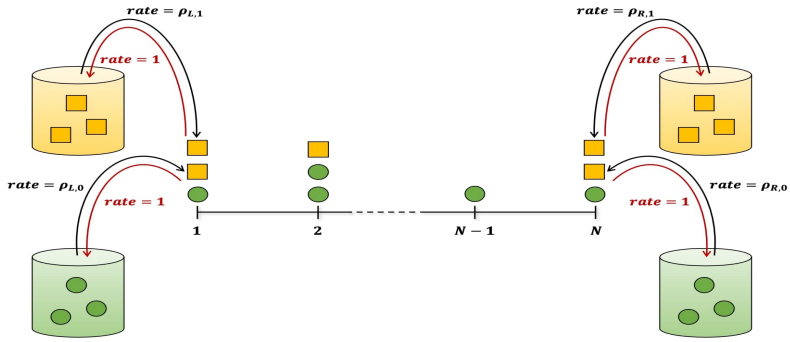


Figure 1.7: Representation of the switching process with boundary reservoirs when $\sigma = 0$. Green particles are active and yellow particles are dormant. Here, $\rho_{L,0}, \rho_{R,0}, \rho_{L,1}, \rho_{R,1}$ are positive parameters controlling the rates at which reservoirs put or remove particles at the boundary sites.

reservoirs at certain rates. To avoid technicalities we refrain from giving the precise mathematical definition of the dual here (see Section 5.3.2). The absorbing nature of the dual immensely simplify the analysis of the switching process with reservoirs and allows for a partial characterisation of the unique non-equilibrium steady state μ_{stat} of the process. In particular, we obtain explicit expressions for the stationary microscopic profile $(\theta_0^{(N)}(x), \theta_1^{(N)}(x))_{x\in[N]_*}$ defined by

$$\theta_i^{(N)}(x) := \mathbb{E}_{\mu_{\text{stat}}}[\eta_i(x,t)], \quad x \in \{1,\dots,N\},\, t \geq 0,\, i \in \{0,1\}, \tag{1.63}$$

where $t \mapsto (\eta_0(x,t), \eta_1(x,t))_{x\in[N]_*}$ is the switching process with reservoirs.

By computing the average flux of the particles in the stationary switching process with the help of the dual process, we are able to characterise the stationary microscopic current through each horizontal edge of the graph $\{1,\dots,N\}$. It turns out that in stationarity the total average current through each horizontal edge is the same and is of the order $\frac{1}{N}$. Therefore, an unambiguous notion of *uphill current* is obtained by imposing that the sign of the stationary current through each edge is the same as the sign of the total density gradient of the particles at the two boundary sites.

We also study the macroscopic behaviour of the stationary switching process with reservoirs under the same scaling of the microscopic parameters as was done in the context of hydrodynamic scaling. We derive the stationary macroscopic profiles of the system by taking the pointwise limit of the microscopic stationary profiles. To be more precise, we obtain the stationary macroscopic profile $(\rho_0^{\text{stat},\epsilon}(y), \rho_1^{\text{stat},\epsilon}(y))_{y\in[0,1]}$ by setting

$$\rho_i^{\text{stat},\epsilon}(y) := \lim_{N\to\infty} \theta_i^{(N)}(\lceil yN \rceil), \quad y \in [0,1],\, i \in \{0,1\}. \tag{1.64}$$

When $\epsilon > 0$, i.e., microscopic particles only admit a mild dormant state, it turns out that the stationary macroscopic profiles $\rho_0^{\text{stat},\epsilon}(\cdot), \rho_1^{\text{stat},\epsilon}(\cdot)$ constitute the unique

smooth strong solution of the boundary value problem

$$\begin{cases} 0 = \Delta u_0 + \Upsilon(u_1 - u_0), \\ 0 = \epsilon \Delta u_1 + \Upsilon(u_0 - u_1), \end{cases} \tag{1.65}$$

with boundary conditions

$$\begin{cases} u_0(0) = \rho_{L,0},\ u_0(1) = \rho_{R,0}, \\ u_1(0) = \rho_{L,1},\ u_1(1) = \rho_{R,1}, \end{cases} \tag{1.66}$$

However, when $\epsilon = 0$, i.e., microscopic particles only admit a pure dormant state, in the non-equilibrium situation (i.e., the two reservoirs at a boundary site are not in thermal equilibrium) the stationary macroscopic profile $\rho_1^{\mathrm{stat},0}$ for the dormant particles has a discontinuity near the boundary sites. By taking $\epsilon \downarrow 0$ and analysing the limiting behaviour of the stationary macroscopic profile $\rho_1^{\mathrm{stat},\epsilon}$, we find that the discontinuity of $\rho_1^{\mathrm{stat},0}$ appears as a sudden bump in the smooth stationary profile $\rho_1^{\mathrm{stat},\epsilon}$ at a distance of order $\sqrt{\epsilon}\log(1/\epsilon)$ from the boundary sites (see Proposition 5.3.20 for a precise statement).

The precise microscopic parameter regime for an uphill current is difficult to describe. However, in the macroscopic setting, the uphill regime becomes simpler and can be described by a continuous manifold determined by the parameters $a_0 := \rho_{R,0} - \rho_{L,0}$, $a_1 := \rho_{R,1} - \rho_{L,1}$ and ϵ. In particular, we show that a macroscopic uphill current takes place in the non-equilibrium situation if and only if

$$a_0^2 + (1+\epsilon)\, a_0 a_1 + \epsilon a_1^2 < 0. \tag{1.67}$$

§1.3 Further research

Finite-systems scheme. In Part I of this thesis, we study an interacting particle system $t \mapsto Z(t)$ that approximates the behaviour of genetic evolution in structured populations with seed-banks. In our model, the populations are assumed to be located on the d-dimensional integer lattice $\mathbb{Z}^d$. Although in general such infinite systems reasonably well approximate real-world populations distributed over a large geographic space, real-world geographic regions are never infinite. Therefore, from the applied point of view, one is usually interested in the behaviour of the process $t \mapsto Z_\Lambda(t)$ evolving on a finite geographic space $\Lambda \subset \mathbb{Z}^d$. The corresponding process Z_Λ restricted to Λ clusters almost surely in a finite random time τ_Λ regardless of the starting configuration. Understanding the asymptotic behaviour of the time τ_Λ and the process Z_Λ as the size $|\Lambda|$ tends to infinity is crucial for any practical use of the model. In the so-called finite-systems scheme studied in e.g., [38, 40, 74], the aim is to provide mathematically precise statements on the comparison between Z_Λ and Z as $\Lambda \uparrow \mathbb{Z}^d$.

In the coexistence regime of the infinite-volume process Z, there is a one-parameter family of non-trivial equilibria $\{\nu_\theta : \theta \in [0,1]\}$ parametrised by the density θ of a fixed gene type. If the infinite-volume process Z is in the coexistence regime, then we expect that, as $\Lambda \uparrow \mathbb{Z}^d$, the law of the finite-volume process Z_Λ on a deterministic time scale t_Λ close to τ_Λ locally approximates the law ν_θ, where the density parameter

θ is a random macroscopic quantity $Y_t \in [0,1]$ for any $t > 0$ such that $\frac{t_\Lambda}{|\Lambda|} \overset{|\Lambda|\uparrow\infty}{\longrightarrow} t$. Depending on whether the average relative strengths of the seed-banks are finite or infinite, the behaviour of $t \mapsto Y_t$ is expected to fall in different *universality classes*. For instance, in the case when seed-banks have finite relative strength on average, we expect the evolution of $t \mapsto Y_t$ to be governed by the Wright-Fisher diffusion, with a diffusion constant that is slowed down by an extra factor capturing the finite average relative seed-bank strength. However, if the average relative strength of the seed-banks is infinite, then different universality classes may appear depending on how fast the seed-bank strengths grow as $\Lambda \uparrow \mathbb{Z}^d$ compared to the time scale t_Λ. It may happen that the evolution of $t \mapsto Y_t$ is no longer a diffusion, but rather a jump process.

Interplay of dormancy, selection and mutation. In Part I, we considered a stochastic model for the genetic evolution of spatially structured populations under the influence of migration, resampling and dormancy. As mentioned earlier in Section 1.1.1, two other important evolutionary forces are selection and mutation. It would be interesting to incorporate these into our model and see how dormancy competes with them. In [57] the authors introduced a Moran model with selection and mutation where the process admits a dual with a similar hypergeometric duality function as in our model. Although their model is concerned with a single finite population, it can be seamlessly extended to the spatial setting with seed-banks similarly like in our context without loss of the duality property. The corresponding dual process is expected to be a branching coalescing interacting particle system, where particles can be active or dormant. Active particles can migrate (due to migration), coalesce with another active particle to form a single active particle (due to resampling), branch into two active particles (due to selection), die (due to mutation), and fall asleep (due to dormancy). In the presence of mutation, we obtain a Feynman-Kac type duality relation between the original process and the dual process.

A typical trend in population genetic models that incorporate mutation but no dormancy is *ergodicity*, i.e., the process converges to a unique equilibrium starting from any initial state (see e.g., [140, Theorem 1.1]). However, in the modified spatial model with mutation and dormancy, seed-banks with an infinite average relative strength may prevent ergodicity altogether and cause a *phase transition* depending on the mutation rate and the relative seed-bank strength in different colonies. The reason behind such speculation is that ergodicity of the original process arises from the annihilating nature of the particles in the dual. If the relative seed-bank strengths are infinite on average, then the dual particles spend most of their time in the dormant state and therefore annihilation events, which happen only when the particles are active, become rare. It will be interesting to turn these heuristics into precise mathematical statements and see how seed-banks give rise to qualitatively different equilibrium behaviour.

Dormancy in fluctuating random environment. In Chapter 4 we study the spatial Moran process with seed-banks in a *static* random environment. The random environment is obtained by sampling the constituent population sizes from a translation-invariant ergodic random field and remains static throughout the evolution of the process. However, in real-world scenario the population sizes are more likely

to change over time. This calls for a model with seed-banks evolving in a *dynamic* random environment. In this setting, the corresponding process becomes a time-inhomogeneous Markov process that is relatively difficult to analyse. Furthermore, we typically lose the stochastic duality property that is crucial for the analysis of the process. These difficulties make the model in the dynamic random environment more interesting from a mathematical point of view because it requires the development of novel techniques.

Systems with multi-layer seed-banks. In the stochastic systems considered in Part I and Part II the constituent seed-banks consist of only one layer.

In the spatial Moran process introduced in Part I the seed-bank in each colony has a finite size that depends on the location of the colony. Because of the location-dependent population sizes, the state space of the process is not translation invariant. On the one hand, the lack of translation invariance makes the analysis of ergodic properties of the process more complicated. On the other hand, if we recover the translation invariance by considering equal population sizes in each colony, then we no longer see the effect of seed-banks on the equilibrium behaviour of the process. To be more precise, the process in the homogeneous state space behaves exactly like the process without seed-banks, where dichotomy of coexistence vs clustering is solely determined by the migration kernel. A solution to this problem can be obtained by extending our model to a multi-layer setting. In particular, following the second model introduced in [76], we can preserve both the translation invariance and the effect of seed-banks by incorporating seed-banks with infinitely many layers at each colony. More precisely, we keep the sizes of the active populations constant and put infinitely many seed-banks of equal size at each colony. Active individuals adopt a colour before entering into a seed-bank, which determines the average of their wake-up time from the dormant state. The advantage of this extension is that we do not destroy the duality property and keep the translation-invariance of the state-space of the underlying process. We expect a similar dichotomy between clustering vs coexistence, but the criterion determining which of the two occurs will heavily rely on the strength of the deep seed-banks and the migration mechanism.

A similar extension for the switching process in Part II to the multi-layer setting is available, where we preserve the self-duality property of the original process. It will be interesting to see if uphill diffusion indeed can appear in such a setting and, if so, in what manner it changes the qualitative behaviour of the system.

§1.4 Outline of the thesis

Part I of this thesis is based on [46, 47, 125] and consists of Chapters 2–4. In Chapter 2 we introduce the interacting particle system describing genetic evolution of spatially structured populations with seed-banks and state our main results on the well-posedness of the model, sampling duality relation with a dual interacting particle system, and the dichotomy between mono-type equilibria (clustering regime) and multi-type equilibria (coexistence regime). In Chapter 3 we refine the criterion

for the clustering regime given in Chapter 2 and identify the precise parameter regime for clustering, which is determined by the relative seed-bank strengths and the migration kernel. In Chapter 4 we extend the model in Chapter 2 to a static random environment setting. Under mild assumptions on the law of the environment and the migration kernel, we state and prove homogenisation results on the equilibrium behaviour of the process in the clustering regime.

Part II of this thesis is based on [62] and consists of Chapter 5. In Chapter 5 we introduce a switching interacting particle system, where particles can be in an active state or a (mild/pure) dormant state. We state and prove results on the hydrodynamic scaling limit, the stochastic duality property of the process etc. Furthermore, we study the non-equilibrium behaviour of the process in the presence of boundary reservoirs and state results on uphill diffusion of the particles, a phenomenon that manifests itself as an outcome of the reaction-diffusion type interactions between active and (mild or pure) dormant particles.