Spatial populations with seed-bank

Oomen, M.

Spatial populations with seed-bank

Margriet Oomen

The research in this thesis was supported by the NWO Gravitation Grant no 024.002.003-NETWORKS of Frank den Hollander

Spatial populations with seed-bank

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 18 november 2021
klokke 12.30 uur

door

Margriet Oomen
geboren te Roosendaal
in 1990

Samenstelling van de promotiecommissie:

1^e Promotor:
Prof. dr. F. den Hollander (Universiteit Leiden)

2^e Promotor:
Prof. dr. A. Greven (Friedrich-Alexander Universität Erlangen-Nürnberg)

Voorzitter:
Prof. dr. E.R. Eliel (Universiteit Leiden)

Secretaris:
Prof. dr. A. Doelman (Universiteit Leiden)

Overige Leden:
Prof. dr. M. Wilke Berenguer (Humboldt-Universität zu Berlin)
Prof. dr. A. Wakolbinger (Goethe Universität Frankfurt am Main)
Prof. dr. F.H.J. Redig (Technische Universiteit Delft)

Contents

CHAPTER 1

Introduction

§1.1 Spatial populations with seedbank

In populations with a *seed-bank*, individuals can temporarily become dormant and refrain from reproduction, until they can become active again. Seed-banks are observed in many taxa, including plants, bacteria and other micro-organisms. Typically, they arise as a response to unfavourable environmental conditions. The dormant state of an individual is characterised by low metabolic activity and interruption of phenotypic development (see e.g. [55]). After a varying and possibly large number of generations, a dormant individual can be resuscitated under more favourable conditions and reprise reproduction after having become active again. This strategy is known to have important implications for population persistence, maintenance of genetic variability and stability of ecosystems, (see e.g. [54]). It acts as a *buffer* against evolutionary forces such as genetic drift, selection and environmental variability.

Various attempts were made to include a seed-bank in already existing mathematical models that describe the genetic evolution of populations (see [50], [11], [10] and [70].) However, after inclusion of the seed-bank these models become complex, because they have long memory. In [12] the so-called "two-type Fisher-Wright model with seed-bank" was introduced. This was the first model that describes the evolution of a population with seed-bank as a Markov process. In this model individuals move in and out of the seed-bank at prescribed rates. Outside the seed-bank individuals are subject to *resampling*, while inside the seed-bank their resampling is *suspended*. Both the long-time behaviour and the genealogy of the population were analysed in detail. In particular, it was shown that the seed-bank increases the genetic variability of the population.

The goal of this thesis is to extend the seed-bank model introduced in [12] to the *spatial* setting where individuals can migrate between different colonies. We analyse the long-time behaviour of the evolution of a spatial population with seed-bank in different settings. We show how the seed-bank increases the genetic variability, compared to spatial population models without seed-bank. In particular, we show how certain types of seed-banks can even prevent the loss of genetic variability altogether.

§1.2 Modeling population genetics

We give a short introduction to modeling genetic evolution of populations. We also introduce two important tools to analyse the genetic evolution in populations, namely, the *Kingman coalescent* and *duality*.

§1.2.1 The Fisher-Wright model

One of the driving forces in the genetic evolution of populations is genetic drift. Genetic drift is the evolutionary mechanism that selects genes randomly. To illustrate the concept of genetic drift, consider a population of turtles. Each year the turtles lay plenty of eggs on the beach, but only a few of these eggs grow into a mature turtle. Which of the eggs will do so is random. In this way, randomness plays an important part in the genetic evolution of populations, and this randomness is called *genetic drift*.

In mathematics genetic drift is modeled through the *Fisher-Wright* model. In the Fisher-Wright model we consider a population of N individuals. Each individual can carry one of two genetic types, denoted by $\heartsuit$ and $\diamondsuit$. In each generation all the individuals will reproduce themselves according to the following rule:

- Each individual chooses uniformly at random an individual from the population and adopts its type. The chosen individual may be the individual itself.

This way of modeling reproduction is called *resampling*. Note that the number of individuals remains fixed during the evolution. Since we mostly consider very large populations without external evolutionary forces, we may assume the population size to be approximately constant. Therefore the assumption that the population size remains fixed is reasonable. An example of a population of 5 individuals is shown in Fig. 1.1. The resampling mechanism can be interpreted as follows: Each individual in the population gets a random number of offspring between 1 and N, and the total number of offspring in the next generation is N. This gives a more natural interpretation of resampling, but the way resampling is phrased above makes the mathematical analysis easier.

Evolution equation. To describe the genetic evolution in the population, we analyse the fraction of individuals of type $\heartsuit$. Label the N individuals by $[N] = \{1, \dots, N\}$. Define $\xi(k) = (\xi_j(k))_{j\in[N]} \in \{0,1\}^{[N]}$ as the random vector where $\xi_j(k) = 1$ if the j'th individual is of type $\heartsuit$ at time k and $\xi_j(k) = 0$ if the j'th individual is of type $\diamondsuit$ at time k. Then

$$X^N(k) = \frac{1}{N} \sum_{j\in[N]} 1_{\{\xi_j(k)=1\}} \tag{1.1}$$

is the fraction of individuals of type $\heartsuit$ in generation k. Since there are two types of individuals in the population, the fraction of individuals of type $\diamondsuit$ in generation k is given by $1 - X^N(k)$. The distribution of $X^N(k+1)$ given $X^N(k)$ is $\mathrm{BIN}(N, X^N(k))$. A key question is whether eventually there is only one type of individual left in the

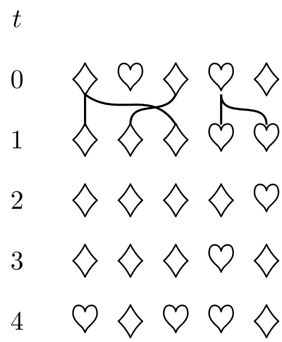

Figure 1.1: Example of the evolution for a population with $N = 5$ individuals in 5 generations. The solid lines within the active population represent resampling in the first generation.

population, or there are always two types of individuals in the population. If only one of the two types is left, then we say that *genetic variability* is lost. For a finite population, genetic variability is lost eventually. The expected time until this happens is of order N, the size of the population [30].

In this thesis we look at genetic evolution in populations where the number of individuals tends to infinity. This large population model is referred to as the continuum model. To obtain this continuum model, we let the number of individuals in the population tend to infinity and speed up time proportionally to the number of individuals in the population, i.e.,

$$\lim_{N\to\infty} \mathcal{L}[\big(X^N(\lfloor Nt \rfloor)\big)_{t\geq 0}]. \tag{1.2}$$

Thus, we observe larger and larger populations on time scales where these populations start to loose their genetic variability. The limit in (1.2) is the law of the continuous-time process $(x(t))_{t\geq 0}$ that evolves according to the stochastic differential equation (SDE)

$$\mathrm{d}x(t) = \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t), \tag{1.3}$$

where $(w(t))_{t\geq 0}$ is a standard Brownian motion. Its initial law is given by $\mathcal{L}[x(0)] = \lim_{N\to\infty} \mathcal{L}[X^N(0)]$. The process $(x(t))_{t\geq 0}$ evolving according to (1.3) is called the *Fisher-Wright diffusion* and has state space $[0,1]$. The stochastic differential equation (SDE) in (1.3) has a unique solution that is a Markov process (see [72]). The fixed points of (1.3) are 0 (only individuals of type $\diamondsuit$ are left) and 1 (only individuals of type $\heartsuit$ are left). The Fisher-Wright diffusion reaches its fixed points in finite time [30].

Geneology. After sampling $n \in \mathbb{N}$ individuals from a large population at some large time, we can ask ourselves what the lineages of these n individuals are. If two of the n sampled individuals have a common ancestor a time s backwards in time, then their lineages coalesce, (see Figure 1.2). It turns out that if we sample individuals from the continuum model, then any two lineages coalesce at rate 1, independently

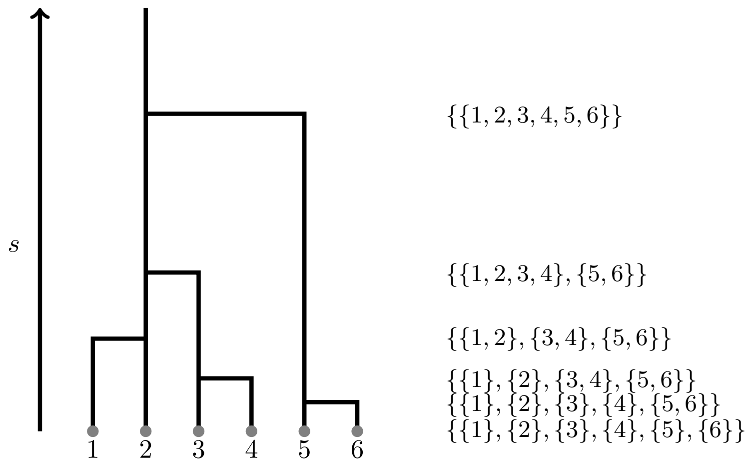

Figure 1.2: Example of a genealogy of 6 *individuals sampled from a Fisher-Wright diffusion. The corresponding Kingman coalescent is written on the right. Each time two lineages merge, the corresponding partition elements merge. Time is indicated by* s *and is running backwards.*

of the other lineages. The ancestral lineages together are called the *genealogy* of the n individuals.

The process that formally describes the genealogy of the n sampled individuals in the continuum model is called the *Kingman coalescent*. The Kingman coalescent is a partition-valued process that at time $s = 0$ assigns to each of the n individuals a partition element, i.e., at time $s = 0$ the Kingman coalescent starts from state $\{\{1\}, \{2\}, \dots, \{n\}\}$. If two lineages coalesce, then the corresponding two partition elements of the Kingman coalescent merge (see Fig. 1.2). Thus, any two partition elements merge at rate 1, independently of the other partition elements. The Kingman coalescent describes how the genetic evolution of a population took place in the past. The Kingman coalescent runs backwards in time. For this reason it is sometimes called *the backward process*. In contrast, the Fisher-Wright diffusion is called *the forward process*.

Since in the Fisher-Wright model individuals inherit their type from their parents, any two individuals whose lineages have a common ancestor are of the same type. If the number of sampled individuals tends to infinity, then the related ancestral lineages still have a common ancestor a finite time s backwards, (see [30]). Consequently, all the individuals in the population are of the same type and genetic variability is lost. We say that *the Kingman coalescent comes down from infinity in finite time*. This is the backward counterpart of the fact that the Fisher-Wright diffusion hits its fixed points in finite time.

Duality. Related to the coalescent process is the *block-counting process*. Suppose that at some large time $t > 0$ we sample n individuals from a population evolving according to the Fisher-Wright diffusion in (1.3). The block counting process $(N(s))_{s\geq 0}$ counts the number of ancestral lineages when we traverse s backwards in time,

$$N(s) = \#\text{ lineages left at time } s. \tag{1.4}$$

Since any two lineages merge at rate 1, the process $(N(s))_{s\geq 0}$ has transition rates

$$n \to n-1 \text{ at rate } \binom{n}{2}. \tag{1.5}$$

Therefore the block-counting process is a death-process. Like the Kingman coalescent, the block-counting process is a backward process.

Let $(x(t))_{t\geq 0}$ be the Fisher-Wright diffusion starting from state $x \in [0,1]$. Let $(N(t))_{t\geq 0}$ be the block-counting process starting from $n \in \mathbb{N}$. Then, for all $n \in \mathbb{N}$, $x \in [0,1]$ and $t \geq 0$, the following relation holds:

$$\mathbb{E}^x[(x(t))^n] = \mathbb{E}^n[x^{N(t)}]. \tag{1.6}$$

Here the expectation on the left-hand side is taken over the Fisher-Wright diffusion $(x(t))_{t\geq 0}$ and the expectation on the right-hand side is taken over the block-counting process $(N(t))_{t\geq 0}$. The relation is called *moment duality.* This moment duality allows us to calculate all the moments of the Fisher-Wright diffusion at a given time t in terms of the death process at time t, which is simple to analyse. Note that the duality relation also expresses a relation between the backward and the forward processes.

State-dependent resampling rates. In the Fisher-Wright model individuals resample at rate 1. However, it is natural to allow for resampling rates that depend on the state of the population. To do this, let $g : [0,1] \to \mathbb{R}_{\geq 0}$ be any function that satisfies

- $g(0) = g(1) = 0$,
- $g(x) > 0$ for $x \in (0,1)$,
- g is Lipschitz continuous on $[0,1]$.

The evolution of the continuum model with resampling function g is given by

$$\mathrm{d}x(t) = \sqrt{g(x(t))}\,\mathrm{d}w(t), \tag{1.7}$$

where $(w(t))_{t\geq 0}$ is a standard Brownian motion. If we choose $g(x) = x(1-x)$, we recognize the Fisher-Wright diffusion in (1.3). Since the Fisher-Wright diffusion resamples at rate 1, the resampling rate in state x for the model in (1.7) is $\frac{g(x)}{x(1-x)}$. The first condition on g ensures that once the genetic diversity is lost, i.e., there are only $\heartsuit$ or $\diamondsuit$ left in the population, it cannot return.

The drawback of the continuum model in (1.7) is that it does not have a duality relation as in (1.6). Therefore this model is more difficult to analyse. Comparing the continuum model in (1.7) with the continuum models where $g = dx(1-x)$ for some constant $d \in (0,\infty)$, we are still able to analyse (1.7). This technique is called *comparison.*

Extensions of the Fisher-Wright model. The Fisher-Wright model can be extended in several ways to include other evolutionary forces. For example, selection of a fitter type and mutation of genes can be included. Also, more than two gene types can be included, which leads to the *multi-type Fisher-Wright model*. The extension to infinitely many gene types is called the *Fleming Viot model*. In the spatial Fisher-Wright model, there are multiple colonies, each evolving according to the Fisher-Wright model, and individuals are allowed to migrate. For all these extensions, extensive research was done. For an overview of the state of the art we refer the reader to [4]. The addition of a seed-bank to the Fisher-Wright model is relatively new and was introduced in 2016 in [12]. The Fisher-Wright model with seed-bank will be the building block of the spatial models considered in this thesis, and is introduced in the next section.

§1.2.2 The Fisher-Wright model with seed-bank

The Fisher-Wright model with (strong) seed-bank defined in [12] consists of a *single colony* with $N \in \mathbb{N}$ active individuals and $M \in \mathbb{N}$ dormant individuals. Each individual can carry one of two types: $\heartsuit$ or $\diamondsuit$. Let $\epsilon \in [0,1]$ be such that ϵN is integer and $\epsilon N \leq M$. Put $\delta = \frac{\epsilon N}{M}$. The evolution of the population is described by a discrete-time Markov chain that undergoes four transitions per step:

(1) From the N active individuals, $(1-\epsilon)N$ are selected uniformly at random without replacement. Each of these selected individuals resamples, i.e., adopts the type of an active individual selected uniformly at random with replacement, and remains active.

(2) Each of the ϵN active individuals not selected resamples, i.e., adopts the type of an active individual selected uniformly at random with replacement, and subsequently becomes dormant.

(3) From the M dormant individuals, $\delta M = \epsilon N$ are selected uniformly at random without replacement. These selected individuals become active. Since these individuals come from the dormant population, they do not resample.

(4) Each of $(1-\delta)M$ dormant individuals not selected remains dormant and retains its type.

Note that the total sizes of the active and the dormant population remain fixed. During the evolution the dormant and the active population *exchange* individuals. We will refer to the repository of the dormant population as the seed-bank. Fig. 1.3 depicts the first five generations of a population with 5 individuals in the active population and 3 individuals in the dormant population. Fig. 1.3 also shows how in the Fisher-Wright model with seed-bank genetic variability in the active population can be lost, but can be reintroduced again due to the exchange with the dormant population.

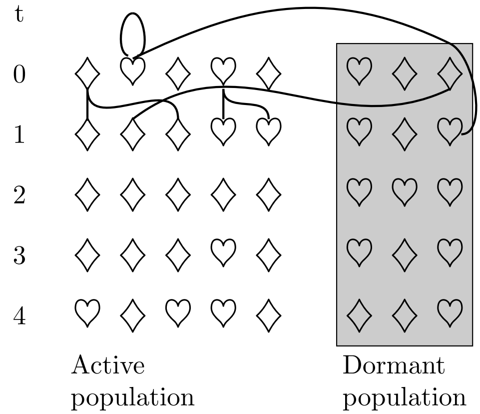

Figure 1.3: Example of the evolution for a population with $N = 5$ active individuals and $M = 3$ dormant individuals. The solid lines within the active population represent resampling, those between the active and the dormant population represent exchange with the seed-bank. Only 1 active individual and 1 dormant individual exchange places per unit of time, which corresponds to $\epsilon = \frac{1}{5}$ and $\delta = \frac{1}{3}$. The relative size of the dormant and the active population is $K = \frac{3}{5}$. Note that the genetic diversity in the active population is lost in generation $t = 2$, but returns in generation $t = 3$ via the seed-bank.

Evolution equation. To formally describe the Fisher-Wright model with seed-bank, we keep track of the fractions of individuals of type $\heartsuit$ in the active and the dormant population. Let $c = \epsilon N = \delta M$, i.e., c is the number of pairs of individuals that change state. Label the N active individuals from 1 to N and the M dormant individuals from 1 up to M. Write $[N] = \{1, \dots, N\}$ and $[M] = \{1, \dots, M\}$. Let $\xi(k) = (\xi_j(k))_{j\in[N]} \in \{0,1\}^{[N]}$ be the random vector where $\xi_j(k) = 1$ if the j'th active individual is of type $\heartsuit$ at time k and $\xi_j(k) = 0$ if the j'th active individual is of type $\diamondsuit$ at time k. Similarly, we let $\eta(k) = (\eta_j(k))_{j\in[M]} \in \{0,1\}^{[M]}$ be the random vector where $\eta_j(k) = 1$ if the j'th dormant individual is of type $\heartsuit$ at time k and $\eta_j(k) = 0$ if the j'th dormant individual is of type $\diamondsuit$ at time k. Define

$$\begin{aligned} X^N(k) &= \frac{1}{N} \sum_{j\in[N]} \mathbf{1}_{\{\xi_j(k)=1\}}, \\ Y^M(k) &= \frac{1}{M} \sum_{j\in[M]} \mathbf{1}_{\{\eta_j(k)=1\}}. \end{aligned} \tag{1.8}$$

Like for the Fisher-Wright model without seed-bank, we can pass to the continuum model. To do so we let both the active and the dormant population size tend to infinity, while keeping their relative sizes fixed, and speed up time proportional by the size of the active population, i.e.,

$$\lim_{N\to\infty} \mathcal{L}\left[\left(X^N(\lfloor Nt\rfloor), Y^M(\lfloor Nt\rfloor)\right)_{t\geq 0}\right] = \mathcal{L}\left[(x(t), y(t))_{t\geq 0}\right]. \tag{1.9}$$

Define

$$K = \frac{\text{size dormant population}}{\text{size active population}} = \frac{M}{N}, \tag{1.10}$$

which is the relative size of the dormant population compared to the active population. It was shown in [12] that the limiting process $(x(t), y(t))_{t\geq 0}$ in (1.9) evolves according to the stochastic differential equation

$$\begin{aligned} \mathrm{d}x(t) &= c\,[y(t) - x(t)]\,\mathrm{d}t + \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t), \\ \mathrm{d}y(t) &= \frac{c}{K}\,[x(t) - y(t)]\,\mathrm{d}t, \end{aligned} \tag{1.11}$$

where $(w(t))_{t\geq 0}$ is a standard Brownian motion. The first term in the first line of (1.11) and the term in the second line of (1.11) describe the exchange of active and dormant individuals in the population. The second term in the first line of (1.11) describes the resampling in the active population. Note that the dormant population does not resample and hence evolves only due the exchange with the active population. In [12] it was shown that in the continuum Fisher-Wright model with seed-bank eventually only one type is left.

For later generalisations, we define the exchange rate

$$e = \frac{c}{K}, \tag{1.12}$$

and rewrite equation (1.11) as

$$\begin{aligned} \mathrm{d}x(t) &= Ke\,[y(t) - x(t)]\,\mathrm{d}t + \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t), \\ \mathrm{d}y(t) &= e\,[x(t) - y(t)]\,\mathrm{d}t. \end{aligned} \tag{1.13}$$

The continuum process $(x(t), y(t))_{t\geq 0}$ evolving according to (1.13) will be the building block of the models analyzed in this thesis. The evolution given in (1.13) is schematically depicted in Fig. 1.4.

The seed-bank coalescent. To describe the genealogy of the continuum Fisher-Wright model with seed-bank, we sample n active individuals and m dormant individuals from the population at some large time and describe their ancestral lineages. We distinguish between active and dormant lineages by giving them labels A for active and D for dormant. The lineages behave according to the following rules:

- Each pair of active lineages coalesces at rate 1, independently of all other lineages.
- Each active lineage becomes dormant at rate Ke.
- Each dormant lineage becomes active at rate e.

Note that dormant lineages cannot coalesce: they can only become active.

Formally the lineages are decribed by the so-called *seed-bank coalescent* that was introduced in [12]. Like the Kingman coalescent, the seed-bank coalescent is

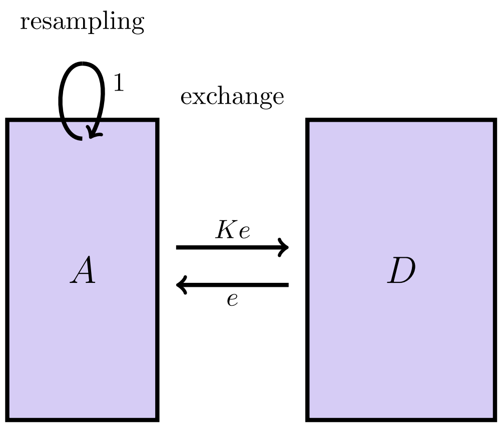

Figure 1.4: Schematic picture of the Fisher-Wright diffusion with seed-bank in (1.13)*. Active individuals resample at rate 1 and become dormant at rate Ke. Dormant individuals become active at rate e.*

a partition-valued process, but in the seed-bank coalescent each partition element is labeled by A or D to indicate whether the corresponding lineage is active or dormant. In [12] it was shown that, as n and m tend to infinity, the ancestral lineages no longer have a common ancestor a finite time back. Hence *the seed-bank coalescent does not come down from infinity in finite time.* This result shows that the Fisher-Wright model with seed-bank behaves qualitatively differently than the Fisher-Wright model seed-bank.

Duality for the seed-bank model. For the Fisher-Wright model with seed-bank we have a similar duality relation as for the Fisher-Wright model without seed-bank. Let

$$\begin{aligned} N(s) &= \#\text{ active lineages left at time } s,\\ M(s) &= \#\text{ dormant lineages left at time } s. \end{aligned} \tag{1.14}$$

Then the block-counting process $(N(s), M(s))_{s\geq 0}$ has transition rates

$$(n,m) \to \begin{cases} (n-1,m), & \text{at rate } \binom{n}{2},\\ (n-1,m+1), & \text{at rate } nKe,\\ (n+1,m-1), & \text{at rate } me. \end{cases} \tag{1.15}$$

Let $(x(t), y(t))_{t\geq 0}$ be the Fisher-Wright diffusion with seed-bank evolving according to (1.13) and starting from state $(x,y) \in [0,1]^2$. Let $(N(t), M(t))_{t\geq 0}$ be the block-counting process starting from state $(n,m) \in \mathbb{N}^2$. Then, for all $(n,m) \in \mathbb{N}^2$, $(x,y) \in [0,1]^2$ and $t \geq 0$, the following relation holds:

$$\mathbb{E}^{(x,y)}[(x(t))^n (y(t))^m] = \mathbb{E}^{(n,m)}[x^{N(t)} y^{M(t)}]. \tag{1.16}$$

Here, the expectation on the left-hand side is taken over the Fisher-Wright diffusion with seed-bank, and the expectation on the right-hand side is taken over the block-counting process starting in (n, m). Thus, also the Fisher-Wright model with seed-bank has a *moment dual*.

Wake-up time distribution of individuals. It has been recognised that qualitatively different behaviour may occur when the wake-up time of individuals in the seed-bank changes from having a thin tail to having a fat tail [55], [50]. Fat-tailed behaviour of the wake-up times is observed in colonies of bacteria. The drawback of the Fisher-Wright model with seed-bank is that it gives thin tails for the wake-up time of individuals. If we define the wake-up time

$$\tau = \text{ time a lineage spends in the seed-bank before it wakes up again,} \tag{1.17}$$

then

$$\tau \stackrel{d}{=} \mathrm{EXP}(e). \tag{1.18}$$

In Section 1.2.3 we will show how we can adapt the Fisher-Wright model with seed-bank to allow for more general wake-up times *without loosing the Markov property.*

Extensions of the Fisher-Wright model with seed-bank. Eventhough the addition of a seed-bank to the Fisher-Wright model is relatively new, in the past five years extensive research was done on extensions of the Fisher-Wright model with seed-bank. An overview of the state of the art is given in [54].

§1.2.3 The Fisher-Wright model with multi-layer seed-bank

A key idea in this thesis is that we can *enrich the seed-bank with internal states* to allow for fat tails and still preserve the Markov property for the evolution. To give the seed-bank an internal structure, we colour the dormant individuals with countably many colours $m \in \mathbb{N}_0$. Thus, instead of one seed-bank we have an infinite sequence of seed-banks, each with its own colour. Active individuals that become dormant are assigned a colour m at rate e_m. If an active individual is assigned a colour m, then it exchanges with a dormant individual of colour m. The colour m-dormant individual loses its colour when it becomes active, but retains its type. Therefore, during the evolution the relative size of the active population and the m-coloured seed-bank is fixed.

Evolution equation. Define

$$K_m = \frac{\text{size } m\text{-dormant population}}{\text{size active population}}, \qquad m \in \mathbb{N}_0, \tag{1.19}$$

which denotes the relative size of the m-dormant population with respect to the active population. Let $(x(t))_{t\geq 0}$ denote the fraction of $\heartsuit$ in the active population at time t, and $(y_m(t))_{t\geq 0}$ the fraction of $\heartsuit$ in the m-dormant population at time t. So, we keep

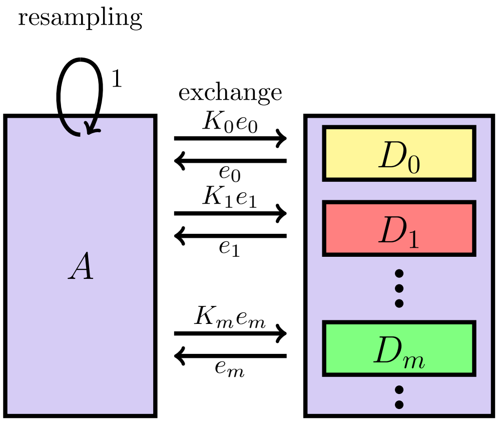

Figure 1.5: Schematic picture of the Fisher-Wright diffusion with layered seed-bank in (1.20)*. Active individuals resample at rate* 1*, but exchange with a countable sequence of dormant populations. At rate K_me_m an active individual becomes dormant with colour m. An m-dormant individual becomes active at rate e_m.*

track of the complete sequence of dormant populations. In the continuum limit, the process $(x(t), (y_m(t))_{m\in\mathbb{N}_0})_{t\geq 0}$ evolves according to

$$\begin{aligned} \mathrm{d}x(t) &= \sum_{m\in\mathbb{N}_0} K_m e_m \,[y_m(t) - x(t)]\,\mathrm{d}t + \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t), \\ \mathrm{d}y_m(t) &= e_m\,[x(t) - y_m(t)]\,\mathrm{d}t, \qquad m \in \mathbb{N}_0. \end{aligned} \tag{1.20}$$

Comparing (1.20) to (1.13), we see that the active population exchanges with the whole sequence of dormant populations. However, each dormant population only evolves due to exchange with the active population. To ensure that active individuals do not become dormant instantaneously, we must assume that

$$\sum_{m\in\mathbb{N}_0} K_m e_m < \infty. \tag{1.21}$$

The evolution in (1.20) is depicted in Fig. 1.5.

Genealogy. Like for the Fisher-Wright model with seed-bank, we can describe the genealogy of the population. At a large time we sample from the population n active individuals, labeled by A, and k_m m-dormant individuals, labeled by D_m, for $m \in \mathbb{N}_0$. Then the lineages of the sampled individuals evolve according to the following rules.

- Each pair of active lineages coalesces at rate 1, independently of all other lineages.
- Each active lineage becomes m-dormant at rate K_me_m.
- Each m-dormant lineage becomes active at rate e_m.

Similarly as for the Fisher-Wright model with (non-layered) seed-bank, we can define a layered seed-bank coalescent and a corresponding block-counting process.

Fat-tailed wake-up times. We define χ to be the total rate at which an active lineage becomes dormant, i.e.,

$$\chi = \sum_{m\in\mathbb{N}_0} K_m e_m. \tag{1.22}$$

Note that $\chi < \infty$ by (1.21). The distribution of the wake-up time τ defined in (1.17) for a lineage in the multi-layer seed-bank is given by

$$\mathbb{P}(\tau > t) = \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{\chi}\, \mathrm{e}^{-e_m t}. \tag{1.23}$$

Choosing the relative sizes of the seed-banks $(K_m)_{m\in\mathbb{N}_0}$ and the rates of exchange $(e_m)_{m\in\mathbb{N}_0}$ properly, we can mimic different wake-up time distributions. For example, we can choose

$$\begin{aligned} &K_m \sim Am^{\alpha}, \qquad e_m \sim Bm^{-\beta},\ m \to \infty \\ &A,\, B \in (0,\infty),\, \alpha,\, \beta \in \mathbb{R},\, \alpha \le 1 < \alpha+\beta, \end{aligned} \tag{1.24}$$

where $\sim$ means asymptotically equal. Then

$$\mathbb{P}(\tau > t) \sim Ct^{-\gamma}, \qquad t \to \infty, \tag{1.25}$$

where $\gamma = \frac{\alpha+\beta-1}{\beta}$ and $C = \frac{A}{\chi\beta} B^{1-\gamma}\Gamma(\gamma)$, with Γ the Gamma-function. Therefore we can choose the sizes K_m and the rates e_m such that when we take the colours into account we still have a Markov process, but when we ignore the colours in the seed-bank we have a wake-up time that is fat-tailed.

Using the layered seed-bank model, we can choose the sequences $(K_m)_{m\in\mathbb{N}_0}$ and $(e_m)_{m\in\mathbb{N}_0}$ such that $0 < \gamma < \frac{1}{2}$ in (1.25). It turns out that in this case the Fisher-Wright diffusion with seed-bank in (1.20) no longer eventually reach its fixed points. (This was also observed in [11] for a non-Markovian seed-bank model.) Hence, in this parameter regime, *the layered seed-bank can prevent loss of genetic variability.*

A key quantity in the Fisher-Wright model with layered seed-bank is the relative size of the total seed-bank with respect to the active population:

$$\rho = \sum_{m\in\mathbb{N}_0} K_m. \tag{1.26}$$

The case $\rho = \infty$ shows different behaviour than the case $\rho < \infty$ (also this was observed in [11]). For example, for the expected wake-up time τ is we find

$$\mathbb{E}[\tau] = \frac{\rho}{\chi}. \tag{1.27}$$

In the next section we will turn to the layered Fisher-Wright model with seed-bank in the spatial setting. Also there the cases $\rho < \infty$ and $\rho = \infty$ give rise to qualitatively different long-term behaviour.

§1.3 Summary of Part I

In Part I of this thesis we consider a *spatial* version of the continuum Fisher-Wright models with seed-bank introduced in Sections 1.2.2–1.2.3. In the spatial version individuals live in colonies, each with their own seed-bank, and are allowed to *migrate* between colonies. The underlying geographic space is a countable Abelian group $\mathbb{G}$.

The spatial Fisher-Wright model without seed-bank has been the object of intense study. A sample of relevant papers and overviews is [66], [17], [20], [25], [22], [33], [29], [27], [24], [41]. In these papers the convergence of the system to equilibrium was proven. Parameter regimes were identified in which the spatial system converges to a mono-type equilibrium, i.e., the system grows locally mono-type clusters of only $\heartsuit$ or of only $\diamondsuit$, or in which the system converges to a multi-type equilibrium, i.e., locally both types are present. The first type of long-term behaviour is called *clustering*, the second type is called *coexistence.* It was shown that the dichotomy between clustering and coexistence for the spatial Fisher-Wright model without seed-bank is completely determined by the *migration kernel* according to which individuals migrate between colonies. If the migration kernel is transient, then coexistence prevails, while if the migration kernel is recurrent, then clustering prevails.

We expect that the presence of the seed-bank affects the long-time behaviour of the spatial system not only quantitatively but also qualitatively. To understand how this comes about, we must find ways to deal with the *richer behaviour* of the population caused by the motion in and out of the seed-bank. In [28] a spatial model with seed-bank, migration and mutation was analysed. There the probability for two individuals drawn randomly from two colonies to be identical by decent was computed as a function of the distance between the colonies.

The first goal in Part I is to prove convergence to equilibrium for the spatial Fisher-Wright model with seed-bank, and to identify the parameter regimes for clustering and coexistence. The second goal is to identify the role of the wake-up time. We will show that if the expected wake-up time is finite, then the dichotomy between clustering and coexistence is completely determined by the migration kernel and the seed-bank has only a quantitative effect on the long-term behaviour. However, if we allow the wake-up time to have infinite mean and moderately fat tails, then both the exchange rates with the seed-bank and the migration kernel determine the dichotomy. In that case the seed-bank has both a quantitative and a qualitative effect on the long-term behaviour. If the wake-up time has very fat tails, then the dichotomy is completely determined by the seed-bank, independently of the migration kernel.

In what follows we first introduce three models of increasing generality that are studied in Part I. After that we state the main results of Part I in words and briefly comment on the techniques used.

Basic ingredients for the models in Part I. We extend the continuum Fisher-Wright models with seed-bank introduced in Sections 1.2.2–1.2.3 to three spatial models of increasing generality. In each of the three models, we consider populations of individuals of two types – either $\heartsuit$ or $\diamondsuit$ – located in colonies on a *geographic space* $\mathbb{G}$ that is a countable Abelian group. In each of the three models, the population in a colony consist of an *active* part and a *dormant* part. The *repository* of the dormant population at colony $i \in \mathbb{G}$ is called the seed-bank at $i \in \mathbb{G}$. Individuals in the active part of colony $i \in \mathbb{G}$ can resample, migrate and exchange with a dormant population. Individuals in the dormant part of colony $i \in \mathbb{G}$ can only exchange with the active population. An active individual that resamples chooses uniformly at random another individual from its colony and adopts its type. The rate of resampling can be state-dependent and is controlled by a diffusion function $g : [0,1] \to \mathbb{R}_{\geq 0}$ that satisfies the criteria for state-dependent resampling in Section 1.2.1. When an active individual at colony $i \in \mathbb{G}$ migrates, it chooses a parent from another colony $j \in \mathbb{G}$ and adopts its type. In each of the three models the migration is described by a *migration kernel* $a(\cdot,\cdot)$, which is an irreducible $\mathbb{G} \times \mathbb{G}$ matrix of transition rates satisfying

$$a(i,j) = a(0,j-i) \quad \forall\, i,j \in \mathbb{G}, \qquad \sum_{i \in \mathbb{G}} a(0,i) < \infty. \tag{1.28}$$

Here, $a(i,j)$ has to be interpreted as the rate at which an active individual at colony $i \in \mathbb{G}$ chooses an individual in the active part of colony $j \in \mathbb{G}$ and adopts its type. An active individual that becomes dormant *exchanges* with a randomly chosen dormant individual that becomes active. The dormant part of the population only evolves due to exchange of individuals with the active part of the population.

The three models we introduce below differ in the way the active population exchanges with the dormant population. However, in each of the three models the exchange mechanism guarantees that the sizes of the active and the dormant population stay fixed over time.

Since we look at continuum models obtained from individual-based models, we are interested in the fraction of individuals of type $\heartsuit$ in the different colonies.

Model 1: single-layer seed-bank. In this model we consider a multi-colony version of the continuum model in (1.13). Each colony $i \in \mathbb{G}$ has an active part A and a dormant part D. For $i \in \mathbb{G}$ and $t \geq 0$, let $x_i(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are active at time t, and $y_i(t)$ the fraction of individuals in colony i of type $\heartsuit$ that are dormant at time t. Like in (1.10) in Section 1.2.2, let $K \in (0,\infty)$ be the relative size of the dormant population w.r.t. the active population, and let e be the rate at which active and dormant individuals exchange. We assume K and e to be the same for all colonies. The fractions of individuals of type $\heartsuit$ in the population evolve according to the system of stochastic differential equations (SSDE)

$$\begin{aligned} \mathrm{d}x_i(t) &= \sum_{j \in \mathbb{G}} a(i,j)\,[x_j(t) - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) && (1.29)\\ &\quad + Ke\,[y_i(t) - x_i(t)]\,\mathrm{d}t, \\ \mathrm{d}y_i(t) &= e\,[x_i(t) - y_i(t)]\,\mathrm{d}t, \qquad i \in \mathbb{G}, && (1.30) \end{aligned}$$

where $(w_i(t))_{t\geq 0}$, $i \in \mathbb{G}$, are independent standard Brownian motions. The first term in (1.29) describes the *migration* of active individuals from i to j at rate $a(i,j)$. The second term in (1.29) describes the *resampling* of individuals at rate $\frac{g(x)}{x(1-x)}$ in state x. The third term in (1.29) together with the term in (1.30) describe the *exchange* of active and dormant individuals at rate $e \in (0, \infty)$.

Model 2: multi-layer seed-bank. In this model we consider a multi-colony version of the continuum model in (1.20). Therefore, for each $i \in \mathbb{G}$ a colony now consists of an active part A and a sequence $(D_m)_{m \in \mathbb{N}_0}$ of dormant parts, labeled by their colour $m \in \mathbb{N}_0$. As before, for $i \in \mathbb{G}$, let $x_i(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are active at time t, but now let $y_{i,m}(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are dormant with colour m at time t. Let $e_m \in (0, \infty)$ be the rate at which active individuals exchange with dormant individuals of colour m, and let $K_m \in (0, \infty)$ denote the relative size of the m-dormant population w.r.t. the active population as in (1.19). We assume e_m and K_m to be the same for all colonies. Then the fraction of $\heartsuit$ in the population evolves according to the SSDE

$$\begin{aligned} \mathrm{d}x_i(t) &= \sum_{j \in \mathbb{G}} a(i,j)\,[x_j(t) - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) \\ &\quad + \sum_{m \in \mathbb{N}_0} K_m e_m\,[y_{i,m}(t) - x_i(t)]\,\mathrm{d}t, \end{aligned} \tag{1.31}$$

$$\mathrm{d}y_{i,m}(t) = e_m\,[x_i(t) - y_{i,m}(t)]\,\mathrm{d}t, \qquad m \in \mathbb{N}_0,\ i \in \mathbb{G}, \tag{1.32}$$

where $(w_i(t))_{t \geq 0}, i \in \mathbb{G}$, are independent standard Brownian motions. Comparing (1.31)–(1.32) with (1.29)–(1.30), we see that active individuals migrate (the first term in (1.31)) and resample (the second term in (1.31)) in the same way, but now interact with a sequence of dormant populations (the third term in (1.31) and the term (1.32)). The dormant individuals only exchange with the active individuals.

Model 3: multi-layer seed-bank with displaced seeds. We can extend the mechanism of Model 2 by allowing active individuals that become dormant to do so in a randomly chosen colony. This amounts to introducing a sequence of irreducible *displacement kernels* $a_m(\cdot,\cdot)$, $m \in \mathbb{N}_0$, satisfying

$$a_m(i,j) = a_m(0, j-i) \quad \forall\, i,j \in \mathbb{G}, \qquad \sum_{i \in \mathbb{G}} a_m(0,i) = 1 \quad \forall\, m \in \mathbb{N}_0, \tag{1.33}$$

and replacing (1.31)–(1.32) by

$$\begin{aligned} \mathrm{d}x_i(t) &= \sum_{j \in \mathbb{G}} a(i,j)\,[x_j(t) - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) \\ &\quad + \sum_{j \in \mathbb{G}} \sum_{m \in \mathbb{N}_0} K_m e_m\, a_m(j,i)\,[y_{j,m}(t) - x_i(t)]\,\mathrm{d}t, \end{aligned} \tag{1.34}$$

$$\mathrm{d}y_{i,m}(t) = \sum_{j \in \mathbb{G}} e_m\, a_m(i,j)\,[x_j(t) - y_{i,m}(t)]\,\mathrm{d}t, \qquad m \in \mathbb{N}_0,\ i \in \mathbb{G}. \tag{1.35}$$

Here, the third term in (1.34) together with the term in (1.35) describe the *switch of colony* when individuals exchange between active and dormant. Namely, with probability $a_m(i,j)$ simultaneously an active individual in colony i becomes dormant with colour m in colony j and a randomly chosen dormant individual with colour m in colony j becomes active in colony i.

Two key quantities. Like in Section 1.2.3, in Models 2 and 3 we must assume that

$$\chi = \sum_{m \in \mathbb{N}_0} K_m e_m < \infty \tag{1.36}$$

in order to make sure that active individuals do not become dormant instantly. Like in Section 1.2.3, in Models 2 and 3 we set

$$\rho = \sum_{m \in \mathbb{N}_0} K_m = \frac{\text{size dormant population}}{\text{size active population}}. \tag{1.37}$$

It turns out that χ and ρ are key quantities of the system. In particular, we will see that the long-time behaviour of Models 2 and 3 is different for $\rho < \infty$ and $\rho = \infty$.

Main results of Part I.

(1) For all three models the system convergences to a unique equilibrium that exhibits a dichotomy between *clustering* and *coexistence*. In all three models the density of $\heartsuit$ in the population is preserved over time, and in both the clustering case and the coexistence case the equilibrium depends on the initial density of $\heartsuit$. In the coexistence case the limiting equilibrium also depends on the resampling function, the exchange rates with the seed-bank and the migration kernel.

(2) For all three models we identify the parameter regimes for clustering and coexistence. These parameter regimes do not depend on the resampling function g.

   (2a) In Model 1 the wake-up time has *finite mean.* The dichotomy between coexistence and clustering is controlled by the migration only and the seed-bank has no effect on the dichotomy. In particular, clustering prevails when the symmetrised migration kernel is recurrent, while coexistence prevails when it is transient. This result is the classical dichotomy for populations without seed-bank [14].

   (2b) In Model 2 the wake-up time can have both *finite mean* and *infinite mean.* If the wake-up time has finite mean, then the dichotomy between coexistence and clustering is controlled by the migration only and the seed-bank has no effect, similarly as for Model 1. If the wake-up time has *infinite mean* with *moderately fat tails*, then the dichotomy is controlled by both the migration and the seed-bank. In particular, the parameter regimes for clustering and coexistence reveal an interesting interplay between rates for migration and rates for exchange with the seed-bank. If the wake-up time has *infinite mean* with *very fat tails*, then the dichotomy is controlled by the seed-bank only and the migration has no effect. For *infinite mean* wake-up times it turns out that in the coexistence regime the seed-banks with colour $m \to \infty$ are in a state that is almost surely equal to the initial density of $\heartsuit$ in the population. Therefore *deep seed-banks become deterministic.*

(2c) In Model 3, the extra migration of active individuals that become dormant makes coexistence more likely. This extra migration can be incorperated into the dichotomy criterion obtained for model 2.

Techniques used in Part I. To prove the dichotomy, we first consider the three models when the resampling function is $g(x) = dx(1-x)$, $d \in \mathbb{R}_+$. For these diffusion functions we have duality relations similarly as those introduced in Section 1.2. As will be explained in detail in Part I, the lineages in the dual of the spatial model behave like a set of coalescing random walks. Afterwards, we can use comparison techniques to extend the results to a general resampling function g. To prove convergence to a unique equilibrium in the case of coexistence, we make use of *coupling* techniques.

§1.4 Summary of Part II

Part II of this thesis focuses on spatial populations with seed-bank where the underlying geographic space is the so-called *hierarchical group*, which we introduce next.

Hierachical lattice. The hierarchical lattice of order N is given by

$$\Omega_N = \left\{ \xi = (\xi_k)_{k\in\mathbb{N}_0} \colon \xi_k \in \{0,1,\dots,N-1\}, \sum_{k\in\mathbb{N}_0} \xi_k < \infty \right\}, \tag{1.38}$$

which with addition modulo N becomes the hierarchical group of order N (see Fig. 1.6). The *hierarchical distance* is defined by

$$d_{\Omega_N}(\xi,\eta) = d_{\Omega_N}(0,\xi-\eta) = \min\{k \in \mathbb{N}_0 \colon \xi_l = \eta_l \ \forall\, l \geq k\}, \qquad \xi,\eta \in \Omega_N. \tag{1.39}$$

Intuitively, depicting Ω_N as the leaves of an infinite tree as in Fig. 1.6, the distance between two points on Ω_N is the number of branches we have to travel upwards in the tree to find a common node.

The choice of Ω_N as geographic space plays an important role for population models, and was first exploited in [65] in an attempt to formalise ideas coming from ecology. One interpretation is that the sequence $(\xi_k)_{k\in\mathbb{N}_0}$ encodes the 'address' of colony ξ: ξ_0 is the 'house', ξ_1 is the 'street', ξ_2 is the 'village', ξ_3 is the 'province', ξ_4 is the 'country', and so on.

The goal of Part II of this thesis is to study the spatial Fisher-Wright model with seed-bank on the hierarchical lattice when the order N of the hierarchical group tends to infinity. This limit is called the *hierarchical mean-field limit*. To analyse the limiting system, it turns out that we have to consider different *space-time scales*. In what follows we first set up the model. After that we explain how the different space-time scales come into play in a natural way.

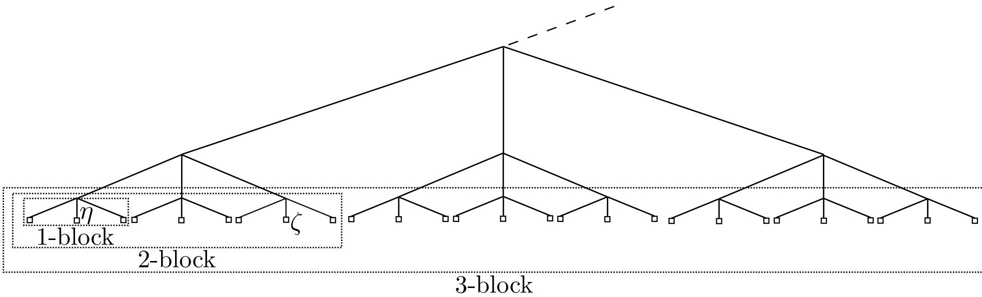

Figure 1.6: Close-ups of a 1-block, a 2-block and a 3-block in the hierarchical group of order $N = 3$. The elements of the group are the leaves of the tree (indicated by $\Box$'s). The hierarchical distance between two elements in the group is the graph distance to the most recent common ancestor in the tree: $d_{\Omega_3}(\eta, \zeta) = 2$ for η and ζ in the picture.

Hierarchical migration. We construct a migration kernel $a^{\Omega_N}(\cdot, \cdot)$ on the hierarchical group Ω_N built from a sequence of migration rates

$$\underline{c} = (c_k)_{k\in\mathbb{N}_0} \in (0, \infty)^{\mathbb{N}_0} \tag{1.40}$$

that do not depend on N. Individuals migrate as follows:

- For all $k \in \mathbb{N}$, each individual chooses at rate c_{k-1}/N^{k-1} the block of radius k around its present location and selects a colony uniformly at random from that block. Subsequently it selects an individual in this colony uniformly at random and adopts its type.

Since the block of radius k contains N^k colonies, the migration kernel is given by

$$a^{\Omega_N}(\eta, \xi) = \sum_{k \geq d_{\Omega_N}(\eta,\xi)} \frac{c_{k-1}}{N^{k-1}} \frac{1}{N^k}, \quad \eta, \xi \in \Omega_N, \eta \neq \xi, \qquad a^{\Omega_N}(\eta, \eta) = 0, \quad \eta \in \Omega_N. \tag{1.41}$$

We assume that

$$\sum_{\xi\in\Omega_N} a^{\Omega_N}(\eta, \xi) < \infty \tag{1.42}$$

to guarantee that the total migration rate per individual is finite.

Evolution on the hierarchical lattice. The evolution of the single colonies in Part II is similar to the evolution of the single colonies in Part I in Model 2. The difference is that in the hierarchical setting we let the migration rates and the exchange rates depend on the order N of the hierarchical group. For $\xi \in \Omega_N$, define

$$\begin{aligned} x_\xi(t) &= \text{ the fraction of active individuals of type } \heartsuit \text{ at colony } \xi \text{ at time } t, \\ y_{\xi,m}(t) &= \text{ the fraction of } m\text{-dormant individuals of type } \heartsuit \text{ at colony } \xi \text{ and time } t. \end{aligned} \tag{1.43}$$

Active individuals exchange with dormant individuals with colour m at rate $\frac{e_m}{N^m}$, where $e_m \in (0, \infty)$. Like in (1.19), let K_m be the relative size of the m-dormant population with respect to the active population. Like in Part I, the sequences $(K_m)_{m\in\mathbb{N}_0}$

and $(e_m)_{m\in\mathbb{N}_0}$ are the same for all colonies. Also here we allow for a general diffusion function $g : [0, 1] \to \mathbb{R}_{\geq 0}$ satisfying the conditions in Section 1.2.1.

The fraction of individuals of type $\heartsuit$ in the population evolves according to the SSDE

$$\begin{aligned}
\mathrm{d}x_\xi(t) &= \sum_{\eta\in\Omega_N} a^{\Omega_N}(\xi,\eta)[x_\eta(t) - x_\xi(t)]\,\mathrm{d}t + \sqrt{g(x_\xi(t))}\,\mathrm{d}w_\xi(t)\\
&\qquad + \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{N^m}\,[y_{\xi,m}(t) - x_\xi(t)]\,\mathrm{d}t, \qquad (1.44)\\
\mathrm{d}y_{\xi,m}(t) &= \frac{e_m}{N^m}\,[x_\xi(t) - y_{\xi,m}(t)]\,\mathrm{d}t, \quad m \in \mathbb{N}_0, \qquad \xi \in \Omega_N,
\end{aligned}$$

where we assume that $\sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{N^m} < \infty$. The first term in the first equation describes the evolution of the active population at colony ξ due to migration, the second term due to resampling. The third term in the first equation and the term in the second equation describe the exchange between the active and the dormant population at colony ξ (see Fig. 1.7). Like the migration rates, the exchange rates between the active and the dormant population depend on the order N of the hierarchical group. Since dormant individuals are not subject to resampling or migration, the dynamics of the dormant population is completely determined by the exchange with the active population.

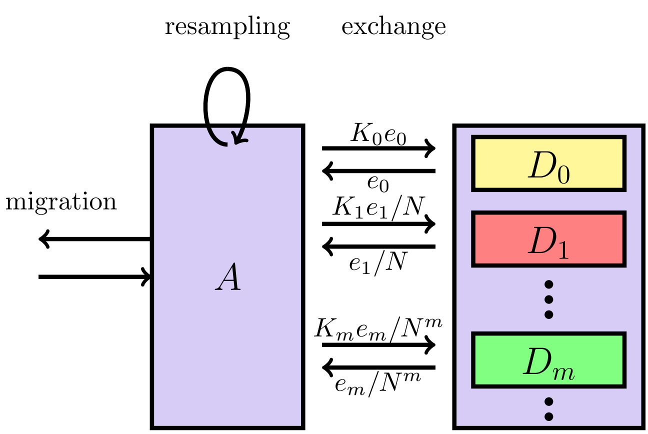

Figure 1.7: Active individuals (A) are subject to migration, resampling and exchange with dormant individuals (D). When active individuals become dormant they are assigned a colour (D_m, $m \in \mathbb{N}_0$), which they lose when they become active again. The resampling rate in the active popluation depends on the type-$\heartsuit$ frequency x and equals $g(x)/x(1-x)$.

Evolution of block averages. The choice of the migration kernel in (1.41) implies that, for every $k \in \mathbb{N}$, at rate $\asymp \frac{1}{N^k}$ individuals choose a space horizon of distance $k+1$ and subsequently choose a random colony from that space horizon. Therefore, in order to see interactions over a distance $k+1$ for large N, we need to speed up time by a factor N^k. A similar observation applies to the interaction with the seed-bank. Dormant individuals with colour k become active at rate $\asymp \frac{1}{N^k}$. Therefore, in order to see interactions with the k-dormant population for large N, we need to speed up time by a factor N^k. To analyse the effective interaction on different time scales N^k, $k \in \mathbb{N}_0$ we introduce *successive block averages.*

Definition 1.4.1. For $k \in \mathbb{N}_0$, let

$$B_k(0) = \{\eta \in \Omega_N \colon\ d_{\Omega_N}(0,\eta) \le k\}$$

denote the k-block around 0. Define the k-block average around 0 at time $N^k t$ by

$$\begin{aligned} x_k^{\Omega_N}(t) &= \frac{1}{N^k} \sum_{\eta \in B_k(0)} x_\eta(N^k t), \\ y_{m,k}^{\Omega_N}(t) &= \frac{1}{N^k} \sum_{\eta \in B_k(0)} y_{\eta,m}(N^k t), \qquad m \in \mathbb{N}_0. \end{aligned} \tag{1.45}$$

■

The k-block average represents the dynamics of the system averaged over the N^k colonies around 0 with time speeded up by a factor N^k. Therefore we say that the k-block average represents the dynamics of the system *on space-time scale* k, or *on hierarchical level* k.

To obtain the *hierarchical mean-field limit*, we analyse the block averages defined above in the limit as $N \to \infty$, for which we expect a seperation of space-time scales. It turns out that in this limit each of the block averages preforms an autonomous diffusion, similar to the diffusion preformed by a single colony.

Main results Part II.

(1) For fixed N the results obtained in Part I are applied to the hierarchical model with seed-bank. For two classes of parameters the clustering regime is identified. In case the wake-up time has infinite mean, the clustering regime exhibits a trade off between the exchange rates and the migration rates.

(2) The hierarchical mean-field limit is identified, i.e., the evolution of the k-block averages defined in (1.45) is determined in the limit $N \to \infty$. For all $k \in \mathbb{N}_0$, the limiting k-block averages evolve according to a k-dependent SSDE, from which we can read off the following results:

  (2a) For each $k \in \mathbb{N}_0$, if $N \to \infty$, then the migration terms in the evolution of the active k-block average can be replaced by a drift towards the active block one level up, i.e., a drift towards the active $(k+1)$-block average. This phenomenon is called *decoupling.*

  (2b) For each $k \in \mathbb{N}_0$, if $N \to \infty$, then the resampling rate of the active k-block average is the average resampling rate of the colonies within the k-block. The resampling rate of the active k-block is given by the k-fold iteration of a *renormalisation transformation* $\mathcal{F}$ applied to the original diffusion function g. The resulting diffusion function $\mathcal{F}^{(k)}g$ is called the *renormalised diffusion function*.

(2c) For each $k \in \mathbb{N}_0$, if $N \to \infty$, then the active k-block exchanges only with the dormant k-block of colour k. Therefore we say that the k-dormant population is the *effective seed-bank* on level k. The k-block averages of dormant populations of colours $m < k$ start equalising with the active k-block averages and are called *fast seed-banks*. The seed-banks of colours $m > k$ do not change and stay in a fixed state. These seed-banks are refered to as *slow seed-banks*.

(2d) For each $k \in \mathbb{N}_0$, if $N \to \infty$, then the migration terms in the evolution of the active k-block averages, induce a drift towards the $(k+1)$-block average. It is in this way that subsequent block averages are connected. The connection between different hierarchical levels is captured by what is called the *interaction chain*.

(3) With the help of the interaction chain, the attracting orbit of the renormalisation transformation acting on the space of diffusion functions is analysed. In the *clustering regime* and after appropriate scaling, the renormalised diffusion function $\mathcal{F}^{(k)}g$ converges to the Fisher-Wright diffusion function as $k \to \infty$, irrespective of the diffusion function g controlling the resampling in the single colonies. This convergence shows that the hierarchical system exhibits *universality* on large space-time scales in terms of the scaling limit. For several subclasses of parameters the scaling of the renormalised diffusion function is identified. This scaling reveals a delicate interplay between the parameters controlling the migration and the seed-bank and also determines the speed at which *mono-type clusters grow in space and time.*

Techniques used in Part II. To prove the results in Part II we use the abstract schemes for mean-field analysis that were introduced in [21], [25] and [15]. To analyse the behaviour of the seed-bank, we use coupling results and a random walk interpretation of the system that is proven in Part I. For the analysis of the renormalisation transformation we proceed like in [5].

§1.5 Further research

Two topics that would be interesting to study are the finite-systems scheme and the genealogy and cluster formation.

Finite-systems scheme. In Part I we analyse a spatial version of the Fisher-Wright model with seed-bank where the underlying space is a countable Abelian group. A key question is how well the infinite systems introduced in Part I can approximate “real-world” finite systems. To answer this question, we use the so-called finite-systems scheme that was introduced in [15]. In the *finite-systems scheme* we trucate both the geographic space and the seed-bank, and subsequently let both the truncation level and the time tend to infinity, properly tuned together.

In an upcoming paper [44], we focus on Model 2, introduced in Section 1.3 and evolving according to (1.31)–(1.32). We study the truncated system in the parameter regime of coexistence for the non-truncated system. This parameter regime is identified in one of the main theorems in Part I. For truncated finite systems, we always get clustering. To obtain a meaningful scaling limit, the truncation level and the time are scaled in such a way that we are just at the verge of seeing the clustering of the finite system coming in. We find that, with this scaling, the finite system behaves like the infinite system in equilibrium in the coexistence regime, but with a density of $\heartsuit$ that is random.

Again we see a difference in behaviour between $\rho < \infty$ and $\rho = \infty$. In case $\rho < \infty$, we can adapt classical techniques to analyse the finite-systems scheme and the results lead to a single universality class. In case $\rho = \infty$, we can tune the speed at which the seed-bank tends to infinity relative to the speed at which the geographic space tends to infinity. This leads to new phenomena, and different universality classes appear.

Cluster formation and genealogy. It would be interesting to study the growth of monotone clusters in the spatial setting. A short introduction to cluster formation will be given in the setting of the hierarchical group in Part II. However, it would be interesting to see what the effect of the seed-bank is on the cluster formation. Closely related to the cluster formation is the genealogy. Scaling of the genealogy leads to continuum random trees with dormancy. We have not yet started this part of the research.

§1.6 Outline of the thesis

Part I of this thesis is based on [43]. It treats the general setting of spatial populations with seed-bank. In Chapter 2 we formally introduce the models of Section 1.3, and subsequently state the main results about well-posedness, duality and long-term behaviour. Chapter 3 is devoted to the proofs of the main results stated in Chapter 2.

Part II is based on the upcoming paper [45]. It treats the spatial hierarchical seed-bank model introduced in Section 1.4. In Chapter 4 we formally introduce the hierarchical seed-bank model and subsequently state the main results about multi-scaling renormalisation and universality. Chapters 5–10 are devoted to the proofs of the main results stated in Chapter 4.

PART I

SPATIAL POPULATIONS WITH SEED-BANK: WELL-POSEDNESS, DUALITY AND EQUILIBRIUM

This part is based on:
A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: well-posedness, duality and equilibrium. *Preprint*, 2020

CHAPTER 2

Spatial populations with seed-bank, models and results

§2.1 Background and outline

§2.1.1 Background and goals

In populations with a seed-bank, individuals can become dormant and stop reproducing themselves, until they can become active and start reproducing themselves again. In [10] and [12], the evolution of a population evolving according to the Fisher-Wright model with a *seed-bank* was studied. In this model individuals are subject to *resampling* and can move in and out of a seed-bank. While in the seed-bank they suspend resampling, i.e., the seed-bank acts as a repository for the genetic information of the population. Individuals that do not reside in the seed-bank are called *active*, those that do are called *dormant*. In the present paper we extend the single-colony Fisher-Wright model with seed-bank introduced in [12] to a multi-colony setting in which individuals live in different colonies and move between colonies. In other words, we introduce *spatialness*.

Seed-banks are observed in many taxa, including plants, bacteria and other micro-organisms. Typically, they arise as a response to unfavourable environmental conditions. The dormant state of an individual is characterised by low metabolic activity and interruption of phenotypic development (see e.g. Lennon and Jones [55]). After a varying and possibly large number of generations, dormant individuals can be resuscitated under more favourable conditions and reprise reproduction after having become active. This strategy is known to have important implications for population persistence, maintenance of genetic variability and stability of ecosystems. It acts as a *buffer* against evolutionary forces such as genetic drift, selection and environmental variability. The importance of this evolutionary trait has led to several attempts to model seed-banks from a mathematical perspective, see e.g. [50], [11], [40], [9]. In [12] it was shown that the continuum model obtained by taking the large-colony-size limit of the individual-based model with seed-bank is the Fisher-Wright diffusion with seed-bank. Also the long-time behaviour and the genealogy of the continuum model with seed-bank were analysed in [12].

In the present paper we consider a *spatial* version of the continuum model with seed-bank, in which individuals live in colonies, each with their own seed-bank, and are allowed to *migrate* between colonies. Our goal is to understand the change in behaviour compared to the spatial model without seed-bank. The latter has been the object of intense study. A sample of relevant papers and overviews is [66], [17], [20], [25], [22], [33], [29], [27], [24], [41]. We expect the presence of the seed-bank to affect the long-time behaviour of the system not only quantitatively but also qualitatively. To understand how this comes about, we must find ways to deal with the *richer behaviour* of the population caused by the motion in and out of the seed-bank. Earlier work on a spatial model with seed-bank, migration and mutation was carried out in [28], where the probability to be identical by decent for two individuals drawn randomly from two colonies was computed as a function of the distance between the colonies.

It has been recognised that qualitatively different behaviour may occur when the wake-up time in the seed-bank changes from having a thin tail to having a fat tail [55]. One challenge in modelling seed-banks has been that fat tails destroy the Markov property for the evolution of the system. A key idea of the present paper is that we can *enrich the seed-bank with internal states* – which we call colours – to allow for fat tails and still preserve the Markov property for the evolution. We will see that fat tails induce *new universality classes.*

The main goals of the present paper are the following:

(1) Identify the typical features of the long-time behaviour of populations with a seed-bank. In particular, prove convergence to equilibrium, and identify the parameter regimes for *clustering* (= convergence towards locally mono-type equilibria) and *coexistence* (= convergence towards locally multi-type equilibria).

(2) Identify the role of finite versus infinite mean wake-up time. Identify the *critical dimension* in case the geographic space is $\mathbb{Z}^d$, $d \geq 1$, i.e., the dimension at which the crossover between clustering and coexistence occurs for migration with finite variance.

    (2a) Show that if the wake-up time has *finite mean*, then the dichotomy between coexistence and clustering is controlled by the migration only and the seed-bank has no effect. In particular, clustering prevails when the symmetrised migration kernel is recurrent while coexistence prevails when it is transient. This is the classical dichotomy for populations without seed-bank [14]. The critical dimension is $d = 2$.

    (2b) Show that if the wake-up time has *infinite mean* with *moderately fat tails*, then the dichotomy is controlled by both the migration and the seed-bank. In particular, the parameter regimes for clustering and coexistence reveal an interesting interplay between rates for migration and rates for exchange with the seed-bank. The critical dimension is $1 < d < 2$.

    (2c) Show that if the wake-up time has *infinite mean* with *very fat tails*, then the dichotomy is controlled by the seed-bank only and the migration has no effect. The critical dimension is $d = 1$.

We focus on the situation where the individuals can be of two types. The extension to infinitely many types, called the Fleming-Viot measure-valued diffusion, only requires standard adaptations and will not be considered here (see [25]). Also, instead of Fisher-Wright resampling we will allow for state-dependent resampling, i.e., the rate of resampling in a colony depends on the fractions of the two types in that colony. In what follows we only work with *continuum* models, in which the components represent *type frequencies* in the colonies labelled by a discrete geographic space.

The techniques of proof that we use include duality, moment relations, semigroup comparisons and coupling. These techniques are standard, but have to be adapted to the fact that individuals move into and out of seed-banks. Since there is no resampling and no migration in the seed-bank, the motion of ancestral lineages in the dual process looses part of the random-walk structure that is crucial in models without seed-bank. Moreover, for seed-banks with infinite mean wake-up times, we encounter *fat-tailed* wake-up time distributions in the dual process, and we need to deal with lineages that are dormant most of the time and therefore are much slower to coalesce. The coupling arguments also change. Already in a single colony, if the seed-bank has infinitely many internal states, then we are dealing with an infinite system in which the manipulation of Lyapunov functions and the construction of successful couplings from general classes of initial states is hard. In the multi-colony setting this becomes even harder, and conceptually challenging issues arise.

§2.1.2 Outline

In Section 2.2 we introduce three models of increasing generality, establish their well-posedness via a martingale problem, and introduce their dual processes, which play a crucial role in the analysis. In Section 2.3 we state our main results. We focus on the long-time behaviour, prove convergence to equilibrium, and establish a *dichotomy* between clustering and coexistence. We show that this dichotomy is *affected by the presence of the seed-bank*, namely, the dichotomy depends not only on the migration rates, but can also depend on the relative sizes of the active and the dormant population and their rates of exchange. In particular, if the dormant population is much larger than the active population, then the residence time in the seed-bank has a fat tail that *enhances genetic diversity significantly*.

Sections 3.1–3.4 are devoted to the proofs of the theorems stated in Sections 2.2–2.3. In Appendix A.1 we give the derivation of the single-colony continuum model from the single-colony individual-based Fisher-Wright model in the large-colony-size limit. In the individual-based model active individuals *exchange* with dormant individuals, i.e., for each active individual that becomes dormant a dormant individual becomes active. In Appendix A.2 we look at the continuum limit of the single-colony individual-based Moran model in which active and dormant individuals no longer exchange state but rather change state independently. We show that change instead of exchange does not affect the long-time behaviour. Appendices A.3 and A.4 contain the proof of technical lemmas that are needed in the proof of the convergence to equilibrium.

In three companion papers we deal with three further aspects:

(I) In [44] we establish the *finite-systems scheme*, i.e., we identify in the coexistence regime how a finite truncation of the system behaves as both the time and the truncation level tend to infinity, properly tuned together. This underlines the relevance of systems with an infinite geographic space and a seed-bank with infinitely many colours for the description of systems with a large finite geographic space and a seed-bank with a large finite number of colours. We show that there is a *single universality class* for the scaling limit, represented by a Fisher-Wright diffusion whose volatility constant is *reduced* by the seed-bank. We show that if the wake-up time has finite mean, then the scaling time is *proportional* to the geographical volume of the system, while if the wake-up time has infinite mean, then the scaling time grows *faster* than the geographical volume of the system. We also investigate what happens for systems with a large finite geographic space and a seed-bank with infinitely many colours, where the behaviour turns out to be different.

(II) In [45] we consider the special case where the colonies are organised in a *hierarchical fashion*, i.e., the geographic space is the hierarchical group Ω_N of order N. We identify the parameter regime for clustering for all $N < \infty$, and analyse the *multi-scale behaviour* of the system in the hierarchical mean-field limit $N \to \infty$ by looking at block averages on successive hierarchical space-time scales. Playing with the migration kernel, we can choose the migration to be close to critically recurrent in the sense of potential theory. By letting $N \to \infty$ we can approach the *critical dimension*, so that the migration becomes similar to migration on the two-dimensional Euclidean geographic space. With the help of *renormalisation arguments* we show that, close to the critical dimension, the scaling behaviour on large space-time scales is *universal*.

(III) Our goal for the fourth paper is to identify the *pattern of cluster formation* in the clustering regime (= how fast mono-type clusters grow in time) and describe the *genealogy* of the population. The latter provides further insight into how the seed-bank enhances genetic diversity.

In these papers too we will see that the seed-bank can cause not only quantitative but also qualitative changes in the scaling behaviour of the system.

§2.2 Introduction of the three models and their basic properties

In Section 2.2.1 we give a formal definition of the three models of increasing generality. In Section 2.2.2 we comment on their biological significance. In Section 2.2.3 we establish their well-posedness via a martingale problem (Theorem 2.2.4). In Section 2.2.4 we introduce the associated dual processes and state the relevant duality relations (Theorems 2.2.5, 2.2.8 and 2.2.10). In Section 2.2.5 we use these duality relations to formulate a criterion for clustering versus coexistence (Theorems 2.2.11 and 2.2.13).

§2.2.1 Migration, resampling and seed-bank: three models

In this section we extend the model for a population with seed-bank from [12] to three models of increasing generality for spatial populations with seed-bank. In each of the three models, we consider populations of individuals of two types – either $\heartsuit$ or $\diamondsuit$ – located in a *geographic space* $\mathbb{G}$ that is a countable Abelian group endowed with the discrete topology. In each of the three models, the population in a colony consist of an *active* part and a *dormant* part. The *repository* of the dormant population at colony $i \in \mathbb{G}$ is called the seed-bank at $i \in \mathbb{G}$. Individuals in the active part of a colony $i \in \mathbb{G}$ can resample, migrate and exchange with a dormant population. Individuals in the dormant part of a colony $i \in \mathbb{G}$ only exchange with the active population. An active individual that resamples chooses uniformly at random another individual from its colony and adopts its type. (Alternatively, resampling may be viewed as the active individual being replaced by a copy of the active individual chosen. Because individuals carry a type and not a label, this gives the same model.) When an active individual at colony $i \in \mathbb{G}$ migrates, it chooses a parent from another colony $j \in \mathbb{G}$ and adopts its type. In each of the three models the migration is described by a *migration kernel* $a(\cdot,\cdot)$, which is an irreducible $\mathbb{G} \times \mathbb{G}$ matrix of transition rates satisfying

$$a(i,j) = a(0,j-i) \quad \forall\, i,j \in \mathbb{G}, \qquad \sum_{i\in\mathbb{G}} a(0,i) < \infty. \tag{2.1}$$

Here, $a(i,j)$ is to be interpreted as the rate at which an active individual at colony $i \in \mathbb{G}$ chooses a parent in the active part of colony $j \in \mathbb{G}$ and adopts its type. An active individual that becomes dormant *exchanges* with a randomly chosen dormant individual that becomes active. The three models we discuss in the present paper differ in the way the active population exchanges with the dormant population. However, in each of the three models the exchange mechanism guarantees that the sizes of the active and the dormant population stay fixed over time. The dormant part of the population only evolves due to exchange of individuals with the active part of the population.

Since we look at continuum models obtained from individual-based models, we are interested in the frequencies of type $\heartsuit$ in the different colonies. In Appendix A.1 we discuss the individual-based models underlying the continuum models described below.

Remark 2.2.1 (Notation). Throughout the paper we use lower case letters for *components* and upper case letters for *systems of components*. ■

Model 1: single-layer seed-bank. Each colony $i \in \mathbb{G}$ has an active part A and a dormant part D. Therefore we say that *the effective geographic space* is given by $\mathbb{G} \times \{A, D\}$. For $i \in \mathbb{G}$ and $t \geq 0$, let $x_i(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are active at time t, and $y_i(t)$ the fraction of individuals in colony i of type $\heartsuit$ that are dormant at time t. Then the system is described by the process

$$(Z(t))_{t\geq 0}, \qquad Z(t) = \big(z_i(t)\big)_{i\in\mathbb{G}}, \qquad z_i(t) = (x_i(t), y_i(t)), \tag{2.2}$$

on the state space

$$E = ([0,1] \times [0,1])^{\mathbb{G}}, \tag{2.3}$$

and $(Z(t))_{t\geq 0}$ evolves according to the following SSDE:

$$\begin{aligned} \mathrm{d}x_i(t) &= \sum_{j\in\mathbb{G}} a(i,j)\,[x_j(t)-x_i(t)]\,\mathrm{d}t + \sqrt{dx_i(t)[1-x_i(t)]}\,\mathrm{d}w_i(t) \qquad (2.4)\\ &\quad + Ke\,[y_i(t)-x_i(t)]\,\mathrm{d}t,\\ \mathrm{d}y_i(t) &= e\,[x_i(t)-y_i(t)]\,\mathrm{d}t, \qquad i\in\mathbb{G}, \qquad (2.5)\end{aligned}$$

where $(w_i(t))_{t\geq 0}$, $i \in \mathbb{G}$, are independent standard Brownian motions. As initial state $Z(0) = z$ we may pick any $z \in E$. The first term in (2.4) describes the *migration* of active individuals at rate $a(i,j)$. The second term in (2.4) describes the *resampling* of individuals at rate $d \in (0,\infty)$. The third term in (2.4) together with the term in (2.5) describe the *exchange* of active and dormant individuals at rate $e \in (0,\infty)$.

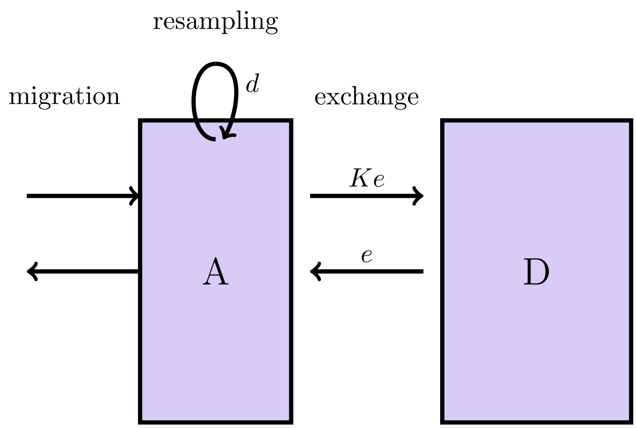

Figure 2.1: The evolution in model 1. Individuals are subject to migration, resampling and exchange with the seed-bank.

The factor $K \in (0,\infty)$ is defined by

$$K = \frac{\text{size dormant population}}{\text{size active population}}, \tag{2.6}$$

and is the same for all colonies $i \in \mathbb{G}$. The factor K turns up in the scaling limit of the individual-based model when there is an *asymmetry* between the sizes of the active and the dormant population (see Appendix A.1). In Fig. 2.1 we give a schematic illustration of the process (2.4)–(2.5). A detailed description of the underlying individual-based model, as well as a derivation of the continuum limit (2.4)–(2.5) from the individual-based model following [12], can be found in Appendix A.1. The continuum limit is also referred to as the frequency limit or the diffusion limit.

Remark 2.2.2 (Interpretation of the state space.). Note that the *state space* of the system can also be written as

$$E = [0,1]^{\mathbb{S}}, \qquad \mathbb{S} = \mathbb{G} \times \{A, D\}, \tag{2.7}$$

where A denotes the reservoir of the active population and D the repository of the dormant population. With that interpretation, the process is denoted by

$$(Z(t))_{t\geq 0}, \qquad Z(t) = \big(z_u(t)\big)_{u\in\mathbb{S}} \tag{2.8}$$

with $z_u(t) = x_i(t)$ if $u = (i, A)$ and $z_u(t) = y_i(t)$ if $u = (i, D)$. To analyse the system we need both interpretations of the state space. ■

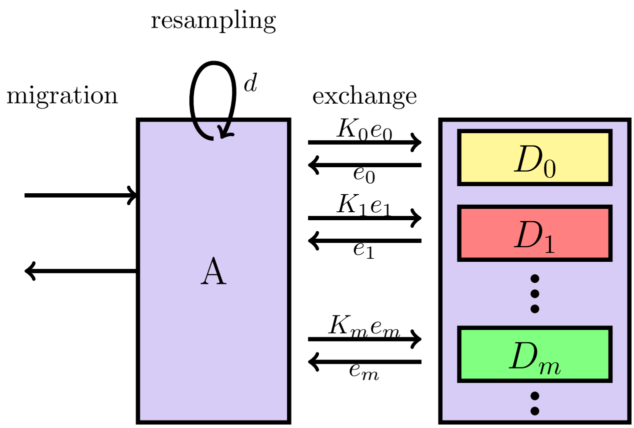

Figure 2.2: The evolution in model 2. Individuals are subject to migration, resampling and exchange with the seed-bank, as in model 1. Additionally, when individuals become dormant they get a colour and when they become active they loose their colour.

Model 2: multi-layer seed-bank. In this model we give the seed-bank an internal structure by colouring the dormant individuals with countably many colours $m \in \mathbb{N}_0$. Active individuals that become dormant are assigned a colour m that is drawn randomly from an infinite sequence of colours labeled by $\mathbb{N}_0$ (see Fig. 2.2 for an illustration). As will be explained in Section 2.2.2, this captures the *different ways in which individuals can enter into the seed-bank.* In Section 2.2.4 we will show how this internal structure allows for fat tails in the wake-up times of individuals while preserving the Markov property.

For each $i \in \mathbb{G}$ a colony now consists of an active part A and a whole sequence $(D_m)_{m\in\mathbb{N}_0}$ of dormant parts, labeled by their colour $m \in \mathbb{N}_0$. Therefore in this model the *effective geographic space* is given by $\mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}$.

As before, for $i \in \mathbb{G}$, let $x_i(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are active at time t, but now let $y_{i,m}(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are dormant with colour m at time t. Then the system is described by the process

$$(Z(t))_{t\geq 0}, \qquad Z(t) = \big(z_i(t)\big)_{i\in\mathbb{G}}, \qquad z_i(t) = (x_i(t), (y_{i,m}(t))_{m\in\mathbb{N}_0}), \tag{2.9}$$

on the state space

$$E = ([0,1]\times[0,1]^{\mathbb{N}_0})^{\mathbb{G}}. \tag{2.10}$$

Suppose that active individuals exchange with dormant individuals with colour m at rate $e_m \in (0,\infty)$, and let the factor $K_m \in (0,\infty)$ capture the asymmetry between the size of the active population and the m-dormant population, i.e., similarly as in (2.6),

$$K_m = \frac{\text{size } m\text{-dormant population}}{\text{size active population}}, \qquad m\in\mathbb{N}_0, \tag{2.11}$$

where $K_m \in (0,\infty)$ is the same for all colonies. Then the process $(Z(t))_{t\geq 0}$ evolves according to the SSDE

$$\begin{aligned}
\mathrm{d}x_i(t) &= \sum_{j\in\mathbb{G}} a(i,j)\,[x_j(t)-x_i(t)]\,\mathrm{d}t + \sqrt{dx_i(t)[1-x_i(t)]}\,\mathrm{d}w_i(t) \qquad (2.12)\\
&\quad + \sum_{m\in\mathbb{N}_0} K_m e_m\,[y_{i,m}(t)-x_i(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,m}(t) &= e_m\,[x_i(t)-y_{i,m}(t)]\,\mathrm{d}t, \qquad m\in\mathbb{N}_0,\ i\in\mathbb{G}, \qquad (2.13)
\end{aligned}$$

where we have to assume that

$$\sum_{m\in\mathbb{N}_0} K_m e_m < \infty, \tag{2.14}$$

since otherwise active individuals become dormant instantly. Comparing (2.12)–(2.13) with the SSDE of model 1 in (2.4)–(2.5), we see that active individuals migrate (the first term in (2.12)), resample (the second term in (2.12)), but now interact with a whole sequence of dormant populations (the third term in (2.12) and the term in (2.13)). As initial state $Z(0) = z$ we may again take any $z \in E$.

Remark 2.2.3 (Interpretation of the state space.). Note that, like in Remark 2.2.2, the *state space* of the system can also be written as

$$E = [0,1]^{\mathbb{S}}, \qquad \mathbb{S} = \mathbb{G}\times\{A, (D_m)_{m\in\mathbb{N}_0}\}. \tag{2.15}$$

With this interpretation, the process is denoted by

$$(Z(t))_{t\geq 0}, \qquad Z(t) = \big(z_u(t)\big)_{u\in\mathbb{S}}, \tag{2.16}$$

with $z_u(t) = x_i(t)$ if $u = (i,A)$ and $z_u(t) = y_{i,m}(t)$ if $u = (i, D_m)$ for $m \in \mathbb{N}_0$. ■

Model 3: multi-layer seed-bank with displaced seeds. We can extend the mechanism of model 2 by allowing active individuals that become dormant to do so in a randomly chosen colony. This amounts to introducing a sequence of irreducible *displacement kernels* $a_m(\cdot,\cdot)$, $m \in \mathbb{N}_0$, satisfying

$$a_m(i,j) = a_m(0,j-i) \quad \forall\, i,j \in \mathbb{G}, \qquad \sum_{i\in\mathbb{G}} a_m(0,i) = 1 \quad \forall\, m \in \mathbb{N}_0, \tag{2.17}$$

and replacing (2.12)–(2.13) by

$$\begin{aligned} \mathrm{d}x_i(t) &= \sum_{j\in\mathbb{G}} a(i,j)\,[x_j(t)-x_i(t)]\,\mathrm{d}t + \sqrt{dx_i(t)[1-x_i(t)]}\,\mathrm{d}w_i(t) \qquad (2.18)\\ &\quad + \sum_{j\in\mathbb{G}}\sum_{m\in\mathbb{N}_0} K_m e_m\, a_m(j,i)\,[y_{j,m}(t)-x_i(t)]\,\mathrm{d}t,\\ \mathrm{d}y_{i,m}(t) &= \sum_{j\in\mathbb{G}} e_m\, a_m(i,j)\,[x_j(t)-y_{i,m}(t)]\,\mathrm{d}t, \qquad m\in\mathbb{N}_0,\ i\in\mathbb{G}. \qquad (2.19) \end{aligned}$$

Here, the third term in (2.18) together with the term in (2.19) describe the *switch of colony* when individuals exchange between active and dormant. Namely, with probability $a_m(i,j)$ simultaneously an active individual in colony i becomes dormant with colour m in colony j and a randomly chosen dormant individual with colour m in colony j becomes active in colony i. The state space E is the same as in (2.10). Also (2.9), (2.11), (2.14) and (2.16) remain the same.

Two key quantities. In models 2 and 3 we must assume that

$$\chi = \sum_{m\in\mathbb{N}_0} K_m e_m < \infty \tag{2.20}$$

in order to make sure that active individuals do not become dormant instantly. Define

$$\rho = \sum_{m\in\mathbb{N}_0} K_m = \frac{\text{size dormant population}}{\text{size active population}}. \tag{2.21}$$

It turns out that ρ and χ are two key quantities of our system. In particular, we will see that the long-time behaviour of model 2 and model 3 is different for $\rho < \infty$ and $\rho = \infty$.

§2.2.2 Comments

(1) Models 1–3 are increasingly more general. Model 2 is the special case of model 3 when $a_m(0,0) = 1$ for all $m \in \mathbb{N}_0$, while model 1 is the special case of model 2 when $e_0 = e$, $K_0 = K$ and $e_m = K_m = 0$ for all $m \in \mathbb{N}$. Nonetheless, in what follows we prefer to state our main theorems for each model separately, in order to exhibit the increasing level of complexity. In Appendix A.1 we explain how (2.4)–(2.5), (2.12)–(2.13) and (2.18)–(2.19) arise as the large-colony-size limit of individual-based Fisher-Wright models.

(2) As geographic space $\mathbb{G}$ we allow *any* countable Abelian group endowed with the discrete topology. Key examples are the Euclidean lattice $\mathbb{G} = \mathbb{Z}^d$, $d \in \mathbb{N}$, and the hierarchical lattice $\mathbb{G} = \Omega_N$, $N \in \mathbb{N}$. In this paper we will focus $\mathbb{G} = \mathbb{Z}^d$. The case $\mathbb{G} = \Omega_N$ will be considered in more detail in [45].

(3) In model 1, each colony has a seed-bank that serves as a *repository* for the genetic information (type $\heartsuit$ or $\diamondsuit$) carried by the individuals. Because the active and the dormant population exchange individuals, the genetic information can be temporarily stored in the seed-bank and thereby be withdrawn from the resampling. We may think of dormant individuals as seeds that drop into the soil and preserve their type until they come to the surface again and grow into a plant.

In model 2, the seed-bank is a repository for seeds with one of infinitely many colours. *The colours provide us with a tool to model different distributions for the time an individual stays dormant without loosing the Markov property for the evolution of the system.* Tuning the parameters K_m and e_m properly and subsequently forgetting about the colours, we can mimic different distributions for the time an individual stays dormant. This is of biological significance, especially in colonies of bacteria, where individuals stay dormant for random times whose distribution is fat-tailed (see [55]).

In model 3, the seed may even be blown elsewhere. Individuals that displace before becoming dormant are observed in plant-species as well as in bacteria populations (see [55]).

(4) In Appendix A.2 we comment on what happens when the rates to become active or dormant are decoupled, i.e., individuals are no longer subject to exchange but move in and out of the seed-bank independently. This leads to a Moran model where the sizes of the active and the dormant population can fluctuate. We will show that, modulo a change of variables and a short transient period in which the sizes of the active and the dormant population establish equilibrium, this model has the same behaviour as the model with exchange.

(5) In (2.4), (2.12) and (2.18) we may replace the diffusion functions dg_{FW}, $d \in (0, \infty)$, where

$$g_{\mathrm{FW}}(x) = x(1-x), \qquad x \in [0,1], \tag{2.22}$$

is the Fisher-Wright diffusion function, by a *general diffusion function* in the class $\mathcal{G}$ defined by

$$\mathcal{G} = \Big\{ g\colon [0,1] \to [0,\infty)\colon\ g(0) = g(1) = 0,\ g(x) > 0\ \forall\, x \in (0,1),\ g \text{ Lipschitz} \Big\}. \tag{2.23}$$

This class is appropriate because a diffusion with a diffusion function $g \in \mathcal{G}$ stays confined to $[0,1]$, yet can go everywhere in [0,1] (Breiman [13, Chapter 16, Section 7]). Picking $g \neq g_{\mathrm{FW}}$ amounts to allowing the resampling rate to be *state-dependent*, i.e., the resampling rate in state x equals $g(x)/x(1-x)$, $x \in (0,1)$. An example is the Kimura-Ohta diffusion function $g(x) = [x(1-x)]^2$,

$x \in [0,1]$, for which the resampling rate is equal to the genetic diversity of the colony. *In the sequel we allow for general diffusion functions $g \in \mathcal{G}$ in all three models, unless stated otherwise.*

§2.2.3 Well-posedness

For every law on E, with E depending on the choice of model, we want the SSDE for models 1, 2 and 3 to define a Borel Markov process, i.e., the law of the path is a Borel measurable function of the initial state for every starting point in the state space [17, p.62]. We use a martingale problem, in the sense of [32, p.173], to characterize the SSDE. Let

$$\begin{aligned}\mathcal{F} = \Big\{ & f \in C_b(E,\mathbb{R})\colon\ f \text{ depends on finitely many components} \\ & \text{and is twice continuously differentiable in each component}\Big\}.\end{aligned} \tag{2.24}$$

The generator G of the process acting on $\mathcal{F}$ reads for model 1 ((2.4)–(2.5)),

$$\begin{aligned}G = \sum_{i\in\mathbb{G}} \Bigg(& \Bigg[\sum_{j\in\mathbb{G}} a(i,j)(x_j - x_i)\Bigg] \frac{\partial}{\partial x_i} + \frac{1}{2} g(x_i) \frac{\partial^2}{\partial x_i^2} \\ & + Ke\,(y_i - x_i)\frac{\partial}{\partial x_i} + e\,(x_i - y_i)\frac{\partial}{\partial y_i}\Bigg),\end{aligned} \tag{2.25}$$

for model 2 ((2.12)–(2.13)),

$$\begin{aligned}G = \sum_{i\in\mathbb{G}} \Bigg(& \Bigg[\sum_{j\in\mathbb{G}} a(i,j)(x_j - x_i)\Bigg] \frac{\partial}{\partial x_i} + \frac{1}{2} g(x_i) \frac{\partial^2}{\partial x_i^2} \\ & + \sum_{m\in\mathbb{N}_0} \Bigg[K_m e_m\,(y_{i,m} - x_i)\frac{\partial}{\partial x_i} + e_m\,(x_i - y_{i,m})\frac{\partial}{\partial y_{i,m}}\Bigg]\Bigg),\end{aligned} \tag{2.26}$$

while for model 3 ((2.18)–(2.19)) the last term in the right-hand side of (2.26) is to be replaced by

$$\sum_{i,j\in\mathbb{G}} \sum_{m\in\mathbb{N}_0} \Bigg[K_m e_m\, a_m(j,i)\,(y_{j,m} - x_i)\frac{\partial}{\partial x_i} + e_m\, a_m(i,j)\,(x_j - y_{i,m})\frac{\partial}{\partial y_{i,m}}\Bigg]. \tag{2.27}$$

Theorem 2.2.4 (Well-posedness: models 1–3). *For each of the three models the following holds:*
(a) The SSDE has a unique strong solution in $C([0,\infty),E)$. Its law is the unique solution of the $(G,\mathcal{F},\delta_u)$-martingale problem for all $u \in E$.
(b) The process starting in $u \in E$ is Feller and strong Markov. Consequently, the SSDE defines a unique Borel Markov process starting from any initial law on E.

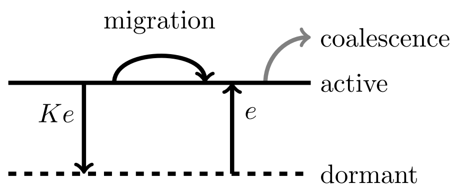

Figure 2.3: Transition scheme for an ancestral lineage in the dual, which moves according to the transition kernel $b(\cdot,\cdot)$ in (2.31)*. Two active ancestral lineages that are at the same colony coalesce at rate d.*

§2.2.4 Duality

For $g = dg_{\mathrm{FW}}$ the three models have a tractable dual, which will be seen to play a crucial role in the analysis of their long-time behaviour. For $g \neq dg_{\mathrm{FW}}$ the three models do not have a tractable dual. However, we compare them with models that do and determine their long-time behaviour. In [12, Sections 2.2 and 3] it was shown that the non-spatial Fisher-Wright diffusion with seed-bank is dual to the so-called *block-counting process* of a seed-bank coalescent. The latter describes the evolution of the number of partition elements in a partition of $n \in \mathbb{N}$ individuals, sampled from the current population, into subgroups of individuals with the same ancestor (i.e., individuals that are identical by descent). The enriched dual generates the ancestral lineages of the individuals evolving according to a Fisher-Wright diffusion with seed-bank, i.e., generates their full genealogy. The corresponding block-counting process counts the number of ancestral lineages left when traveling backwards in time. In this section we will extend the duality results in [12] to the spatial setting.

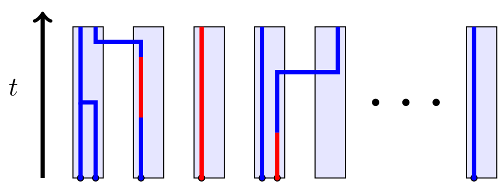

Figure 2.4: Picture of the evolution of lineages in the spatial coalescent. The purple blocks depict the colonies, the blue lines the active lineages, and the red lines the dormant lineages. Blue lineages can migrate and become dormant, (i.e., become red lineages). Two blue lineages can coalesce when they are at the same colony. Red dormant lineages first have to become active (blue) before they can coalesce with other blue and active lineages or migrate. Note that the dual runs backwards in time*. The collection of all lineages determines the genealogy of the system.*

Model 1. Recall that for model 1, $\mathbb{S} = \mathbb{G} \times \{A, D\}$ is the effective geographic space. For $n \in \mathbb{N}$ the state space of the n-spatial seed-bank coalescent is the set of partitions of $\{1, \ldots, n\}$, where the partition elements are marked with a position vector giving their location. A state is written as π, where

$$\begin{aligned}
&\pi = ((\pi_1,\eta_1),\dots,(\pi_{\bar n},\eta_{\bar n})), \quad \bar n = |\pi|,\\
&\pi_\ell \subset \{1,\dots,n\},\ \{\pi_1,\cdots\pi_{\bar n}\} \text{ is a partition of } \{1,\dots,n\},\\
&\eta_\ell \in \mathbb{S}, \quad \ell\in\{1,\dots,\bar n\}, \quad 1\le \bar n\le n.
\end{aligned} \tag{2.28}$$

A marked partition element (π_ℓ,η_ℓ) is called active if $\eta_\ell = (j,A)$ and called dormant if $\eta_\ell=(j,D)$ for some $j\in\mathbb{G}$. The n-spatial seed-bank coalescent is denoted by

$$(\mathcal{C}^{(n)}(t))_{t\geqslant 0}, \tag{2.29}$$

and starts from

$$\mathcal{C}^{(n)}(0)=\pi(0), \qquad \pi(0)=\{(\{1\},\eta_{\ell_1}),\dots,(\{n\},\eta_{\ell_n})\}, \qquad \eta_{\ell_1},\dots,\eta_{\ell_n}\in\mathbb{S}. \tag{2.30}$$

The n-spatial seed-bank coalescent is a Markov process that evolves according to the following two rules.

(a) Each partition element moves independently of all other partition elements according the kernel

$$b^{(1)}((i,R_i),(j,R_j)) = \begin{cases} a(i,j), & \text{if } R_i=R_j=A,\\ Ke, & \text{if } i=j,\ R_i=A,\ R_j=D,\\ e, & \text{if } i=j,\ R_i=D,\ R_j=A,\\ 0, & \text{otherwise,}\end{cases} \tag{2.31}$$

where $a(\cdot,\cdot)$ is the migration kernel defined in (2.1), K is the relative size of the dormant population defined in (2.6), and e is the rate of exchange between the active and the dormant population shown in (2.4)–(2.5). Therefore an active partition element migrates according to the transition kernel $a(\cdot,\cdot)$ and becomes dormant at rate Ke, while a dormant partition element can only become active and does so at rate e. In (2.31), the notation $b^{(1)}$ marks that the kernel refers to model 1. Later we will use the notation $b^{(2)}$ for model 2 and $b^{(3)}$ for model 3.

(b) Independently of all other partition elements, two partition elements that are at the same colony and are both active coalesce with rate d, i.e., the two partition elements merge into one partition element.

The spatial seed-bank coalescent $(\mathcal{C}(t))_{t\geq 0}$ is defined as the projective limit of the n-spatial seed-bank coalescents $(\mathcal{C}^{(n)}(t))_{t\geq 0}$ as $n\to\infty$. This object is well-defined by Kolmogorov's extension theorem (see [12, Section 3]).

For $n\in\mathbb{N}$ we define the block-counting process $(L(t))_{t\geq 0}$ corresponding to the n-spatial seed-bank coalescent as the process that counts at each site $(i,R_i)\in\mathbb{G}\times\{A,D\}$ the number of partition elements of $\mathcal{C}^{(n)}(t)$, i.e.,

$$\begin{aligned}
&L(t) = \big(L_{(i,A)}(t), L_{(i,D)}(t)\big)_{i\in\mathbb{G}},\\
&L_{(i,A)}(t) = L_{(i,A)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar n} 1_{\{\eta_\ell(t)=(i,A)\}},\\
&L_{(i,D)}(t) = L_{(i,D)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar n} 1_{\{\eta_\ell(t)=(i,D)\}}.
\end{aligned} \tag{2.32}$$

Therefore $(L(t))_{t\geq 0}$ has state space $E' = (\mathbb{N}_0 \times \mathbb{N}_0)^{\mathbb{G}}$. We denote the elements of E' by sequences $(m_i, n_i)_{i\in\mathbb{G}}$, and define $\delta_{(j,R_j)} \in E'$ to be the element of E' that is 0 at all sites $(i, R_i) \in \mathbb{G} \times \{A, D\}\backslash(j, R_j)$, but 1 at the site (j, R_j). From the evolution of $\mathcal{C}^{(n)}(t)$ described below (2.29) we see that the block-counting process has the following transition kernel:

$$(m_i, n_i)_{i\in\mathbb{G}} \to \begin{cases} (m_i, n_i)_{i\in\mathbb{G}} - \delta_{(j,A)} + \delta_{(k,A)}, & \text{at rate } m_j a(j,k) \text{ for } j,k \in \mathbb{G}, \\ (m_i, n_i)_{i\in\mathbb{G}} - \delta_{(j,A)}, & \text{at rate } d\binom{m_j}{2} \text{ for } j \in \mathbb{G}, \\ (m_i, n_i)_{i\in\mathbb{G}} - \delta_{(j,A)} + \delta_{(j,D)}, & \text{at rate } m_j Ke \text{ for } j \in \mathbb{G}, \\ (m_i, n_i)_{i\in\mathbb{G}} + \delta_{(j,A)} - \delta_{(j,D)}, & \text{at rate } n_j e \text{ for } j \in \mathbb{G}. \end{cases} \tag{2.33}$$

The process $(Z(t))_{t\geq 0}$ defined in (2.4)–(2.5) is dual to the block-counting process $(L(t))_{t\geq 0}$. The duality function $H\colon\, E \times E' \to \mathbb{R}$ is defined by

$$H\Big((x_i, y_i)_{i\in\mathbb{G}}, (m_i, n_i)_{i\in\mathbb{G}}\Big) = \prod_{i\in\mathbb{G}} x_i^{m_i} y_i^{n_i}. \tag{2.34}$$

The *duality relation* reads as follows.

Theorem 2.2.5 (Duality relation: model 1). *Let H be defined as in* (2.34)*. Then for all $(x_i, y_i)_{i\in\mathbb{G}} \in E$ and $(m_i, n_i)_{i\in\mathbb{G}} \in E'$,*

$$\begin{aligned} &\mathbb{E}_{(x_i,y_i)_{i\in\mathbb{G}}}\Big[H\Big((x_i(t), y_i(t))_{i\in\mathbb{G}}, (m_i, n_i)_{i\in\mathbb{G}}\Big)\Big] \\ &\qquad = \mathbb{E}_{(m_i,n_i)_{i\in\mathbb{G}}}\Big[H\Big((x_i, y_i)_{i\in\mathbb{G}}, (L_{(i,A)}(t), L_{(i,D)}(t))_{i\in\mathbb{G}}\Big)\Big] \end{aligned} \tag{2.35}$$

with $\mathbb{E}$ the generic symbol for expectation (on the left over the original process, on the right over the dual process).

Since the duality function H gives all the mixed moments of $(Z(t))_{t\geq 0}$, the duality relation in Theorem 2.2.5 is called a *moment dual.*

Remark 2.2.6 (Effective geographic space). Interpreting $(Z(t))_{t\geq 0}$ as a process on the effective geographic space $\mathbb{S}$, recall Remark 2.2.2, we can rewrite the duality relation. Let the block-counting process $(L(t))_{t\geq 0} = (L(\mathcal{C}(t))_{t\geq 0}$ count at each site $u \in \mathbb{S}$ the number of partition elements of $\mathcal{C}(t)$, i.e.,

$$\begin{aligned} L(t) &= (L_u(t))_{u\in\mathbb{S}}, \\ L_u(t) &= L_u(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=u\}}, \end{aligned} \tag{2.36}$$

and rewrite the duality function H in (2.34) as

$$H((z_u, l_u)_{u\in\mathbb{S}}) = \prod_{u\in\mathbb{S}} z_u^{l_u}. \tag{2.37}$$

Then, for $z \in \mathbb{E}$ and $l \in \mathbb{E}'$, the *duality relation* reads

$$\mathbb{E}\big[H(z_u(t), l_u)\big] = \mathbb{E}\big[H(z_u, L_u(t))\big]. \tag{2.38}$$

Interpreting the duality relation in terms of the effective geographic space $\mathbb{S}$, we see that each ancestral lineage in the dual is a Markov chain that moves according to the transition kernel $b^{(1)}(\cdot,\cdot)$. Interpreting the duality relation in terms of the geographic space $\mathbb{G}$, we see that an ancestral lineage is a random walk moving on $\mathbb{G}$, with internal states A and D. Both interpretations turn out to be useful in analysing the long-time behaviour of the system. ■

Remark 2.2.7 (Wake-up times). Define (see Fig. 2.3)

$$\begin{aligned} \sigma &= \text{typical time spent by an ancestral lineage in state } A \\ &\quad\ \text{before switching to state } D, \\ \tau &= \text{typical time spent by an ancestral lineage in state } D \\ &\quad\ \text{before switching to state } A. \end{aligned} \tag{2.39}$$

(Here, the word typical refers to what happens to an ancestral lineage each time it switches state at some geographic location. For a more precise definition we refer to Section 3.2.2 and Fig. 3.1.) It follows from (2.31) that

$$\begin{aligned} \mathbb{P}(\sigma > t) &= \mathrm{e}^{-Ket}, \\ \mathbb{P}(\tau > t) &= \mathrm{e}^{-et}. \end{aligned} \tag{2.40}$$

An ancestral lineage in the dual of the spatial seed-bank process behaves as an ancestral lineage in the dual of a spatial Fisher-Wright diffusion without seed-bank (see e.g. [36]), but becomes dormant every once in a while. On the long run we expect an ancestral lineage to be active only a fraction $\frac{1}{1+K}$ of the time. We will see in Section 3.2 that the effect of the seed-bank on the long-time behaviour of the ancestral lineages in the dual is a slow down by a factor $\frac{1}{1+K}$ compared to the long-time behaviour of the ancestral lineages in the dual of interacting Fisher-Wright diffusions without seed-bank. ■

Model 2. The dual for model 2 arises naturally from the dual for model 1 by adding internal states to the seed-bank and adapting the rates of becoming active and dormant accordingly. Recall that for model 2 the effective geographic space is $\mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}$. Migration and coalescence are as before, but at every colony switches between an active copy A and a dormant copy D_m now occur at rates e_m, respectively, $K_m\, e_m$. The spatial coalescent $(\mathcal{C}(t))_{t\geq 0}$ in (2.29) starts from an initial configuration like (2.30) and evolves according to the same two rules, but the transition kernel $b(\cdot,\cdot)$ must be replaced by

$$b^{(2)}((i,R_i),(j,R_j)) = \begin{cases} a(i,j), & R_i = R_j = A, \\ K_m e_m, & i = j,\ R_i = A,\ R_j = D_m, m \in \mathbb{N}_0, \\ e_m, & i = j,\ R_i = D_m,\ m \in \mathbb{N}_0,\ R_j = A, \\ 0, & \text{otherwise.} \end{cases} \tag{2.41}$$

The corresponding block-counting process becomes

$$\begin{aligned} L(t) &= \Big(L_{(i,A)}(t), \big(L_{(i,D_m)}(t)\big)_{m\in\mathbb{N}_0}\Big)_{i\in\mathbb{G}}, \\ L_{(i,A)}(t) &= L_{(i,A)}(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=(i,A)\}}, \\ L_{(i,D_m)}(t) &= L_{(i,D_m)}(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=(i,D_m)\}}, \; m\in\mathbb{N}_0. \end{aligned} \tag{2.42}$$

The state space is now given by $E' = \Big(\mathbb{N}_0 \times \mathbb{N}_0^{\mathbb{N}_0}\Big)^{\mathbb{G}}$, and the transition kernel becomes

$$\begin{aligned} &(m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} \\ &\quad\to \begin{cases} (m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} - \delta_{(j,A)} + \delta_{(k,A)}, & \text{at rate } m_j a(j,k) \text{ for } j,k\in\mathbb{G}, \\ (m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} - \delta_{(j,A)}, & \text{at rate } d\binom{m_j}{2} \text{ for } j\in\mathbb{G}, \\ (m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} - \delta_{(j,A)} + \delta_{(j,D_m)}, & \text{at rate } m_j K_m e_m \text{ for } j\in\mathbb{G}, \\ (m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} + \delta_{(j,A)} - \delta_{(j,D_m)}, & \text{at rate } n_{j,m} e_m \text{ for } j\in\mathbb{G}. \end{cases} \end{aligned} \tag{2.43}$$

The duality function $H\colon\ E\times E' \to \mathbb{R}$ is defined by

$$H\Big((x_i,y_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0},(m_i,n_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0}\Big) = \prod_{i\in\mathbb{G}}\prod_{m\in\mathbb{N}_0} x_i^{m_i} y_{i,m}^{n_{i,m}}. \tag{2.44}$$

Theorem 2.2.8 (Duality relation: model 2). *For $(x_i,y_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0}\in E$ and $(m_i,n_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0}\in E'$,*

$$\begin{aligned} &\mathbb{E}_{(x_i,y_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0}}\Big[H\Big((x_i(t),y_{i,m}(t))_{i\in\mathbb{G},m\in\mathbb{N}_0},(m_i,n_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0}\Big)\Big] \\ &\quad= \mathbb{E}_{(m_i,n_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0}}\Big[H\Big((x_i,y_{i,m})_{i\in\mathbb{G},m\in\mathbb{N}_0},(L_{(i,A)}(t),L_{(i,D_m)}(t))_{i\in\mathbb{G},m\in\mathbb{N}_0}\Big)\Big]. \end{aligned} \tag{2.45}$$

By rewriting the block-counting process as in Remark 2.2.6, the duality function can be rewritten as

$$H((z_u,l_u)_{u\in\mathbb{S}}) = \prod_{u\in\mathbb{S}} z_u^{l_u} \tag{2.46}$$

and the *duality relation* reads

$$\mathbb{E}\Big[H\Big((z_u(t))_{u\in\mathbb{S}},(l_u)_{u\in\mathbb{S}}\Big)\Big] = \mathbb{E}\Big[H\Big((z_u)_{u\in\mathbb{S}},(L_u(t))_{u\in\mathbb{S}}\Big)\Big]. \tag{2.47}$$

Remark 2.2.9 (Fat-tailed wake-up times.). Recall the definition of χ in (2.20) and the definition of ρ in (2.21). Define

$$\begin{aligned} \sigma = {} & \text{typical time spent by an ancestral lineage in the active state } A \\ & \quad \text{before switching to a dormant state } \cup_{m\in\mathbb{N}_0} D_m, \\ \tau = {} & \text{typical time spent by an ancestral lineage in the dormant state } \cup_{m\in\mathbb{N}_0} D_m \\ & \quad \text{before switching to the active state } A. \end{aligned} \tag{2.48}$$

Note that τ does not look at the colour of the dormant state. It follows from (2.41) that

$$\begin{aligned}\mathbb{P}(\sigma > t) &= \mathrm{e}^{-\chi t},\\ \mathbb{P}(\tau > t) &= \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{\chi} \mathrm{e}^{-e_m t},\end{aligned} \tag{2.49}$$

independently of the colony $i \in \mathbb{G}$. Hence

$$\mathbb{E}\left[\tau\right] = \frac{\rho}{\chi}. \tag{2.50}$$

If $\rho < \infty$, then we invoke the seed-bank colours and use the balance equations for recurrent Markov chains to see that each ancestral lineage in the dual in the long run spends a fraction $\frac{\rho}{1+\rho}$ of the time in the dormant state. Like in model 1, an ancestral lineage in the dual behaves like an ancestral lineage in the dual of interacting Fisher-Wright diffusions, but is slowed down by a factor $\frac{\rho}{1+\rho}$. However, if $\rho = \infty$, then (2.41) together with (2.50) imply that each ancestral lineage in the dual behaves like a null-recurrent Markov chain on $\{A, (D_m)_{m\in\mathbb{N}_0}\}$, and consequently the probability to be active tends to 0 as $t \to \infty$. Therefore we may expect that the long-time behaviour of the system is affected by the seed-bank. In particular, choosing

$$\begin{aligned}&K_m \sim A\, m^{-\alpha}, \quad e_m \sim B\, m^{-\beta}, \quad m \to \infty,\\ &A, B \in (0,\infty), \quad \alpha, \beta \in \mathbb{R}\colon\ \alpha \leq 1 < \alpha + \beta,\end{aligned} \tag{2.51}$$

we see that (2.49) implies

$$\mathbb{P}(\tau > t) \sim C t^{-\gamma}, \qquad t \to \infty, \tag{2.52}$$

with $\gamma = \frac{\alpha+\beta-1}{\beta}$ and $C = \frac{A}{\chi\beta} B^{1-\gamma}\,\Gamma(\gamma)$, where Γ is the Gamma-function. The conditions on α, β guarantee that $\rho = \infty$, $\chi < \infty$ (recall (2.20) and (2.21)). Examples are: $\alpha = 0$, $\beta > 1$ and $\alpha \in (0,1)$, $\beta > 1-\alpha$. Thus, for $\rho = \infty$ we can model individuals with a fat-tailed wake-up time simply by not taking their colours into account. *The internal structure of the seed-bank captured by the colours allows us to model fat-tailed wake-up times without loosing the Markov property for the evolution.* ■

Model 3. The effective geographic space is again $\mathbb{S} = \mathbb{G}\times\{A, (D_m)_{m\in\mathbb{N}_0}\}$. On top of migration and coalescence, each switch from A to D_m and vice versa is accompanied by a displacement according to the displacement kernel $a_m(\cdot,\cdot)$ defined in (2.17). Therefore each lineage in the dual evolves according to

$$b^{(3)}((i,R_i),(j,R_j)) = \begin{cases} a(i,j), & R_i = R_j = A,\\ K_m e_m a_m(j,i), & R_i = A,\ R_j = D_m, m \in \mathbb{N}_0,\\ e_m a_m(i,j), & R_i = D_m,\ m \in \mathbb{N}_0\, R_j = A.\end{cases} \tag{2.53}$$

Again, when two ancestral lineages are active at the same site they coalesce at rate 1 and the corresponding block-counting process evolves according to the transition kernel

$$
\begin{aligned}
&(m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}}\\
&\to \begin{cases}
(m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}}-\delta_{(j,A)}+\delta_{(k,A)}, & \text{at rate } m_j a(j,k) \text{ for } j,k\in\mathbb{G},\\
(m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}}-\delta_{(j,A)}, & \text{at rate } d\binom{m_j}{2} \text{ for } j\in\mathbb{G},\\
(m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}}-\delta_{(j,A)}+\delta_{(k,D_m)}, & \text{at rate } m_j K_m e_m a_m(k,j) \text{ for } j\in\mathbb{G},\\
(m_i,(n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}}+\delta_{(k,A)}-\delta_{(j,D_m)}, & \text{at rate } n_{j,m} e_m a_m(j,k) \text{ for } j\in\mathbb{G}.
\end{cases}
\end{aligned}
\tag{2.54}
$$

Theorem 2.2.10 (Duality relation: model 3). *The same duality relation holds as in* (2.45)*, where now the dual dynamics includes not only the exchange between active and dormant but also the accompanying displacement in space.*

§2.2.5 Dichotomy criterion

For $g = dg_{\mathrm{FW}}$ the duality relations in Theorems 2.2.5, 2.2.8 and 2.2.10 provide us with the following criterion to characterise the long-term behaviour. If, in the limit as $t \to \infty$, locally only one type survives in the population, then we say that the system exhibits *clustering*. If, in the limit as $t \to \infty$, locally both types survive in the population, then we say that the system exhibits *coexistence*. For model 1 the criterion reads as follows.

Theorem 2.2.11 (Dichotomy criterion: model 1). *Suppose that $\mu(0)$ is invariant and ergodic under translations. Let $d \in (0,\infty)$. Then the system with $g = dg_{\mathrm{FW}}$ clusters if and only if in the dual two partition elements coalesce with probability* 1*.*

The idea behind Theorem 2.2.11 is as follows. If in the dual two partition elements coalesce with probability 1, then a random sample of n individuals drawn from the current population has a common ancestor some finite time backwards in time. Since individuals inherit their type from their parent individuals, this means that all n individuals have the same type. A formal proof will be given in Section 3.1.3.

For model 2–3 we need an extra assumption on $\mu(0)$ when $\rho = \infty$.

Definition 2.2.12 (Colour regular initial measures). We say that $\mu(0)$ is *colour regular* when

$$
\lim_{N\to\infty} \mathbb{E}_{\mu(0)}[y_{0,N}] \quad \text{exists,} \tag{2.55}
$$

i.e., $\mu(0)$ has asymptotically converging colour means. □

Thus, colour regularity is a condition on the deep seed-banks (where deep means $m \to \infty$). This condition is needed because as time proceeds lineages starting from deeper and deeper seed-banks become active for the first time, and bring new types into the active population. Without control on the initial states of the deep seed-banks, there may be no convergence to equilibrium.

Theorem 2.2.13 (Dichotomy criterion: models 2–3). *The same as in Theorem 2.2.11 is true for $\rho < \infty$, but for $\rho = \infty$ additionally requires that $\mu(0)$ is colour regular.*

Remark 2.2.14 (Clustering criterion general $g \in \mathcal{G}$). In Section 2.3 we will see that the dichotomy criterion in Theorems 2.2.11 and 2.2.13 for $g = dg_{\mathrm{FW}}$ does not depend on d, the rate of resampling. We will use *duality comparison arguments* to carry over the dichotomy criterion in Theorems 2.2.11 and 2.2.13 to $g \in \mathcal{G}$. We will see later that for all three models the system with g exhibits clustering if and only if the system with g_{FW} exhibits clustering. ■

Remark 2.2.15 (Liggett conditions). We will see in Section 3.3.3 that, for model 2 with $\rho = \infty$, if an initial measure μ is invariant and ergodic under translations and is colour regular, then the Markov chain evolving according to $b^{(2)}(\cdot,\cdot)$ satisfies the following two conditions:

(1)

$$\lim_{t\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\,\mathbb{E}_\mu[z_{(k,R_k)}] = \theta, \tag{2.56}$$

(2)

$$\begin{aligned}\lim_{t\to\infty} \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}} & b_t^{(2)}\big((i,R_i),(k,R_k)\big)\, b_t^{(2)}\big((j,R_j),(l,R_l)\big) \\ & \times \mathbb{E}_\mu[z_{(k,R_k)} z_{(l,R_l)}] = \theta^2.\end{aligned} \tag{2.57}$$

These are precisely the conditions in [56, Chapter V.1] necessary to determine the dichotomy in the long-time behaviour of the voter model. We show that (1) and (2) imply convergence to a unique equilibrium that is invariant and ergodic under translations. It is difficult to identify exactly which initial measures μ satisfy (1) and (2). This is the reason why we work with sufficient conditions and need the notion of colour regularity.

For model 2 with $\rho < \infty$, conditions (1) and (2) are satisfied when $\mu(0)$ is invariant and ergodic under translations, and colour regularity is not needed. The same holds for model 1, once the state space is replaced by $\mathbb{G} \times \{A, D\}$ and $b^{(2)}(\cdot,\cdot)$ is replaced by $b^{(1)}(\cdot,\cdot)$. Also for model 3 conditions (1) and (2) hold after replacing $b^{(2)}(\cdot,\cdot)$ by $b^{(3)}(\cdot,\cdot)$. If $\rho = \infty$ in model 3 we need to assume colour regularity, if $\rho < \infty$, this is not needed. ■

§2.3 Long-time behaviour

In this section we study the long-time behaviour of models 1–3. In Sections 2.3.1–2.3.3 we prove convergence to a unique equilibrium measure, establish the dichotomy between clustering and coexistence, and identify which of the two occurs in terms of the migration kernel and the rates governing the exchange with the seed-bank (Theorems 2.3.1–2.3.6).

Throughout the sequel, g is a general diffusion function from the class $\mathcal{G}$ defined in (2.23). Special cases are the multiples of the standard Fisher-Wright diffusion function: $g = dg_{\mathrm{FW}}$, $d \in (0,\infty)$, with $g_{\mathrm{FW}}(x) = x(1-x)$, $x \in [0,1]$. We use the following notation (with $\mathcal{P}(E)$ denotes the set of probability measures on E):

$$\begin{aligned} \mathcal{T} &= \big\{\mu \in \mathcal{P}(E)\colon \ \mu \text{ is invariant under translations in } \mathbb{G}\big\},\\ \mathcal{T}^{\mathrm{erg}} &= \big\{\mu \in \mathcal{T}\colon \ \mu \text{ is ergodic under translations in } \mathbb{G}\big\},\\ \mathcal{I} &= \big\{\mu \in \mathcal{T}\colon \ \mu \text{ is invariant under the evolution}\big\}. \end{aligned} \tag{2.58}$$

§2.3.1 Long-time behaviour of Model 1

Let $a(\cdot,\cdot)$ be as in (2.1). Define the *symmetrized* migration kernel

$$\hat{a}(i,j) = \tfrac{1}{2}[a(i,j) + a(j,i)], \qquad i,j \in \mathbb{G}, \tag{2.59}$$

which describes the difference of two independent copies of the migration each driven by $a(\cdot,\cdot)$. Let $\hat{a}_t(0,0)$ denote the time-t transition kernel of the random walk with migration kernel $\hat{a}(\cdot,\cdot)$, and suppose that

$$t \mapsto \hat{a}_t(0,0) \text{ is regularly varying at infinity.} \tag{2.60}$$

(Examples can be found in [47, Chapter 3].) Define

$$I_{\hat{a}} = \int_1^\infty \mathrm{d}t\, \hat{a}_t(0,0). \tag{2.61}$$

Note that $I_{\hat{a}} = \infty$ if and only if $\hat{a}(\cdot,\cdot)$ is recurrent (see e.g. [69, Chapter 1]). Define

$$\theta = \mathbb{E}_{\mu(0)}\left[\frac{x_0 + Ky_0}{1+K}\right]. \tag{2.62}$$

If $\mu(0)$ is invariant and ergodic under translations, then θ is the initial density of $\heartsuit$ in the population.

From the SSDE in (2.4)–(2.5) we see that

$$\left(\frac{x_0(t) + Ky_0(t)}{1+K}\right)_{t\geq 0} \tag{2.63}$$

is a martingale. In particular,

$$\theta = \mathbb{E}_{\mu(t)}\left[\frac{x_0 + Ky_0}{1+K}\right] \qquad \forall\, t \geq 0. \tag{2.64}$$

For $\theta \in [0,1]$, we define

$$\mathcal{T}^{\mathrm{erg}}_\theta = \left\{\mu \in \mathcal{T}^{\mathrm{erg}}\colon\ \mathbb{E}_{\mu(0)}\left[\frac{x_0 + Ky_0}{1+K}\right] = \theta\right\}. \tag{2.65}$$

Write $\mu(t)$ to denote the law of $(Z(t))_{t\geq 0}$, defined in (2.2). Recall that associated means that increasing functions of the configuration are positively correlated, i.e., if $f\colon E \to \mathbb{R}$ and $g\colon E \to \mathbb{R}$ depend on only finitely many coordinates and are coordinate-wise increasing, then

$$\mathbb{E}_{\nu_\theta}[f(x)g(x)] \geq \mathbb{E}_{\nu_\theta}[f(x)]\,\mathbb{E}_{\nu_\theta}[g(x)]. \tag{2.66}$$

Theorem 2.3.1 (Long-time behaviour: model 1). *Suppose that* $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$.

(a) *(Coexistence regime) If* $\hat{a}(\cdot,\cdot)$ *is transient, i.e.,* $I_{\hat{a}} < \infty$, *then*

$$\lim_{t\to\infty} \mu(t) = \nu_\theta, \tag{2.67}$$

where

$$\nu_\theta \text{ is an equilibrium measure for the process on } E, \tag{2.68}$$
$$\nu_\theta \text{ is invariant, ergodic and mixing under translations}, \tag{2.69}$$
$$\nu_\theta \text{ is associated}, \tag{2.70}$$
$$\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_0] = \theta, \tag{2.71}$$

with $\mathbb{E}_{\nu_\theta}$ *denoting expectation over* ν_θ.

(b) *(Clustering regime) If* $\hat{a}(\cdot,\cdot)$ *is recurrent, i.e.,* $I_{\hat{a}} = \infty$, *then*

$$\lim_{t\to\infty} \mu(t) = \theta\,[\delta_{(1,1)}]^{\otimes\mathbb{G}} + (1-\theta)\,[\delta_{(0,0)}]^{\otimes\mathbb{G}}. \tag{2.72}$$

The results in (2.67)–(2.72) say that the system converges to an equilibrium whose density of type ♡ equals θ in (2.62), a parameter that is controlled by the initial state $\mu(0)$ and the asymmetry parameter K. The equilibrium can be either locally mono-type or locally multi-type, depending on whether the symmetrised migration kernel is recurrent or transient. If the equilibrium is mono-type, then the system grows large mono-type clusters (= clustering). If the equilibrium is multi-type, then the system allows ♡ and ♢ to mix (= coexistence). In the case of coexistence, the equilibrium measure ν_θ also depends on the migration kernel $a(\cdot,\cdot)$, the values of the parameters e, K, and the diffusion function $g \in \mathcal{G}$ (recall (2.23)). The dichotomy itself, however, is *controlled by* $I_{\hat{a}}$ *only*. In particular, $g \in \mathcal{G}$ plays no role, a fact that will be shown with the help of a *duality comparison* argument. In view of Theorem 2.2.11, if $g = dg_{\mathrm{FW}}$, then $I_{\hat{a}} = \infty$ implies that with probability 1 two ancestral lineages in the dual coalesce. Therefore $I_{\hat{a}} = \infty$ is said to be *the total hazard of coalescence*. Remarkably, this dichotomy is the same as the dichotomy observed for systems without seed-bank (see [14]): clustering prevails for recurrent migration; coexistence prevails for transient migration; for $\mathbb{G} = \mathbb{Z}^d$ the critical dimension is $d = 2$. From the proof in Section 3.2.2 it will become clear that in the dual the ancestral lineages in the long run behave like the ancestral lineages without seed-bank, but are slowed down by a factor $\frac{1}{1+K}$. Consequently, the dormant periods of the ancestral lineages do not affect the dichotomy of the system. In particular, it does not affect the critical dimension separating clustering from coexistence.

Remark 2.3.2 (Ergodic decomposition). Because $\mathcal{T}$ is a Choquet simplex, Theorem 2.3.1 carries over from $\mu(0) \in \mathcal{T}^{\mathrm{erg}}$ to $\mu(0) \in \mathcal{T}$, after decomposition into ergodic components. ■

§2.3.2 Long-time behaviour of Model 2

For model 2 we need the extra condition that $a(\cdot,\cdot)$ is *symmetric*, i.e.,

$$a(i,j) = a(j,i) \qquad \forall\, i,j \in \mathbb{G}. \tag{2.73}$$

Note that $\hat{a}_t(0,0) = a_t(0,0)$ because of (2.73). Below we comment on what happens when we drop this assumption. Recall (2.20)–(2.21). It turns out that the long-time behaviour of model 2 is different for $\rho < \infty$ and $\rho = \infty$.

Case $\rho < \infty$. For a finite seed-bank, we define the initial density as

$$\theta = \mathbb{E}_{\mu(0)}\left[\frac{x_0 + \sum_{m\in\mathbb{N}_0} K_m\, y_{0,m}}{1+\rho}\right], \tag{2.74}$$

which is the counter part of (2.62) in model 1. Like in model 1, it follows from the SSDE in (2.12)–(2.13) that

$$\left(\frac{x_0(t) + \sum_{m\in\mathbb{N}_0} K_m\, y_{0,m}(t)}{1+\rho}\right)_{t\geq 0} \tag{2.75}$$

is a martingale. Hence also here the density is a preserved quantity under the evolution of the system. The dichotomy is controlled by the same integral $I_{\hat{a}}$ as defined in (2.61) for model 1.

Case $\rho = \infty$. For an infinite seed-bank, we assume that (recall Remark 2.2.9)

$$\begin{aligned} &K_m \sim A\, m^{-\alpha}, \quad e_m \sim B\, m^{-\beta}, \quad m \to \infty,\\ &A, B \in (0,\infty), \quad \alpha,\beta \in \mathbb{R}\colon\ \alpha \leq 1 < \alpha+\beta, \end{aligned} \tag{2.76}$$

for which

$$P(\tau > t) \sim C\, t^{-\gamma}, \quad t \to \infty, \tag{2.77}$$

with $\gamma = \frac{\alpha+\beta-1}{\beta} \in (0,1)$ and $C = \frac{A}{\beta}\, B^{1-\gamma}\, \gamma\Gamma(\gamma) \in (0,\infty)$, where Γ is the Gamma-function. In addition, we assume that the initial measure $\mu(0)$ is colour regular (recall Definition 2.2.12), and define

$$\theta = \lim_{m\to\infty} \mathbb{E}[y_{0,m}]. \tag{2.78}$$

This ensures the existence of the initial density

$$\theta = \lim_{M\to\infty} \mathbb{E}_{\mu(0)}\left[\frac{x_0 + \sum_{m=0}^{M} K_m\, y_{0,m}}{1+\sum_{m=0}^{M} K_m}\right]. \tag{2.79}$$

It turns out that the dichotomy is controlled by the integral

$$I_{\hat{a},\gamma} = \int_1^\infty \mathrm{d}t\ t^{-(1-\gamma)/\gamma}\, \hat{a}_t(0,0) \tag{2.80}$$

instead of the integral $I_{\hat{a}}$ for $\rho < \infty$.

For $\theta \in (0,1)$, define (both for $\rho < \infty$ and $\rho = \infty$)

$$\mathcal{T}_\theta^{\mathrm{erg}} = \left\{ \mu \in \mathcal{T}^{\mathrm{erg}} \colon \lim_{M\to\infty} \mathbb{E}_{\mu(0)} \left[\frac{x_0 + \sum_{m=0}^M K_m y_{0,m}}{1 + \sum_{m=0}^M K_m} \right] = \theta \right\}. \tag{2.81}$$

Theorem 2.3.3 (Long-time behaviour: model 2). (I) *Let $\rho < \infty$. Assume* (2.60) *and* (2.73). *Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$.*

(a) *(Coexistence regime) If $I_{\hat{a}} < \infty$, then*

$$\lim_{t\to\infty} \mu(t) = \nu_\theta, \tag{2.82}$$

where

$$\nu_\theta \text{ is an equilibrium measure for the process on } E, \tag{2.83}$$
$$\nu_\theta \text{ is invariant, ergodic and mixing under translations}, \tag{2.84}$$
$$\nu_\theta \text{ is associated}, \tag{2.85}$$
$$\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta \;\forall\, m \in \mathbb{N}_0, \tag{2.86}$$

with $\mathbb{E}_{\nu_\theta}$ denoting expectation over ν_θ. Moreover,

$$\begin{aligned} \liminf_{m\to\infty} e_m > 0 \colon \; \liminf_{m\to\infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) > 0, \\ \limsup_{m\to\infty} e_m = 0 \colon \; \limsup_{m\to\infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) = 0. \end{aligned} \tag{2.87}$$

(b) *(Clustering regime) If $I_{\hat{a}} = \infty$, then*

$$\lim_{t\to\infty} \mu(t) = \theta\, [\delta_{(1,1^{\mathbb{N}_0})}]^{\otimes\mathbb{G}} + (1-\theta)\, [\delta_{(0,0^{\mathbb{N}_0})}]^{\otimes\mathbb{G}}. \tag{2.88}$$

(II) *Let $\rho = \infty$. Assume* (2.60), (2.73) *and* (2.76). *Suppose that $\mu(0) \in \mathcal{T}^{\mathrm{erg}}$ and, in addition, is colour regular with initial density θ given by* (2.79). *Then the same results as in* (I) *hold after $I_{\hat{a}}$ in* (2.61) *is replaced by $I_{\hat{a},\gamma}$ in* (2.80). *Moreover,*

$$\lim_{M\to\infty} \mathbb{E}_{\nu_\theta} \left[\frac{x_0 + \sum_{m=0}^M K_m\, y_{0,m}}{1 + \sum_{m=0}^M K_m} \right] = \theta, \tag{2.89}$$

and ν_θ is colour regular.

The result in part (I) shows that for $\rho < \infty$ the long-time behaviour is similar to that of model 1. Like in model 1, the results in (2.82)–(2.88) say that the system converges to an equilibrium whose density of type ♡ equals θ in (2.62), the density of ♡ under the initial measure $\mu(0)$. Again, the equilibrium can be either mono-type or multi-type, depending on whether the symmetrised migration kernel is recurrent or transient. Like in model 1, in both cases the equilibrium measure depends on θ. In the case of coexistence, the equilibrium measure ν_θ also depends on the migration kernel $a(\cdot,\cdot)$, the sequences of parameters $(e_m)_{m\in\mathbb{N}_0}$ and $(K_m)_{m\in\mathbb{N}_0}$, and the diffusion function $g \in \mathcal{G}$ (recall (2.23)). Again, the dichotomy itself is *controlled by $I_{\hat{a}}$ only*, and the resampling rate given by $g \in \mathcal{G}$ plays no role. Therefore if $g = dg_{FW}$, in view of Theorem 2.2.11, whether or not two ancestral lineages in the dual coalesce with probability 1 is still only determined by the migration kernel $a(\cdot,\cdot)$. The same dichotomy holds as for systems without seed-bank (see [14]). Therefore part (I) of Theorem 2.3.3 indicates that, as long as the dormant periods of the ancestral lineages in the dual have a finite mean ($\frac{\rho}{1+\rho}$; recall Remark 2.2.9), the seed-bank does not affect the dichotomy of the system.

Even so, (2.87) indicates that there is interesting behaviour in the deep seed-banks. Indeed, when the exchange rate e_m between the m-dormant and the active population is bounded away from zero as $m \to \infty$ the deep seed-banks are asymptotically *random*, while when e_m tend to zero as $m \to \infty$ the deep seed-banks are asymptotically *deterministic*. The latter means that the deep seed-banks serve as a reservoir, containing a fixed mixture of types. For $\rho < \infty$ this reservoir is too small to influence the dichotomy of the system, but not for $\rho = \infty$.

For $\rho = \infty$ the system again converges to an equilibrium whose density of type $\heartsuit$ equals θ in (2.79), the density of $\heartsuit$ under the initial measure $\mu(0)$. The equilibrium can be mono-type or multi-type, but *the dichotomy criterion has changed.* Instead of $I_{\hat{a}}$, the dichotomy is now controlled by the integral $I_{\hat{a},\gamma}$ (recall (2.80)), where γ is the parameter determined by relative sizes K_m of the colour m-dormant populations with respect to the active population and the exchanges rates $(e_m)_{m\in\mathbb{N}_0}$ with the seed-bank, recall (2.76)–(2.77). If $g = dg_{FW}$, γ is the parameter of the tail of the wake-up time of an ancestral lineages in the dual (recall (2.2.9)). Therefore if $g = dg_{Fw}$, in view of Theorem 2.2.13, we see that the dormant periods of the ancestral lineages in the dual do affect whether or not two ancestral lineages in the dual coalesce with probability 1. For general $g \in \mathcal{G}$, the integral $I_{\hat{a},\gamma}$ in (2.80) shows a *competition* between migration and exchange. The smaller γ is, the longer the individuals remain dormant in the seed-bank, the smaller $I_{\hat{a},\gamma}$ is, and the more coexistence becomes likely. As a consequence clustering requires more stringent conditions than recurrent migration; for $\mathbb{G} = \mathbb{Z}^d$ the critical dimension is $1 < d < 2$ for $\gamma \in [\frac{1}{2}, 1]$ and $d = 1$ for $\gamma \in (0, \frac{1}{2})$. The *seed-bank enhances genetic diversity.* Note that $\gamma \uparrow 1$ links up with the case $\rho < \infty$, where coexistence occurs if and only if the migration is transient. Also note that for $\gamma \in (0, \frac{1}{2})$ there is always coexistence *irrespective of the migration.*

In the case of clustering the equilibrium measure only depends on θ, while in the case of coexistence, like for $\rho < \infty$, ν_θ depends on the migration kernel $a(\cdot,\cdot)$, the sequences of parameters $(e_m)_{m\in\mathbb{N}_0}$, $(K_m)_{m\in\mathbb{N}_0}$, and the diffusion function $g \in \mathcal{G}$. Since we assumed (2.76), we have $\limsup_{m\to\infty} e_m = 0$, and so we are automatically in the second case of (2.87). Hence the deep seed-banks are asymptotically deterministic, i.e., the m-dormant population converges in law to a deterministic state θ as $m \to \infty$. Roughly speaking, in case $g = dg_{FW}$, in equilibrium the volatility of a colour is inversely proportional to its average wake-up time in the dual. Since $\rho = \infty$, for each $M \in \mathbb{N}_0$ we have $\sum_{m=M}^{\infty} K_m = \infty$, and in the coexistence regime the effect of the seed-bank can be interpreted as a migration towards an infinite reservoir with deterministic density θ.

Like for model 1, also here $\mathcal{T}$ is a Choquet simplex, and Theorem 2.3.3 carries over from $\mathcal{T}^{\mathrm{erg}}$ to $\mathcal{T}$, after decomposition into ergodic components.

Example of effect of infinite seed-bank. For a symmetric migration kernel with finite second moment the following holds:

- For $\mathbb{G} = \mathbb{Z}^2$, $\hat{a}_t(0,0) \asymp t^{-1}$, $t \to \infty$, and so coexistence occurs for all $\gamma \in (0,1)$.
- For $\mathbb{G} = \mathbb{Z}$, $\hat{a}_t(0,0) \asymp t^{-1/2}$, $t \to \infty$, and so coexistence occurs if and only if $\gamma \in (0, \frac{2}{3})$.

In both cases the migration is recurrent, so that clustering prevails in model 1.

Corollary 2.3.4 (Three regimes). *Under the conditions of Theorem 2.3.3, the system in* (2.12)–(2.13) *has three different parameter regimes:*

(1) $\gamma \in (1,\infty)$*: migration determines the dichotomy.*

(2) $\gamma \in [\frac{1}{2}, 1]$*: interplay between migration and seed-bank determines the dichotomy.*

(3) $\gamma \in (0, \frac{1}{2})$*: seed-bank determines the dichotomy.*

Role of symmetry in migration. Unlike in model 1, it is *not* possible to remove the symmetry assumption in (2.73), as the following counterexample shows. We consider model 2 with $\rho < \infty$ under assumption (2.60), but we do not assume (2.73).

- **Counterexample:** Let $\mathbb{G} = \mathbb{Z}^2$, and for $\eta \in (0,1)$ pick

$$a(i,j) = \begin{cases} \frac{1}{4}(1+\eta), & j = i + (1,0) \text{ or } i + (0,1), \\ \frac{1}{4}(1-\eta), & j = i - (1,0) \text{ or } i - (0,1), \end{cases} \tag{2.90}$$

  i.e., two-dimensional nearest-neighbour random walk with drift upward and rightward. Suppose that τ in (2.77) has a *one-sided stable distribution* with parameter $\gamma \in (1,2)$ (obtained from (2.76) but with $\alpha, \beta \in \mathbb{R}$: $1 < \alpha < 1 + \beta$). Then coexistence occurs while $I_{\hat{a}} = \infty$.

Recall that for the two-dimensional nearest-neighbour random walk without drift we get clustering according to Theorem 2.3.3, independently of the distribution of τ. The key feature of the counterexample is that it corresponds to $\mathbb{E}(\tau) < \infty$ and $\mathbb{E}(\tau^2) = \infty$. Hence the central limit theorem fails for τ. We will see in Section 3.3.5 that the failure of the central limit theorem for τ is responsible for turning clustering into coexistence.

The above raises the question to what extent the equilibrium behaviour depends on the nature of the geographic space. To answer this question, we need a key concept for random walks on countable Abelian groups, which we describe next.

Remark 2.3.5 (Dichotomy criterion and degrees of random walk). We can read the condition $I_{\hat{a},\gamma} < \infty$ for coexistence versus $I_{a,\gamma} = \infty$ for clustering in terms of the *degree* of the random walk. Namely, let $\hat{a}(\cdot,\cdot)$ be the transition kernel of an irreducible random walk on a countable Abelian group. Then the degree δ of $\hat{a}(\cdot,\cdot)$ is defined as

$$\delta = \sup\left\{\zeta > -1\colon\ \int_1^\infty \mathrm{d}t\, t^\zeta\, \hat{a}_t(0,0) < \infty\right\}. \tag{2.91}$$

The degree is defined to be δ^+ when the integral is finite *at* the degree and δ^- when the integral is infinite *at* the degree. Hence we can rephrase the dichotomy criterion in Theorem 2.3.3 as

$$\text{clustering} \quad\Longleftrightarrow\quad \text{either } -\frac{1-\gamma}{\gamma} \geq \delta^- \text{ or } -\frac{1-\gamma}{\gamma} > \delta^+. \tag{2.92}$$

For further details we refer to [18], [19], which relate the degree of the random walk to the tail of its return time to the origin. ■

Modulation of wake-up time with slowly varying function. Under weak conditions it is possible to modulate (2.77) by a slowly varying function. Assume that

$$\frac{\mathbb{P}(\tau \in \mathrm{d}t)}{\mathrm{d}t} \sim \varphi(t)\, t^{-(1+\gamma)}, \quad t \to \infty, \tag{2.93}$$

with φ *slowly varying at infinity.* Define

$$\hat{\varphi}(t) = \begin{cases} \varphi(t), & \gamma \in (0,1), \\ \int_1^t \mathrm{d}s\, \varphi(s) s^{-1}, & \gamma = 1. \end{cases} \tag{2.94}$$

As shown in [8, Section 1.3], without loss of generality we may take $\hat{\varphi}$ to be infinitely differentiable and to be represented by the integral

$$\hat{\varphi}(t) = \exp\left[\int_{(\cdot)}^t \frac{\mathrm{d}u}{u}\, \psi(u)\right] \tag{2.95}$$

for some $\psi\colon [0,\infty) \to \mathbb{R}$ such that $\lim_{u\to\infty} |\psi(u)| = 0$. If we assume that ψ eventually has a sign and satisfies $|\psi(u)| \leq C/\log u$ for some $C < \infty$, then (2.80) needs to be replaced by

$$I_{\hat{a},\gamma,\varphi} = \int_1^\infty \mathrm{d}t\, \hat{\varphi}(t)^{-1/\gamma}\, t^{-(1-\gamma)/\gamma}\, \hat{a}_t(0,0). \tag{2.96}$$

A proof is given in Section 3.3.6. The modulation of the wake-up time by a slowly varying function appears naturally for the model on the hierarchical group, analysed in [45]. There the integral criterion for the dichotomy in (2.96) is needed to apply Theorem 2.3.3.

§2.3.3 Long-time behaviour of Model 3

It remains to see how the switch of colony during the exchange affects the dichotomy. We will focus on the special case where the displacement kernels do not depend on m, i.e.,

$$a_m(\cdot,\cdot) = a^\dagger(\cdot,\cdot) \qquad \forall\, m \in \mathbb{N}_0, \tag{2.97}$$

with $a^\dagger(\cdot,\cdot)$ an irreducible *symmetric* random walk kernel on $\mathbb{G}\times\mathbb{G}$. Let $\hat{a}_t^\dagger(\cdot,\cdot)$ denote the time-t transition kernel of the random walk with symmetrised displacement kernel $\hat{a}^\dagger(\cdot,\cdot)$ $(= a^\dagger(\cdot,\cdot))$ and jump rate 1. Assume that (compare with (2.60))

$$\begin{aligned} &t \mapsto (\hat{a}_t * \hat{a}_t^\dagger)(0,0) \text{ is regularly varying at infinity,} \\ &(\hat{a}_{Ct} * \hat{a}_t^\dagger)(0,0) \asymp (\hat{a}_t * \hat{a}_t^\dagger)(0,0) \text{ as } t\to\infty \text{ for every } C\in(0,\infty), \end{aligned} \tag{2.98}$$

where $*$ stands for convolution. Let

$$I_{\hat{a}*\hat{a}^\dagger} = \int_1^\infty \mathrm{d}t\ (\hat{a}_t * \hat{a}_t^\dagger)(0,0) \tag{2.99}$$

and

$$I_{\hat{a}*\hat{a}^\dagger,\gamma} = \int_1^\infty \mathrm{d}t\ t^{-(1-\gamma)/\gamma}\,(\hat{a}_t * \hat{a}_t^\dagger)(0,0). \tag{2.100}$$

Theorem 2.3.6 (Long-time behaviour: model 3). *Suppose that, in addition to the assumptions of Theorem 2.3.3, both* (2.97) *and* (2.98) *hold. Then the same results as for model* 2 *hold: (I) for $\rho<\infty$ after $I_{\hat{a}}$ in* (2.61) *is replaced by $I_{\hat{a}*\hat{a}^\dagger}$ in* (2.99)*; (II) for $\rho=\infty$ after $I_{\hat{a},\gamma}$ in* (2.80) *is replaced by $I_{\hat{a}*\hat{a}^\dagger,\gamma}$ in* (2.100)*.*

In the case of coexistence the equilibrium measure ν_θ depends on $a(\cdot,\cdot)$, $a^\dagger(\cdot,\cdot)$, $(e_m)_{m\in\mathbb{N}_0}$, $(K_m)_{m\in\mathbb{N}_0}$ and $g\in\mathcal{G}$. The dichotomy itself, however, is controlled by $I_{\hat{a}*\hat{a}^\dagger}$, respectively, $I_{\hat{a}*\hat{a}^\dagger,\gamma}$ alone.

An interesting observation is the following. Since $\hat{a}_t(\cdot,\cdot)$ and $\hat{a}_t^\dagger(\cdot,\cdot)$ are symmetric, we have (by a standard Fourier argument)

$$\hat{a}_t(i,j) \le \hat{a}_t(0,0), \qquad \hat{a}_t^\dagger(i,j) \le \hat{a}_t^\dagger(0,0) \qquad \forall\, i,j\in\mathbb{G}\ \forall\, t\ge 0. \tag{2.101}$$

Hence, $I_{\hat{a}*\hat{a}^\dagger,\gamma} \le I_{\hat{a},\gamma} \wedge I_{\hat{a}^\dagger,\gamma}$. Consequently, the extra displacement in model 3 can only make coexistence more likely compared to model 2, which is intuitively plausible.

If $a(\cdot,\cdot)=a^\dagger(\cdot,\cdot)$, then $(a_t * a_t^\dagger)(0,0) = a_{2t}(0,0)$ and therefore the dichotomy is the same as for model 2. Hence the extra displacement has in this case no effect on the dichotomy. However, if the displacement is transient while the migration is recurrent, then there is a difference. For instance, if $\rho<\infty$, the migration is a simple random walk on $\mathbb{Z}$, and the displacement is a symmetric random walk on $\mathbb{Z}$ with infinite mean, e.g. $a^\dagger(0,x)=a^\dagger(0,-x)\sim D|x|^{-\delta}$, $D\in(0,\infty)$, $\delta\in(1,2)$, then $I_{\hat{a}}=\infty$, $I_{\hat{a}^\dagger}<\infty$ and $I_{\hat{a}*\hat{a}^\dagger}<\infty$ [69, Section 8]. Therefore there is clustering in model 2, but coexistence in model 3.

CHAPTER 3

Spatial populations with seed-bank, proofs

§3.1 Proofs: Well-posedness and duality

In Section 3.1.1 we prove Theorem 2.2.4, in Section 3.1.2 Theorems 2.2.5, 2.2.8 and 2.2.10, and in Section 3.1.3 Theorems 2.2.11 and 2.2.13.

§3.1.1 Well-posedness

In this section we prove Theorem 2.2.4.

Proof. (a) We first prove Theorem 2.2.4(a): existence and uniqueness of solutions to the SSDE. We do this for each of the three models separately.

Model 1. Existence of the process defined in (2.4)–(2.5) for model 1 is a consequence of the assumptions in (2.1), (2.17) and (2.20), in combination with [67, Theorem 3.2], which reads as follows:

Theorem 3.1.1 (Unique strong solution). *Let $\mathbb{S}$ be a countable set, and let $Z = \{z_u\}_{u\in\mathbb{S}} \in [0,1]^{\mathbb{S}}$. Consider the stochastic differential equation*

$$\mathrm{d}z_u(t) = \alpha_u(z_u(t))\,\mathrm{d}B_u(t) + f_u(Z(t))\,\mathrm{d}t, \qquad u\in\mathbb{S}, \tag{3.1}$$

where $\alpha_u\colon [0,1]\to\mathbb{R}$ for all $u\in\mathbb{S}$, $f_u\colon [0,1]^{\mathbb{S}}\to[0,1]$ for all $u\in\mathbb{S}$, and $B=\{B_u\}_{u\in\mathbb{S}}$ is a collection of independent standard Brownian motions. Suppose that:

(1) *The functions α_u, $u\in\mathbb{S}$, are real-valued, $\frac{1}{2}$-Hölder continuous (i.e., there are $C_u\in(0,\infty)$ such that $|\alpha_u(x)-\alpha_u(y)|\leq C_u|x-y|^{\frac{1}{2}}$ for all $x,y\in[0,1]$) and uniformly bounded, with $\alpha_u(0)=\alpha_u(1)=0$, $u\in\mathbb{S}$.*

(2) *The functions f_u, $u\in\mathbb{S}$, are continuous and satisfy:*

- *There exists a matrix $Q=\{Q_{u,v}\}_{u,v\in\mathbb{S}}$ such that $Q_{u,v}\geq 0$ for all $u,v\in\mathbb{S}$, $\sup_{u\in\mathbb{S}}\sum_{v\in\mathbb{S}} Q_{u,v}<\infty$, and*

$$\begin{aligned}&|f_u(Z^1)-f_u(Z^2)|\leq \sum_{v\in\mathbb{S}} Q_{u,v}|z^1_v-z^2_v|,\\ &\textit{for}\quad Z^1=\{z^1_v\}_{v\in\mathbb{S}}\in[0,1]^{\mathbb{S}},\ Z^2=\{z^2_v\}_{v\in\mathbb{S}}\in[0,1]^{\mathbb{S}}.\end{aligned} \tag{3.2}$$

- *For $Z \in [0,1]^{\mathbb{S}}$ and $z_u = 0$,*

$$f_u(Z) \geq 0. \tag{3.3}$$

- *For $Z \in [0,1]^{\mathbb{S}}$ and $z_u = 1$,*

$$f_u(Z) \leq 0. \tag{3.4}$$

Then (3.1) *has a unique* $[0,1]^{\mathbb{S}}$*-valued strong solution with a continuous path.*

To apply Theorem 3.1.1 to model 1, recall that

$$\mathbb{S} = \mathbb{G} \times \{A, D\}, \tag{3.5}$$

where A denotes the active part of a colony and D the dormant part of a colony. Since $\mathbb{G}$ is countable and $\{A, D\}$ is finite, $\mathbb{S}$ is countable. As before, we denote the fraction of active individuals of type $\heartsuit$ at colony $i \in \mathbb{G}$ by x_i and the fraction of dormant individuals of type $\heartsuit$ at colony $i \in \mathbb{G}$ by y_i. Note that for every $u \in \mathbb{S}$ we have either $u = (i, A)$ or $u = (i, D)$ for some $i \in \mathbb{G}$. Therefore

$$Z = \{z_u\}_{u \in \mathbb{S}} = \{x_i \colon i \in \mathbb{G}\} \cup \{y_i \colon i \in \mathbb{G}\}, \tag{3.6}$$

and $z_u = x_i$ when $u = (i, A)$ and $z_u = y_i$ when $u = (i, D)$. We can rewrite (2.4)–(2.5) in the form of (3.1) by picking

$$\alpha_u(z_u) = \begin{cases} \sqrt{g(x_i)}, & u = (i, A), \\ 0, & u = (i, D), \end{cases} \tag{3.7}$$

and

$$f_u(Z) = \begin{cases} \sum_{j \in \mathbb{G}} a(i,j)\,(x_j - x_i) + Ke\,(y_i - x_i), & u = (i, A), \\ e\,(x_i - y_i), & u = (i, D). \end{cases} \tag{3.8}$$

Since $g \in \mathcal{G}$ (recall (2.23)), the conditions in (1) are satisfied. To check the conditions in (2), define the matrix $Q = \{Q_{u,v}\}_{u,v \in \mathbb{S}}$ by

$$Q_{u,v} = \begin{cases} \sum_{j \in \mathbb{G}} a(i,j) + Ke, & u = (i, A),\ v = (i, A), \\ a(i,j), & u = (i, A),\ v = (j, A), \\ Ke, & u = (i, A),\ v = (i, D), \\ e, & u = (i, D),\ v = (i, D) \text{ or } u = (i, D),\ v = (i, A), \\ 0, & \text{otherwise}. \end{cases} \tag{3.9}$$

Then

$$\sum_{v \in S} Q_{u,v} = \begin{cases} 2\sum_{j \in \mathbb{G}} a(i,j) + 2Ke, & u = (i, A), \\ 2e, & u = (i, D). \end{cases} \tag{3.10}$$

Since we have assumed that $\sum_{j \in \mathbb{G}} a(i,j) = \sum_{j \in \mathbb{G}} a(0, j-i) < \infty$, it follows that $\sup_{u \in \mathbb{S}} \sum_{v \in \mathbb{S}} Q_{u,v} < \infty$. Since $x_i \in [0,1]$ and $y_i \in [0,1]$, the requirements on f_u are immediate. Hence we have a unique strong solution with a continuous path.

By Itô's formula, the law of the strong solution solves the martingale problem. Uniqueness of that solution follows from [62, Theorem IX 1.7(i)]. This in turn implies the Markov property.

Model 2. To apply Theorem 3.1.1 to model 2, recall that

$$\mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}. \tag{3.11}$$

Pick

$$\alpha_u(z_u) = \begin{cases} \sqrt{g(x_i)}, & u = (i, A), \\ 0, & u = (i, D_m),\ m \in \mathbb{N}_0, \end{cases} \tag{3.12}$$

and

$$f_u(Z) = \begin{cases} \sum_{j\in\mathbb{G}} a(i,j)\,(x_j - x_i) + \sum_{m\in\mathbb{N}_0} K_m e_m\,(y_{i,m} - x_i), & u = (i, A), \\ e_m\,(x_i - y_{i,m}), & u = (i, D_m). \end{cases} \tag{3.13}$$

Set

$$Q_{u,v} = \begin{cases} \sum_{j\in\mathbb{G}} a(i,j) + \sum_{m\in\mathbb{N}_0} K_m e_m, & u = (i, A),\ v = (i, A), \\ a(i,j), & u = (i, A),\ v = (j, A),\ j \neq i, \\ K_m e_m, & u = (i, A),\ v = (i, D_m), \\ e_m, & u = (i, D_m),\ v = (i, D_m) \\ & \text{ or } u = (i, D_m),\ v = (i, A), \\ 0, & \text{otherwise.} \end{cases} \tag{3.14}$$

Then, by assumptions (2.1) and (2.20), Q, f and α satisfy the conditions of Theorem 3.1.1.

Model 3. The state space $\mathbb{S}$ and the function α are the same as in model 2. When $u \in \mathbb{S}$ is of the form (i, A), we must adapt the function f_u such that it takes the displacement of seeds into account. The matrix Q must be adapted accordingly and, by assumption (2.17), the conditions of Theorem 3.1.1 are again satisfied.

(b) The proof of Theorem 2.2.4(b) is the same for models 1–3. The Feller property can be proved by using duality if $g = dg_{\mathrm{FW}}$, $d \in (0,\infty)$. For general g we use [67, Remark 3.2] (see also [56, Theorem 5.8]). The Feller property in turn implies the strong Markov property. □

§3.1.2 Duality

In this section we prove Theorems 2.2.5, 2.2.8 and 2.2.10.

Model 1: Proof of Theorem 2.2.5.

Proof. We use the generator criterion (see [32, p.190–193] or [48, Proposition 1.2]) to prove the duality relation given in (2.35). Let F be the generator of the spatial block-counting process defined in (2.33), and let $H((m_j, n_j)_{j\in\mathbb{G}})$ be defined as in (2.34), but read as a function of the second sequence only. Then

$$\begin{aligned}
&(FH)\big((m_j,n_j)_{j\in\mathbb{G}}\big)\\
&=\sum_{i\in\mathbb{G}}\Bigg[\sum_{k\in\mathbb{G}} m_i a(i,k)\,\big[H\big((m_j,n_j)_{j\in\mathbb{G}}-\delta_{(i,A)}+\delta_{(k,A)}\big)-H\big((m_j,n_j)_{j\in\mathbb{G}}\big)\big]\\
&\qquad+d\binom{m_i}{2}\big[H\big((m_j,n_j)_{j\in\mathbb{G}}-\delta_{(i,A)}\big)-H\big((m_j,n_j)_{j\in\mathbb{G}}\big)\big]\\
&\qquad+m_iKe\big[H\big((m_j,n_j)_{j\in\mathbb{G}}-\delta_{(i,A)}+\delta_{(i,D)}\big)-H\big((m_j,n_j)_{j\in\mathbb{G}}\big)\big]\\
&\qquad+n_ie\big[H\big((m_j,n_j)_{j\in\mathbb{G}}+\delta_{(i,A)}-\delta_{(i,D)}\big)-H\big((m_j,n_j)_{j\in\mathbb{G}}\big)\big]\Bigg].
\end{aligned}\tag{3.15}$$

Recall that G is the generator of the SSDE (recall (2.24)–(2.25)). Let $\mathcal{D}_G$ denote the domain of G and $\mathcal{D}_F$ the domain of F. Let $(S_t)_{t\geq 0}$ denote the semigroup of the process $(Z(t))_{t\geq 0}$ in (2.2) and $(R_t)_{t\geq 0}$ the semigroup of the process $(L(t))_{t\geq 0}$ in (2.32). Since

$$\frac{d^2}{dt^2}(R_tH)((x_j,y_j,n_j,m_j)_{j\in\mathbb{G}})=(F^2R_tH)((x_j,y_j,n_j,m_j)_{j\in\mathbb{G}}),\tag{3.16}$$

we see that $H((x_j,y_j,n_j,m_j)_{j\in\mathbb{G}})\in\mathcal{D}_G$ and $(R_tH)((x_j,y_j,n_j,m_j)_{j\in\mathbb{G}})\in\mathcal{D}_G$. It is also immediate that $H((x_j,y_j,n_j,m_j)_{j\in\mathbb{G}})\in\mathcal{D}_F$ and
$(S_tH)((x_j,y_j,n_j,m_j)_{j\in\mathbb{G}})\in\mathcal{D}_F$. Applying the generator G in (2.25) with $g=\frac{d}{2}g_{\mathrm{FW}}$

to (2.34), we find

$$\begin{aligned}
&(GH)\big((x_j,y_j)_{j\in\mathbb{G}}\big)\\
&=\sum_{i\in\mathbb{G}}\Bigg\{\Bigg[\sum_{k\in\mathbb{G}}a(i,k)\,(x_k-x_i)\Bigg]\frac{\partial}{\partial x_i}\Big(\prod_{j\in\mathbb{G}}x_j^{m_j}y_j^{n_j}\Big)\\
&\qquad+\tfrac{d}{2}\,x_i(1-x_i)\frac{\partial^2}{\partial x_i^2}\Big(\prod_{j\in\mathbb{G}}x_j^{m_j}y_j^{n_j}\Big)+Ke\,(y_i-x_i)\frac{\partial}{\partial x_i}\Big(\prod_{j\in\mathbb{G}}x_j^{m_j}y_j^{n_j}\Big)\\
&\qquad\qquad+e\,(x_i-y_i)\frac{\partial}{\partial y_i}\Big(\prod_{j\in\mathbb{G}}x_j^{m_j}y_j^{n_j}\Big)\Bigg\}\\
&=\sum_{i\in\mathbb{G}}\Bigg\{\Bigg[\sum_{k\in\mathbb{G}}m_i a(i,k)\prod_{\substack{j\in\mathbb{G}\\ j\neq i\\ j\neq k}}x_j^{m_j}y_j^{n_j}\,\big(x_i^{m_i-1}y_i^{n_i}x_k^{m_k+1}y_k^{n_k}-x_i^{m_i}y_i^{n_i}x_k^{m_l}y_k^{n_k}\big)\Bigg]\\
&\qquad+\prod_{\substack{j\in\mathbb{G}\\ j\neq i}}x_j^{m_j}y_j^{n_j}\,\tfrac{d}{2}\,m_i(m_i-1)\,\big(x_i^{m_i-1}y_i^{n_i}-x_i^{m_i}y_i^{n_i}\big)\,1_{\{m_i\geq 2\}}\\
&\qquad\qquad+m_iKe\prod_{\substack{j\in\mathbb{G}\\ j\neq i}}x_j^{m_j}y_j^{n_j}\,\big(x_i^{m_i-1}y_i^{n_i+1}-x_i^{m_i}y_i^{n_i}\big)\\
&\qquad\qquad+n_ie\prod_{\substack{j\in\mathbb{G}\\ j\neq i}}x_j^{m_j}y_j^{n_j}\,\big(x_i^{m_i+1}y_i^{n_i-1}-x_i^{m_i}y_i^{n_i}\big)\Bigg\}\\
&=(FH)\big((m_j,n_j)_{j\in\mathbb{G}}\big).
\end{aligned}\tag{3.17}$$

Consequently, it follows from the generator criterion that

$$\mathbb{E}\Big[H\Big((X_i(t),Y_i(t),m_i,n_i)_{i\in\mathbb{G}}\Big)\Big]=\mathbb{E}\Big[H\Big((x_i,y_i,M_i(t),N_i(t))_{i\in\mathbb{G}}\Big)\Big].\tag{3.18}$$

This settles Theorem 2.2.5. □

Model 2: Proof of Theorem 2.2.8.

Proof. Theorem 2.2.8 follows after replacing in the above proof the block-counting process in (2.33) by the one in (2.43), the duality function by the one in (2.44), and checking the generator criterion. □

Model 3: Proof of Theorem 2.2.10.

Proof. Theorem 2.2.10 follows after replacing the block-counting process in (2.33) by the one in (2.54), the duality function is by the one in (2.44), and checking the generator criterion. □

§3.1.3 Dichotomy criterion

In this section we prove Theorems 2.2.11 and 2.2.13.

Model 1: Proof of Theorem 2.2.11.

Proof.

"$\Longleftarrow$" The proof uses the duality relation in Theorem 2.2.5. Define $\theta_x = \mathbb{E}_{\mu(0)}[x_0]$ and $\theta_y = \mathbb{E}_{\mu(0)}[y_0]$. Note that, since $\mu(0)$ is invariant under translations, we have $\mathbb{E}_{\mu(0)}[x_i] = \theta_x$ and $\mathbb{E}_{\mu(0)}[y_i] = \theta_y$ for all $i \in \mathbb{G}$. We proceed as in [12, Proposition 2.9]. Let $(m_i, n_i)_{i\in\mathbb{G}} \in E'$ be such that $\sum_{i\in\mathbb{G}}[m_i(0) + n_i(0)] < \infty$, and put

$$T = \inf\left\{ t \geq 0\colon \sum_{i\in\mathbb{G}}[m_i(t) + n_i(t)] = 1 \right\}. \tag{3.19}$$

By assumption, each pair of partition elements coalesces with probability 1, and hence $\mathbb{P}(T < \infty) = 1$. By duality

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}\left[\prod_{i\in\mathbb{G}} x_i(t)^{m_i} y_i(t)^{n_i}\right] \\
&= \lim_{t\to\infty} \mathbb{E}\left[\prod_{i\in\mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)}\right] \\
&= \lim_{t\to\infty} \mathbb{E}\left[\prod_{i\in\mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \;\middle|\; T < \infty\right] \mathbb{P}(T < \infty) \\
&\qquad + \mathbb{E}\left[\prod_{i\in\mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \;\middle|\; T = \infty\right] \mathbb{P}(T = \infty) \\
&= \lim_{t\to\infty} \mathbb{E}\left[\prod_{i\in\mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \;\middle|\; T < \infty,\ m(t) = 1,\ n(t) = 0\right] \mathbb{P}(m(t) = 1,\ n(t) = 0) \\
&\qquad + \lim_{t\to\infty} \mathbb{E}\left[\prod_{i\in\mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \;\middle|\; T < \infty,\ m(t) = 0,\ n(t) = 1\right] \mathbb{P}(m(t) = 0,\ n(t) = 1) \\
&= \theta_x \frac{1}{1+K} + \theta_y \frac{K}{1+K},
\end{aligned} \tag{3.20}$$

where in the last step we use that a single lineage in the dual behaves like the Markov chain with transition kernel $b^{(1)}(\cdot,\cdot)$ defined in (2.31). It follows from (3.20) that, for all $i, j \in \mathbb{G}$,

$$\lim_{t\to\infty} \mathbb{E}\left[\frac{x_i(t) + K y_i(t)}{1+K}\left(1 - \frac{x_j(t) + K y_j(t)}{1+K}\right)\right] = 0. \tag{3.21}$$

Hence, either $\lim_{t\to\infty}(x(t), y(t)) = (0,0)^{\mathbb{G}}$ or $\lim_{t\to\infty}(x(t), y(t)) = (1,1)^{\mathbb{G}}$. Computing $\lim_{t\to\infty} \mathbb{E}[x_i(t)]$ with the help of (3.20), we find

$$\lim_{t\to\infty} \mu(t) = (1-\theta)\,[\delta_{(0,0)}]^{\bigotimes \mathbb{G}} + \theta\,[\delta_{(1,1)}]^{\bigotimes \mathbb{G}} \tag{3.22}$$

with $\theta = \mathbb{E}_{\mu(0)}\left[\frac{x_0 + K y_0}{1+K}\right] = \frac{\theta_x + K\theta_y}{1+K}$, which means that the system clusters.

"$\Longrightarrow$" Suppose that the systems clusters. Then (3.21) holds for all $i, j \in \mathbb{G}$, which means that

$$\lim_{t\to\infty} \mathbb{E}\left[z_u(t)\,(1 - z_v(t))\right] = 0 \qquad \forall\, u, v \in \mathbb{S}. \tag{3.23}$$

Let

$$|L(t)| = \sum_{u\in\mathbb{S}} L_u(t), \tag{3.24}$$

be the total number of lineages left at time t. Applying the duality relation in (2.38) to (3.23), we find

$$\begin{aligned}
0 &= \lim_{t\to\infty} \mathbb{E}\left[z_u(t)(1 - z_v(t))\right] \\
&= \lim_{t\to\infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\right]\right] - \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\right]\right] \\
&= \lim_{t\to\infty} \left[\frac{\theta_x + K\theta_y}{1+K}\left[1 - \mathbb{P}_{\delta_u+\delta_v}\left(|L(t)| = 1\right)\right]\right. \\
&\qquad \left. - \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)} \,\middle|\, |L(t)| = 2\right]\right] \mathbb{P}_{\delta_u+\delta_v}(|L(t)| = 2)\right].
\end{aligned} \tag{3.25}$$

As to the last term in the right-hand side of (3.25), we note that

$$\begin{aligned}
&\limsup_{t\to\infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)} \,\middle|\, |L(t)| = 2\right]\right] \\
&= \limsup_{t\to\infty} \frac{1}{(1+K)^2}\mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)} \,\middle|\, L(t) = \delta_{(i,A)} + \delta_{(j,A)},\, i, j \in \mathbb{G}\right] \\
&\quad + \limsup_{t\to\infty} \frac{2K}{(1+K)^2}\mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)} \,\middle|\, L(t) = \delta_{(i,A)} + \delta_{(j,D)},\, i, j \in \mathbb{G}\right] \\
&\quad + \limsup_{t\to\infty} \frac{K^2}{(1+K)^2}\mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)} \,\middle|\, L(t) = \delta_{(i,D)} + \delta_{(j,D)}, i, j \in \mathbb{G}\right] \\
&< \frac{\theta_x}{(1+K)^2} + \frac{K\theta_x + K\theta_y}{(1+K)^2} + \frac{K^2\theta_y}{(1+K)^2} = \frac{\theta_x + K\theta_y}{1+K} = \theta.
\end{aligned} \tag{3.26}$$

Here, the strict inequality follows from the non-trivial invariant initial distribution (ruling out $z \equiv 0$ and $z \equiv 1$), together with the fact that the swapping between active and dormant is driven by a positive recurrent Markov chain on $\{A, D\}$. Hence (3.23) holds if and only if $\lim_{t\to\infty} \mathbb{P}_{\delta_u+\delta_v}(|L(t)| = 2 |\, |L(0)| = 2) = 0$ for every $u, v \in \mathbb{S}$. Therefore every pair of lineages coalesces with probability 1. Thus, we have proved Theorem 2.2.11.

Model 2: Proof of Theorem 2.2.13.

Case $\rho < \infty$. Like for model 1, we define

$$\theta_x = \mathbb{E}_{\mu(0)}[x_0], \qquad \theta_{y,m} = \mathbb{E}_{\mu(0)}[y_{0,m}], \qquad \theta = \frac{\theta_x + \sum_{m=0}^{\infty} K_m \theta_{y,m}}{1+\rho}. \tag{3.27}$$

For $\rho < \infty$, a lineage in the dual moves as a positive recurrent Markov chain on $\{A, (D_m)_{m\in\mathbb{N}_0}\}$. Therefore the argument for "$\Longleftarrow$" given for model 1 goes through via the duality relation, which gives

$$\lim_{t\to\infty} \mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u(t)^{l_u}\right] = \lim_{t\to\infty} \mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\right] = \frac{\theta_x + \sum_{m\in\mathbb{N}_0} K_m\theta_{y,m}}{1+\sum_{m\in\mathbb{N}_0} K_m}. \tag{3.28}$$

With the duality relation in (2.47), the argument for "$\Longrightarrow$" given for model 1 also goes through directly.

Case $\rho = \infty$. For $\rho = \infty$, a lineage in the dual moves as a null-recurrent Markov chain, which has no stationary distribution, and so (3.28) does not carry over. However, from [58, Section 3] it follows that, for all $u_1, u_2 \in \mathbb{S}$,

$$\lim_{t\to\infty} \left\|\mathbb{P}_{u_1}(L(t) = \delta_{(\cdot)} \mid L(t) = 1) - \mathbb{P}_{u_2}(L(t) = \delta_{(\cdot)} \mid L(t) = 1)\right\|_{tv} = 0. \tag{3.29}$$

Moreover, by null-recurrence,

$$\begin{aligned} \lim_{t\to\infty} \mathbb{P}(L(t) = \delta_{(\cdot,A)}) &= 0, \\ \lim_{t\to\infty} \mathbb{P}(L(t) = \delta_{(\cdot,D_m)}) &= 0 \qquad \forall m \in \mathbb{N}_0, \\ \lim_{t\to\infty} \sum_{m=M}^{\infty} \mathbb{P}(L(t) = \delta_{(\cdot,D_m)}) &= 1 \qquad \forall M \in \mathbb{N}_0. \end{aligned} \tag{3.30}$$

"$\Longleftarrow$" By duality, we have

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u(t)^{l_u}\right] &= \lim_{t\to\infty} \mathbb{E}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\right] \\ &= \lim_{t\to\infty}\left[\theta_x\mathbb{P}(L(t) = \delta_{(\cdot,A)}) + \sum_{m\in\mathbb{N}_0} \theta_{y,m}\mathbb{P}(L(t) = \delta_{(\cdot,D_m)})\right], \end{aligned} \tag{3.31}$$

where we follow an argument similar as in (3.20) and use that $\mathbb{P}(T < \infty) = 1$. Because the initial measure is colour regular, we know that $\lim_{m\to\infty} \theta_{y,m} = \theta$ (recall Definition 2.2.12). But (3.30)–(3.31) imply that all moments tend to θ. In particular,

$$\lim_{t\to\infty} \mathbb{E}[x_i(t)] = \theta = \lim_{t\to\infty} \mathbb{E}[y_{i,m}(t)], \qquad i \in \mathbb{G},\ m \in \mathbb{N}_0. \tag{3.32}$$

"$\Longrightarrow$" By the duality relation in (2.47) and the assumption of clustering, we find

$$\lim_{t\to\infty} \mathbb{E}\left[z_u(t)(1-z_v(t))\right] = 0 \qquad \forall u, v \in \mathbb{S}. \tag{3.33}$$

Therefore

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}\left[z_u(t)(1-z_v(t))\right] \\
&= \lim_{t\to\infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\right]\right] - \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\right]\right] \\
&= \theta \lim_{t\to\infty}\Bigg[\left[1-\mathbb{P}_{\delta_u+\delta_v}\left(|L(t)|=1\right)\right] \\
&\quad - \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\;\middle|\; |L(t)|=2\right]\right]\mathbb{P}_{\delta_u+\delta_v}(|L(t)|=2)\Bigg] = 0.
\end{aligned} \tag{3.34}$$

Suppose that $\lim_{t\to\infty} \mathbb{P}_{\delta_u+\delta_v}(|L(t)|=2) \neq 0$. Then

$$\lim_{t\to\infty} \mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)} \mid |L(t)|=2\right] = \theta. \tag{3.35}$$

However,

$$\limsup_{t\to\infty} \quad \mathbb{E}_{\mu(0)} \quad \left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\;\middle|\; |L(t)|=2\right]\right] \tag{3.36}$$

$$< \quad \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u+\delta_v}\left[\prod_{u\in\mathbb{S}} z_u^{L_u(t)}\;\middle|\; |L(t)|=1\right]\right] = \theta, \tag{3.37}$$

because we start from a nontrivial stationary distribution. Thus, we have proved Theorem 2.2.13.

Model 3: Proof of Theorem 2.2.13. Since the duality relation for model 3 is exactly the same as for model 2, the same results hold by translation invariance and the extra displacement does not affect the dichotomy criterion. □

§3.1.4 Outline remainder of paper

In Sections 3.2–3.4 we prove Theorems 2.3.1, 2.3.3 and 2.3.6, respectively. For each of the three models we split the proof into four parts:

(a) Moment relations.

(b) The clustering case.

(c) The coexistence case.

(d) Proof of the dichotomy.

§3.2 Proofs: Long-time behaviour for Model 1

In Section 3.2.1 we relate the first and second moments of the process $(Z(t))_{t\geq 0}$ in (2.4)–(2.5) to the random walk with internal states $\{A, D\}$ that evolves according to the transition kernel $b^{(1)}(\cdot,\cdot)$ given in (2.31) (Lemma 3.2.1 below). These moment relations hold for all $g \in \mathcal{G}$. In Section 3.2.2 we deal with the clustering case (Lemmas 3.2.4–3.2.5 below), in Section 3.2.3 with the coexistence case (Lemmas 3.2.7–3.2.13 below). In Section 3.2.4 we prove Theorem 2.3.1. In Sections 3.2.2 and 3.2.3 we will see that the moment relations are crucial when no duality is available.

Below we write $\mathbb{E}_z$ for $\mathbb{E}_{\delta_z}$, the expectation when the process starts from the initial distribution δ_z, $z \in E$.

§3.2.1 Moment relations

Lemma 3.2.1 (First and second moment). *For $z \in E$, $t \geq 0$ and $(i,R_i),(j,R_j) \in \mathbb{G}\times\{A,D\}$,*

$$\mathbb{E}_z[z_{(i,R_i)}(t)] = \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big)\, z_{(k,R_k)} \tag{3.38}$$

and

$$\begin{aligned}&\mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)]\\ &= \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big)\, b_t^{(1)}\big((j,R_j),(l,R_l)\big)\, z_{(k,R_k)} z_{(l,R_l)}\\ &\qquad + \int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{(t-s)}^{(1)}((i,R_i),(k,A))\, b_{(t-s)}^{(1)}((j,R_j),(k,A))\, \mathbb{E}_z[g(x_k(s))].\end{aligned} \tag{3.39}$$

Proof. We derive systems of differential equations for the moments and solve these in terms of the random walk. Let $(RW_t)_{t\geq 0}$ denote the semigroup of the random walk with transition kernel $b^{(1)}(\cdot,\cdot)$, and recall that the corresponding generator is given by

$$(G_{RW} f)(i,R_i) = \sum_{(j,R_j)\in\mathbb{G}\times\{A,D\}} b^{(1)}\big((i,R_i),(j,R_j)\big)\,[f(j,R_j) - f(i,R_i)]\,. \tag{3.40}$$

Applying the generator (2.25) of the system in (2.4)–(2.5) to the function

$$f_{(i,R_i)}\colon\ E \to \mathbb{R},\ f_{(i,R_i)}(z) = z_{(i,R_i)}, \tag{3.41}$$

we obtain by standard stochastic calculus

$$\begin{aligned}&\frac{\mathrm{d}\mathbb{E}_z[z_{(i,R_i)}(t)]}{\mathrm{d}t}\\ &= \left[\sum_{j\in\mathbb{G}} a(i,j)\,\big(\mathbb{E}_z[x_j(t)] - \mathbb{E}_z[x_i(t)]\big) + Ke\,(\mathbb{E}_z[y_i(t)] - \mathbb{E}_z[x_i(t)])\right] 1_{(R_i=A)}\\ &\qquad + e\,\big(\mathbb{E}_z[x_i(t)] - \mathbb{E}_z[y_i(t)]\big)\, 1_{(R_i=D)}.\end{aligned} \tag{3.42}$$

Hence, denoting by $(S_t)_{t\geq 0}$ the semigroup of the system in (2.4)–(2.5), we see from (3.42) and the definition of $b^{(1)}(\cdot,\cdot)$ in (2.31) that $(S_t f_{(i,R_i)})$ solves the differential equation

$$F'(t) = (G_{RW} F)(t). \tag{3.43}$$

On the other hand, for each $f \in \mathcal{C}_b(\mathbb{G} \times \{A, D\})$, $RW_t f$ also solves (3.43). In particular, for $z \in \mathbb{E}$ define $f_z \colon \mathbb{G} \times \{A, D\} \to \mathbb{R}$ by $f_z(i, R_i) = z(i, R_i)$ for $z \in E$, then $RW_t f_z$ is a solution to (3.43). Since

$$(RW_0 f_z)(i, R_i) = z(i, R_i) = (S_0 f_{(i,R_i)})(z), \tag{3.44}$$

we see that (3.38) holds. To prove (3.39), we derive a similar system of differential equations and again solve this in terms of the random walk moving according to the kernel $b(\cdot, \cdot)$. Let $f \colon E \to \mathbb{R}$ be given by $f(z) = z_{(i,R_i)} z_{(j,R_j)}$. Using the generator (2.25), we obtain via Itô-calculus that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\
&= \sum_{k \in \mathbb{G}} a(i,k) \left(\mathbb{E}_z[x_k(t) z_{(j,R_j)}(t)] - \mathbb{E}_z[x_i(t) z_{(j,R_j)}(t)] \right) 1_{\{R_i = A\}} \\
&\quad + Ke \left(\mathbb{E}_z[y_i(t) z_{(j,R_j)}(t)] - \mathbb{E}_z[x_i(t) z_{(j,R_j)}(t)] \right) 1_{\{R_i = A\}} \\
&\quad + e \left(\mathbb{E}_z[x_i(t) z_{(j,R_j)}(t)] - \mathbb{E}_z[y_i(t) z_{(j,R_j)}(t)] \right) 1_{\{R_i = D\}} \\
&\quad + \sum_{l \in \mathbb{G}} a(j,l) \left(\mathbb{E}_z[x_l(t) z_{(i,R_i)}(t)] - \mathbb{E}_z[x_j(t) z_{(i,R_i)}(t)] \right) 1_{\{R_j = A\}} \\
&\quad + Ke \left(\mathbb{E}_z[y_j(t) z_{(i,R_i)}(t)] - \mathbb{E}_z[x_j(t) z_{(i,R_i)}(t)] \right) 1_{\{R_j = A\}} \\
&\quad + e \left(\mathbb{E}_z[x_j(t) z_{(i,R_i)}(t)] - \mathbb{E}_z[y_j(t) z_{(i,R_i)}(t)] \right) 1_{\{R_j = D\}} \\
&\quad + \mathbb{E}_z[g(x_i(t))] \, 1_{\{i=j\}} \, 1_{\{R_i = R_j = A\}}.
\end{aligned} \tag{3.45}$$

Let U be the generator of two independent random walks each moving with transition kernel $b^{(1)}(\cdot, \cdot)$, i.e., for all $h \in \mathcal{C}_b((\mathbb{G} \times \{A, D\})^2)$,

$$\begin{aligned}
&(Uh)((i, R_i), (j, R_j)) \\
&= \sum_{k \in \mathbb{G}} a(i,k) \left[h((k, A), (j, R_j)) - h((i, R_i), (j, R_j)) \right] 1_{\{i, R_i = A\}} \\
&\quad + Ke \left[h((i, D), (j, R_j)) - h((i, R_i), (j, R_j)) \right] 1_{\{i, R_i = A\}} \\
&\quad + e \left[h((i, A), (j, R_j)) - h((i, R_i), (j, R_j)) \right] 1_{\{i, R_i = D\}} \\
&\quad + \sum_{l \in \mathbb{G}} a(j,l) \left[h((i, R_i), (l, A)) - h((i, R_i), (j, R_j)) \right] 1_{\{R_j = A\}} \\
&\quad + Ke \left[h((i, R_i), (j, D)) - h((i, R_i), (j, R_j)) \right] 1_{\{R_j = A\}} \\
&\quad + e \left[h((i, R_i), (j, A)) - h((i, R_i), (j, D)) \right] 1_{\{R_j = D\}}.
\end{aligned} \tag{3.46}$$

Let $F(t) = \mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)]$ and $H(t) = 2\mathbb{E}_z[g(x_i(t))] 1_{\{i=j\}} 1_{\{R_i = R_j = A\}}$. Then we can rewrite (3.45) as

$$\frac{\mathrm{d}}{\mathrm{d}t} F(t) = (UF)(t) + H(t). \tag{3.47}$$

Denote by $(RW_t^{(2)})_{t \geq 0}$ the semigroup corresponding to U. Applying [56, Theorem I.2.15], we obtain

$$F(t) = RW_t^{(2)} F(0) + \int_0^t \mathrm{d}s \, RW_{t-s}^{(2)} H(s). \tag{3.48}$$

Hence

$$\begin{aligned}&\mathbb{E}_z[z_{(i,R_i)}(t)z_{(j,R_j)}(t)]\\&=\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big)\, b_t^{(1)}\big((j,R_j),(l,R_l)\big)\,\mathbb{E}_z[z_{(k,R_k)}z_{(l,R_l)}]\\&\qquad+\int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{t-s}^{(1)}\big((i,R_i),(k,A)\big)\, b_{t-s}^{(1)}\big((j,R_j),(k,A)\big)\,\mathbb{E}_z[g(x_k(s))].\end{aligned} \tag{3.49}$$

□

Remark 3.2.2 (Density). From Lemma 3.2.1 we obtain that if μ is a translation invariant measure such that $\mathbb{E}_\mu[x_0(0)]=\theta_x$ and $\mathbb{E}_\mu[y_0(0)]=\theta_y$, then

$$\begin{aligned}\mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta_x &\sum_{(k,R_k)\in\mathbb{G}\times\{A\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big)\\&+\theta_y \sum_{(k,R_k)\in\mathbb{G}\times\{D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big),\end{aligned} \tag{3.50}$$

in particular, $\lim_{t\to\infty}\mathbb{E}_\mu[z_{(i,R_i)}(t)]=\frac{\theta_x+K\theta_y}{1+K}=\theta$, recall (2.62), and

$$\begin{aligned}&\mathbb{E}_\mu[z_{(i,R_i)}(t)z_{(j,R_j)}(t)]\\&=\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big)\, b_t^{(1)}\big((j,R_j),(l,R_l)\big)\,\mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}]\\&\qquad+2\int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{t-s}^{(1)}\big((i,R_i),(k,A)\big)\, b_{t-s}^{(1)}\big((j,R_j),(k,A)\big)\,\mathbb{E}_\mu[g(x_i(s))].\end{aligned} \tag{3.51}$$

□

Remark 3.2.3 (First moment duality). Note that (3.38) shows that even for general $g\in\mathcal{G}$ there is a *first moment duality* between the process $Z(t)$ and the random walk $RW(t)$, that moves according to the kernel $b^{(1)}(\cdot,\cdot)$. The duality function is given by

$$H: E\times\mathbb{G}\times\{A,D\}\to\mathbb{R},\qquad H(z,(i,R_i))=z_{(i,R_i)}. \tag{3.52}$$

Equation (3.38) in Lemma 3.2.1 tells us that $\mathbb{E}[H(Z(t),RW(0))]=\mathbb{E}[H(Z(0),RW(t))]$.

§3.2.2 The clustering case

The proof that the system in (2.4)–(2.5) converges to a unique trivial equilibrium when $\hat{a}(\cdot,\cdot)$ is recurrent goes as follows. We first consider the case where $g=dg_{\mathrm{FW}}$, for which *duality* is available (Lemma 3.2.4). Afterwards we use a *duality comparison argument* to show that the dichotomy between coexistence and clustering does not depend on the choice of $g\in\mathcal{G}$ (Lemma 3.2.5).

- **Case $g = dg_{\mathbf{FW}}$.**

Lemma 3.2.4 (Clustering). *Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ and $g = dg_{\mathrm{FW}}$. Moreover, suppose that $\hat{a}(\cdot,\cdot)$ defined in* (2.59) *is recurrent, i.e., $I_{\hat{a}} = \infty$. Let $\mu(t)$ be the law at time t of the system defined in* (2.4)–(2.5)*. Then*

$$\lim_{t\to\infty} \mu(t) = \theta\,[\delta_{(1,1)}]^{\otimes\mathbb{G}} + (1-\theta)\,[\delta_{(0,0)}]^{\otimes\mathbb{G}}. \tag{3.53}$$

Proof. Since $g = dg_{\mathrm{FW}}$, we can use duality. By the dichotomy criterion in Theorem 2.2.11, it is enough to show that in the dual two partition elements coalesce with probability 1. Recall from Section 2.2.4 that each of the partition elements in the dual moves according to the transition kernel $b^{(1)}(\cdot,\cdot)$ on $\mathbb{G} \times \{A, D\}$ defined by (2.31) (see Fig. 2.3). Recall from Section (2.2.4) that $b^{(1)}(\cdot,\cdot)$ describes a random walk on $\mathbb{G}$ with migration rate kernel $a(\cdot,\cdot)$ that becomes dormant (state D) at rate Ke (after which it stops moving), and becomes active (state A) at rate e (after which it can move again). When two partition elements in the dual are active and are at the same site, they coalesce at rate d, i.e., each time they are active and meet at the same site they coalesce with probability $d/[\sum_{j\in\mathbb{Z}^d} a(i,j) + Ke + d] > 0$. Hence, in order to show that two partition elements coalesce with probability 1, we have to show that *with probability* 1 *two partition elements meet infinitely often while being active.* The latter holds if and only if the expected total time the random walks spend together at the same colony while being active is infinite. We will show that this occurs if and only the random walk with *symmetrised* transition rate kernel $\hat{a}(\cdot,\cdot)$ is recurrent. The proof comes in 4 Steps.

1. Active and dormant time lapses. Consider two copies of the random walk with kernel $b^{(1)}(\cdot,\cdot)$, both starting at 0 and in the active state. Let

$$(\sigma_k)_{k\in\mathbb{N}}, \quad (\sigma'_k)_{k\in\mathbb{N}}, \tag{3.54}$$

denote the successive time lapses during which they are active and let

$$(\tau_k)_{k\in\mathbb{N}}, \quad (\tau'_k)_{k\in\mathbb{N}}, \tag{3.55}$$

denote the successive time lapses during which they are dormant (see Fig. 3.1). These are mutually independent sequences of i.i.d. random variables with marginal laws

$$\begin{array}{rclcll} \mathbb{P}(\sigma_1 > t) &=& \mathbb{P}(\sigma'_1 > t) &=& \mathrm{e}^{-Ke\,t}, & t \geq 0, \\ \mathbb{P}(\tau_1 > t) &=& \mathbb{P}(\tau'_1 > t) &=& \mathrm{e}^{-e\,t} & t \geq 0, \end{array} \tag{3.56}$$

where we use the symbol $\mathbb{P}$ to denote the joint law of the two sequences.

Let $a_t(\cdot,\cdot)$ denote the time-t transition kernel of the random walk with migration kernel $a(\cdot,\cdot)$. Let

$$\begin{aligned} \mathcal{E}(k,t) &= \left\{ \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) \leq t < \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) + \sigma_{k+1} \right\}, \\ \mathcal{E}'(k',t) &= \left\{ \sum_{\ell=1}^{k'} (\sigma'_\ell + \tau'_\ell) \leq t < \sum_{\ell=1}^{k'} (\sigma'_\ell + \tau'_\ell) + \sigma'_{k+1} \right\}, \end{aligned} \tag{3.57}$$

Chapter 3

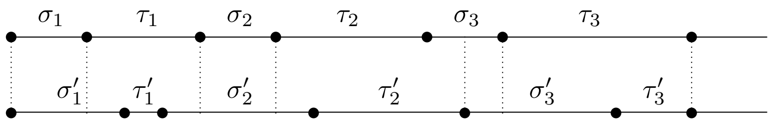

Figure 3.1: Successive periods during which the two random walks are active and dormant. The time lapses between the dotted lines represent periods of joint activity.

be the events that the random walks are active at time t after having become dormant and active exactly k, k' times, and let

$$
\begin{aligned}
T(k,t) &= \sum_{\ell=1}^{k} \sigma_\ell + \left(\left(t - \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) \right) \wedge \sigma_{k+1} \right), \\
T'(k',t) &= \sum_{\ell=1}^{k'} \sigma'_\ell + \left(\left(t - \sum_{\ell=1}^{k'} (\sigma'_\ell + \tau'_\ell) \right) \wedge \sigma_{k+1} \right),
\end{aligned}
\tag{3.58}
$$

be the total accumulated activity times of the random walks on the events in (3.57). Note that the terms between brackets in (3.58) are at most σ_{k+1}, respectively, $\sigma'_{k'+1}$, and therefore are negligible as $k, k' \to \infty$.

Given the outcome of the sequences in (3.54)–(3.55), the probability that at time t both random walks are active and are at the same colony equals

$$
\sum_{k,k' \in \mathbb{N}} \left(\sum_{i \in \mathbb{G}} a_{T(k,t)}(0,i)\, a_{T'(k',t)}(0,i) \right) 1_{\mathcal{E}(k,t)}\, 1_{\mathcal{E}(k',t)}, \tag{3.59}
$$

Therefore the expected total time the random walks are active and are at the same colony equals

$$
I = \int_0^\infty \mathrm{d}t \sum_{k,k' \in \mathbb{N}} \mathbb{E}_{(0,A),(0,A)} \left[\left(\sum_{i \in \mathbb{G}} a_{T(k,t)}(0,i)\, a_{T'(k',t)}(0,i) \right) 1_{\mathcal{E}(k,t)}\, 1_{\mathcal{E}'(k',t)} \right], \tag{3.60}
$$

where $\mathbb{E}$ is the expectation over the sequences in (3.54). Let

$$
\begin{aligned}
N(t) &= \max \left\{ k \in \mathbb{N} \colon \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) \leq t \right\}, \\
N'(t) &= \max \left\{ k' \in \mathbb{N} \colon \sum_{\ell=1}^{k'} (\sigma_\ell + \tau_\ell) \leq t \right\},
\end{aligned}
\tag{3.61}
$$

be the number of times the random walks have become dormant and active up to time t. Let

$$
T(t) = T(N(t),t), \quad T'(t) = T'(N'(t),t), \quad \mathcal{E}(t) = \mathcal{E}(N(t),t), \quad \mathcal{E}'(t) = \mathcal{E}'(N'(t),t), \tag{3.62}
$$

be the total accumulated activity times of the random walks up to time t, respectively, the events that the random walks are active at time t. Then we may write

$$I = \int_0^\infty \mathrm{d}t \; \mathbb{E}_{(0,A),(0,A)} \left[\left(\sum_{i \in \mathbb{G}} a_{T(t)}(0,i)\, a_{T'(t)}(0,i) \right) 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)} \right]. \tag{3.63}$$

We know that coalescence occurs with probability 1 if and only if $I = \infty$.

2. Fourier analysis. Define

$$M(t) = T(t) \wedge T'(t), \qquad \Delta(t) = [T(t) \vee T'(t)] - [T(t) \wedge T'(t)]. \tag{3.64}$$

Then

$$\sum_{i \in \mathbb{G}} a_{T(t)}(0,i)\, a_{T'(t)}(0,i) = \sum_{j \in \mathbb{G}} \hat{a}_{2M(t)}(0,j)\, a_{\Delta(t)}(j,0). \tag{3.65}$$

Indeed, the difference of the two random walks at time $M(t)$ has distribution $\hat{a}_{2M(t)}(0,\cdot)$, and in order for the random walk with the largest activity time to meet the random walk with the smallest activity time at time $2M(t)+\Delta(t)$, it must bridge this difference in time $\Delta(t)$. To work out (3.65), we assume without loss of generality that $\sum_{j \in \mathbb{G}} a(0,j) = 1$, and use *Fourier analysis.* For ease of exposition we focus on the special case where $\mathbb{G} = \mathbb{Z}^d$, but the argument below extends to *any* countable Abelian group endowed with the discrete topology, because these properties ensure that there is a version of Fourier analysis on $\mathbb{G}$ [64, Section 1.2]. For $\phi \in [-\pi,\pi]^d$, define

$$a(\phi) = \sum_{j \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i}(\phi,j)} a(0,j), \qquad \hat{a}(\phi) = \operatorname{Re} a(\phi), \qquad \tilde{a}(\phi) = \operatorname{Im} a(\phi). \tag{3.66}$$

Then

$$\begin{aligned} \hat{a}_t(0,j) &= \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-\mathrm{i}(\phi,j)}\, \mathrm{e}^{-t[1-\hat{a}(\phi)]}, \\ a_t(j,0) &= \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi' \, \mathrm{e}^{\mathrm{i}(\phi',j)}\, \mathrm{e}^{-t[1-\hat{a}(\phi')-\mathrm{i}\tilde{a}(\phi')]}, \end{aligned} \tag{3.67}$$

where we use that $a(\phi) = \hat{a}(\phi) + \mathrm{i}\tilde{a}(\phi)$. Inserting these representations into (3.65), we get

$$\sum_{i \in \mathbb{Z}^d} a_{T(t)}(0,i)\, a_{T'(t)}(0,i) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-[2M(t)+\Delta(t)]\,[1-\hat{a}(\phi)]} \cos(\Delta(t)\tilde{a}(\phi)), \tag{3.68}$$

where we use that $\sum_{j \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i}(\phi'-\phi,j)} = (2\pi)^d \delta(\phi'-\phi)$, with $\delta(\cdot)$ the Dirac distribution (Folland [37, Chapter 7]).

3. Limit theorems. By the strong law of large numbers, we have

$$\lim_{k \to \infty} \frac{1}{k} \sum_{\ell=1}^{k} \sigma_\ell = \frac{1}{Ke} \quad \mathbb{P}\text{-a.s.}, \qquad \lim_{k \to \infty} \frac{1}{k} \sum_{\ell=1}^{k} \tau_\ell = \frac{1}{e} \quad \mathbb{P}\text{-a.s.} \tag{3.69}$$

Therefore, by the standard renewal theorem (Asmussen [3, Chapter I, Theorem 2.2]),

$$\begin{aligned}
&\lim_{t\to\infty} \frac{1}{t}\, N(t) = \lim_{t\to\infty} \frac{1}{t}\, N'(t) = A \quad \mathbb{P}\text{-a.s.},\\
&\lim_{t\to\infty} \frac{1}{t}\, T(t) = \lim_{t\to\infty} \frac{1}{t}\, T'(t) = B \quad \mathbb{P}\text{-a.s.},\\
&\lim_{t\to\infty} \mathbb{P}\big(\mathcal{E}(t)\big) = \lim_{t\to\infty} \mathbb{P}\big(\mathcal{E}'(t)\big) = B,
\end{aligned} \tag{3.70}$$

with

$$A = \frac{1}{\frac{1}{Ke} + \frac{1}{e}} = \frac{K}{1+K}\, e, \qquad B = \frac{\frac{1}{Ke}}{\frac{1}{Ke} + \frac{1}{e}} = \frac{1}{1+K}. \tag{3.71}$$

Moreover, by the central limit theorem, we have

$$\left(\frac{T(t)-Bt}{c\sqrt{t}}, \frac{T'(t)-Bt}{c\sqrt{t}}\right) \quad \Longrightarrow \quad (Z,Z') \quad \text{in } \mathbb{P}\text{-distribution as } t\to\infty \tag{3.72}$$

with (Z,Z') independent standard normal random variables and

$$c^2 = A\left[(1-B)^2\, \mathbb{V}\mathrm{ar}(\sigma_1) + B^2\, \mathbb{V}\mathrm{ar}(\tau_1)\right] \tag{3.73}$$

(see [68] or [3, Theorem VI.3.2]). Since $T(t),\mathcal{E}(t)$ and $T'(t),\mathcal{E}'(t)$ are independent, and each pair is asymptotically independent as well, we find that

$$\mathbb{E}_{(0,A),(0,A)}\left[\left(\sum_{i\in\mathbb{Z}^d} a_{T(t)}(0,i)\, a_{T'(t)}(0,i)\right) 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)}\right] \sim B^2 f(t), \qquad t\to\infty, \tag{3.74}$$

with

$$\begin{aligned}
f(t) &= \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi\, \mathrm{e}^{-[1+o(1)]\,2Bt\,[1-\hat{a}(\phi)]}\, \mathbb{E}\left[\cos\Big([1+o(1)]\, c(Z-Z')\sqrt{t}\,\tilde{a}(\phi)\Big)\right]\\
&= \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi\, \mathrm{e}^{-[1+o(1)]\,2Bt\,[1-\hat{a}(\phi)]}\, \mathrm{e}^{-[1+o(1)]\,c^2 t\, \tilde{a}(\phi)^2},
\end{aligned} \tag{3.75}$$

where we use that cos is symmetric, $Z-Z' = \sqrt{2}\, Z''$ in $\mathbb{P}$-distribution with Z'' standard normal, and $\mathbb{E}(\mathrm{e}^{\mathrm{i}\mu Z''}) = \mathrm{e}^{-\mu^2/2}$, $\mu\in\mathbb{R}$. From (3.63) and (3.74) we have that $I<\infty$ if and only if $t\mapsto f(t)$ is integrable. By Cramér's theorem, deviations of $T(t)/t$ and $T'(t)/t$ away from B are exponentially costly in t. Hence the error terms in (3.75), arising from (3.70) and (3.72), do *not* affect the integrability of $t\mapsto f(t)$. Note that, because $a(\cdot,\cdot)$ is assumed to be *irreducible* (recall (2.1)), $\hat{a}(\phi)=1$ if and only if $\phi=0$. Hence the integrability of $t\mapsto f(t)$ is determined by the behaviour of $\hat{a}(\phi)$ and $\tilde{a}(\phi)$ as $\phi\to 0$.

4. Irrelevance of asymmetric part of migration. We next observe that

$$\tilde{a}(\phi)^2 \le 1-\hat{a}(\phi)^2 \le 2[1-\hat{a}(\phi)]. \tag{3.76}$$

Hence, $t\,\tilde{a}(\phi)^2 \le 2t\,[1-\hat{a}(\phi)]$. Therefore we see from (3.75) that for sufficiently large $T\in\mathbb{R}$ we can bound $t\mapsto f(t)$ on $[T,\infty)$ from above and below by functions of the form $t \mapsto g_C(t)$ with

$$g_C(t) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-\,Ct\,[1-\hat{a}(\phi)]}, \qquad C \in (0,\infty). \tag{3.77}$$

From (3.67) we have

$$g_C(t) = \hat{a}_{Ct}(0,0) \asymp \hat{a}_t(0,0), \tag{3.78}$$

where the last asymptotics uses that $t \mapsto \hat{a}_t(0,0)$ is regularly varying at infinity (recall (2.60)). Combining (3.63), (3.74) and (3.77)–(3.78), we get

$$I = \infty \quad \Longleftrightarrow \quad I_{\hat{a}} = \infty \tag{3.79}$$

with $I_{\hat{a}} = \int_1^\infty \mathrm{d}t \, \hat{a}_t(0,0)$. Thus, if $\hat{a}(\cdot,\cdot)$ is recurrent, then $I = \infty$ and the system clusters. Moreover, we see from the bounds on $f(t)$ (recall (3.75)) that *the asymmetric part of the migration kernel has no effect on the integrability.*

This settles the dichotomy between clustering and coexistence when $g = g_{\mathrm{FW}}$. $\square$

• **Case $g \neq dg_{\mathbf{FW}}$.** For $g \neq dg_{\mathrm{FW}}$ the proof of Lemma 3.2.4 does not go through. However, the *moments relations* in Lemma 3.2.1 hold for general $g \in \mathcal{G}$. Using these moment relations and a technique called duality comparison (see [14]), we prove Lemma 3.2.4 for general $g \in \mathcal{G}$.

Lemma 3.2.5 (Duality comparison). *Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ and $g \in \mathcal{G}$. Moreover, suppose that $\hat{a}(\cdot,\cdot)$ defined in* (2.59) *is recurrent, i.e., $I_{\hat{a}} = \infty$. Let $\mu(t)$ be the law at time t of the system defined in* (2.4)–(2.5)*. Then*

$$\lim_{t\to\infty} \mu(t) = \theta \, [\delta_{(1,1)}]^{\otimes \mathbb{G}} + (1-\theta) \, [\delta_{(0,0)}]^{\otimes \mathbb{G}}. \tag{3.80}$$

Proof. We proceed as in the proof of [14, Theorem]. First assume that $\mu(0) = \delta_z$ for some $z \in E$, that satisfies

$$\lim_{t\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}((i,R_i),(k,R_k)) \, z_{(k,R_k)} = \theta. \tag{3.81}$$

By Lemma 3.2.1, we have

$$\mathbb{E}_z\left[z_{(i,R_i)}(t)\right] = \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}((i,R_i),(k,R_k)) \, z_{(k,R_k)}. \tag{3.82}$$

Hence, by assumption, for all $(i,R_i) \in \mathbb{G} \times \{A,D\}$ we have

$$\lim_{t\to\infty} \mathbb{E}_z\left[z_{(i,R_i)}(t)\right] = \theta. \tag{3.83}$$

Since we have clustering if, for all $(i,R_i),(j,R_j) \in \mathbb{G} \times \{A,D\}$,

$$\lim_{t\to\infty} \mathbb{E}_z\left[z_{(i,R_i)}(t)(1-z_{(j,R_j)}(t))\right] = 0, \tag{3.84}$$

we are left to prove that

$$\lim_{t\to\infty} \mathbb{E}_z\left[z_{(i,R_i)} z_{(j,R_j)}\right] = \theta. \tag{3.85}$$

Since (3.83) implies that $\limsup_{t\to\infty} \mathbb{E}_z[z_{(i,R_i)}z_{(j,R_j)}] \leq \theta$, we are left to prove that

$$\liminf_{t\to\infty} \mathbb{E}_z[z_{(i,R_i)}z_{(j,R_j)}] \geq \theta. \tag{3.86}$$

Like in [14], we will prove (3.86) by *comparison duality.*

Fix $\epsilon > 0$. Since $g \in \mathcal{G}$ we can choose a $c = c(\epsilon) > 0$ such that

$$g(x) \geq \tilde{g}(x) = c(x-\epsilon)(1-(x+\epsilon)),\ x \in [0,1]. \tag{3.87}$$

Note that $\tilde{g}(x) < 0$ for $x \in [0,\epsilon) \cup (1-\epsilon, 1]$, so we cannot replace g by $\tilde{g}$ in the SSDE. Instead we use $\tilde{g}$ as an auxiliary function.

Consider the Markov chain $(B(t))_{t\geq 0}$, with state space

$$\{1,2\} \times (\mathbb{G} \times \{A,D\}) \times (\mathbb{G} \times \{A,D\}) \tag{3.88}$$

and $B(t) = (B_0(t), B_1(t), B_2(t))$, evolving according to

$$\begin{aligned} (1,(i,R_i),(i,R_i)) &\to\ (1,(k,R_k),(k,R_k)), \quad \text{at rate } b^{(1)}((i,R_i),(k,R_k)),\\ (2,(i,R_i),(j,R_j)) &\to \begin{cases} (2,(k,R_k),(j,R_j)), & \text{at rate } b^{(1)}((i,R_i),(k,R_k)),\\ (2,(i,R_i),(l,R_l)), & \text{at rate } b^{(1)}((j,R_j),(l,R_l)),\\ (1,(i,R_i),(i,R_i)), & \text{at rate } c1_{\{i=j\}}1_{\{R_i=R_j=A\}}. \end{cases} \end{aligned} \tag{3.89}$$

This describes two random walks, evolving independently according to the transition kernel $b^{(1)}(\cdot,\cdot)$, that coalesce at rate $c > 0$ when they are at the same site and are active. We put $B_0(t) = 1$ when the two random walks have already coalesced by time t, and $B_0(t) = 2$ otherwise. Let $\mathbb{P}_{(2,(i,R_i),(j,R_j))}$ denote the law of the Markov Chain $B(t)$ that starts in $(2,(i,R_i),(j,R_j))$. Note that

$$\mathbb{P}_{(2,(i,R_i),(j,R_j))}\,(B_1(t) = (k,R_k)) = b_t^{(1)}((i,R_i),(k,R_k)), \tag{3.90}$$

and similarly

$$\mathbb{P}_{(2,(i,R_i),(j,R_j))}\,(B_2(t) = (l,R_l)) = b_t^{(1)}((j,R_j),(l,R_l)). \tag{3.91}$$

Since we have assumed that $\hat{a}(\cdot,\cdot)$ is recurrent, i.e., $I_{\hat{a}} = \infty$, the two random walks meet infinitely often at the same site while being active and hence coalesce with probability 1. Therefore

$$\lim_{t\to\infty} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\,(B_0(t) = 2) = 0. \tag{3.92}$$

We can rewrite the SSDE in (2.4)–(2.5) in terms of $b^{(1)}(\cdot,\cdot)$, namely, for all $(i,R_i) \in \mathbb{G} \times \{A,D\}$,

$$\begin{aligned} \mathrm{d}z_{(i,R_i)}(t) = &\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((i,R_i),(j,R_j))[z_{(j,R_j)}(t) - z_{(i,R_i)}(t)]\,\mathrm{d}t\\ &+ \sqrt{g(z_{i,R_i}(t))}\,1_{\{R_i=A\}}\,\mathrm{d}w_i(t). \end{aligned} \tag{3.93}$$

Using (3.93) and Itô-calculus, we obtain

$$
\begin{aligned}
&\frac{\mathrm{d}\mathbb{E}_z[z_{(i,R_i)}(t)-\epsilon]}{\mathrm{d}t}\\
&=\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((i,R_i),(k,R_k))\,\mathbb{E}\left[(z_{(k,R_k)}(t)-\epsilon)-(z_{(i,R_i)}(t)-\epsilon)\right]
\end{aligned}
\tag{3.94}
$$

and

$$
\begin{aligned}
&\frac{\mathrm{d}\mathbb{E}_z[(z_{(i,R_i)}(t)-\epsilon)(z_{(j,R_j)}(t)+\epsilon)]}{\mathrm{d}t}\\
&=\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((i,R_i),(k,R_k))\\
&\qquad\times\mathbb{E}_z\left[(z_{(j,R_j)}(t)+\epsilon)(z_{(k,R_k)}(t)-\epsilon)-(z_{(j,R_j)}(t)+\epsilon)(z_{(i,R_i)}(t)-\epsilon)\right]\\
&\quad+\sum_{(l,R_l)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,R_j),(k,R_k))\\
&\qquad\times\mathbb{E}_z\left[(z_{(i,R_i)}(t)-\epsilon)(z_{(l,R_l)}(t)+\epsilon)-(z_{(i,R_i)}(t)-\epsilon)(z_{(j,R_j)}(t)+\epsilon)\right]\\
&\quad+\mathbb{E}_z\left[c(z_{(i,R_i)}(t)-\epsilon)(1-(z_{(j,R_j)}(t)+\epsilon))1_{\{i=j\}}1_{\{R_i=R_j=A\}}\right]\\
&\quad+\mathbb{E}_z\left[\left(g(z_{(i,R_i)}(t))-\tilde{g}(z_{(i,R_i)}(t))\right)1_{\{i=j\}}1_{\{R_i=R_j=A\}}\right].
\end{aligned}
\tag{3.95}
$$

For $t\geq 0$, define $F_t\colon\ \{0,1\}\times(\mathbb{G}\times\{A,D\})\times(\mathbb{G}\times\{A,D\})\to\mathbb{R}$ by

$$
\begin{aligned}
F_t(1,(i,R_i),(i,R_i)) &= \mathbb{E}_z\left[z_{(i,R_i)}(t)-\epsilon\right]\\
F_t(2,(i,R_i),(j,R_j)), &= \mathbb{E}_z\left[(z_{(i,R_i)}(t)-\epsilon)(z_{(j,R_j)}(t)+\epsilon)\right],
\end{aligned}
\tag{3.96}
$$

and $H_t\colon\ \{0,1\}\times(\mathbb{G}\times\{A,D\})\times(\mathbb{G}\times\{A,D\})\to\mathbb{R}$ by

$$
\begin{aligned}
&H_t(1,(i,R_i),(i,R_i))=0,\\
&H_t(2,(i,R_i),(j,R_j))=\mathbb{E}_z\left[\left(g(z_{(i,R_i)}(t))-\tilde{g}(z_{(i,R_i)}(t))\right)1_{\{i=j\}}\,1_{\{R_i=R_j=A\}}\right].
\end{aligned}
\tag{3.97}
$$

Let $\mathfrak{B}$ denote the generator of $(B(t))_{t\geq0}$, and let $(V_t)_{t\geq0}$ the associated semigroup. Then

$$
\frac{\mathrm{d}F_t}{\mathrm{d}t}=\mathfrak{B}F_t+H_t. \tag{3.98}
$$

Hence, by [56, Theorem I.2.15], it follows that

$$
F_t=V_tF_0+\int_0^t\mathrm{d}s\,V_{(t-s)}H_s. \tag{3.99}
$$

Since $H_t > 0$ for all $t \geq 0$, we obtain

$$
\begin{aligned}
&F_t(2,(i,R_i),(j,R_j)) \geq V_t F_0(2,(i,R_i),(j,R_j)) \\
&= \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))\right] \\
&= \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))1_{\{B_0(t)=1\}} + F_0(B(t))1_{\{B_0(t)=2\}}\right] \\
&= \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_0(t)=1, B_1(t)=(k,R_k)\right](z_{(k,R_k)}-\epsilon) \\
&\quad + \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))1_{\{B_0(t)=2\}}\right] \\
&= \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_1(t)=(k,R_k)\right](z_{(k,R_k)}-\epsilon) \\
&\quad - \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_0(t)=2, B_1(t)=(k,R_k)\right](z_{(k,R_k)}-\epsilon) \\
&\quad + \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))1_{\{B_0(t)=2\}}\right] \\
&\geq \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}((i,R_i),(k,R_k))\,(z_{(k,R_k)}-\epsilon) \\
&\quad - (1+\epsilon^2)\,\mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_1(t)=2\right].
\end{aligned}
\tag{3.100}
$$

Hence, by (3.92), we obtain

$$
\begin{aligned}
\liminf_{t\to\infty} F_t(2,(i,R_i),(j,R_j)) &\geq \liminf_{t\to\infty} \mathbb{E}_z\left[(z_{(i,R_i)}(t)-\epsilon)(z_{(j,R_j)}(t)+\epsilon)\right] \\
&\geq \theta - \epsilon^2.
\end{aligned}
\tag{3.101}
$$

Letting $\epsilon \downarrow 0$, we get (3.85).

To get rid of the assumption $\mu(0) = \delta_z$, note that for $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ we have (recall Remark 3.2.2)

$$
\lim_{t\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t((i,R_i),(k,R_k))\mathbb{E}_\mu[z_{(k,R_k)}] = \theta. \tag{3.102}
$$

Hence, by the above argument,

$$
\begin{aligned}
&\mathbb{E}_\mu\left[(z_{(i,R_i)}(t)-\epsilon)(z_{j,R_j}(t)+\epsilon)\right] \\
&= \int \mathbb{E}_z\left[(z_{(i,R_i)}(t)-\epsilon)(z_{j,R_j}(t)+\epsilon)\right] \mathrm{d}\mu(z) \\
&\geq \int \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}((i,R_i),(k,R_k))(z_{(k,R_k)}-\epsilon) \\
&\qquad - (1+\epsilon^2)\mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_1(t)=2\right] \mathrm{d}\mu(z)
\end{aligned}
\tag{3.103}
$$

Letting first $t \to \infty$ and then $\epsilon \downarrow 0$, we find that

$$
\lim_{t\to\infty} \mathbb{E}_\mu\left[(z_{(i,R_i)}(t)-\epsilon)(z_{j,R_j}(t)+\epsilon)\right] = \theta, \tag{3.104}
$$

and, for all $(i,R_i),(j,R_j) \in \mathbb{G}\times\{A,D\}$,

$$
\lim_{t\to\infty} \mathbb{E}_\mu\left[z_{(i,R_i)}(t)(1-z_{j,R_j}(t))\right] = 0. \tag{3.105}
$$

□

§3.2.3 The coexistence case

For the coexistence case we proceed as in [14] with small adaptations. For the convenience of the reader we have written out the full proof. The proof relies on the moment relations in Lemma 3.2.1 and no distinction between $g = dg_{\mathrm{FW}}$ and general $g \in \mathcal{G}$ is needed. The proof consist of several lemmas (Lemmas 3.2.7–3.2.13 below), organised into 4 Steps. In Step 1 we use the moment relations in Lemma 3.2.1 to define a set of measures that are preserved under the evolution. In Step 2 we use coupling to prove that, for each given θ, the system converges to a unique equilibrium. In Step 3 we show that, for each given θ, each initial measure under the evolution converges to an invariant measure. In Step 4 we show that the limiting measure is invariant, ergodic and mixing under translations, and is associated.

1. Properties of measures preserved under the evolution. Let θ be defined as in (2.62) such that $\theta = \mathbb{E}_{\mu(0)}\left[\frac{x_0+Ky_0}{1+K}\right] = \frac{\theta_x+K\theta_y}{1+K}$.

Definition 3.2.6 (Preserved class of measure). Let $\mathcal{R}_\theta^{(1)}$ denote the set of measures $\mu \in \mathcal{T}$ satisfying:

(1) For all $(i,R_i) \in \mathbb{G} \times \{A,D\}$,

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta. \tag{3.106}$$

(2) for all (i,R_i), $(j,R_j) \in \mathbb{G} \times \{A,D\}$,

$$\begin{aligned}\lim_{t\to\infty} \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} & b_t^{(1)}\big((i,R_i),(k,R_k)\big)\, b_t^{(1)}\big((j,R_j),(l,R_l)\big)\\ & \times \mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}] = \theta^2.\end{aligned} \tag{3.107}$$

□

Clearly, if $\mu \in \mathcal{R}_\theta^{(1)}$, then (1) and (2) together with Lemma 3.2.1 imply

$$\lim_{t\to\infty} \mathbb{E}_\mu\left[\left(\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}((i,R_i),(k,R_k))\, z_{(k,R_k)} - \theta\right)^2\right] = 0, \tag{3.108}$$

and so $\lim_{t\to\infty} z_{i,R_i}(t) = \theta$ in $L^2(\mu)$.

On the other hand, suppose that (3.108) holds for *some* $(i,R_i) \in \mathbb{G} \times \{A,D\}$. Then, by Lemma (3.2.1), we can rewrite (3.108) as

$$\lim_{t\to\infty} \mathbb{E}_\mu\left[\left(\mathbb{E}_z[z_{(i,R_i)}(t)] - \theta\right)^2\right] = 0. \tag{3.109}$$

This implies

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta, \tag{3.110}$$

and hence, by translation invariance,

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(k,R_i)}(t)] = \theta \qquad \forall\, k \in \mathbb{G}. \tag{3.111}$$

Using that switches between the active state at the dormant state occur at a positive rate, we can use the strong Markov property to obtain that (3.111) holds both for $R_i = A$ and for $R_i = D$. Hence (3.106) holds. Combining (3.106) and (3.108), we see that also (3.107) holds.

Lemma 3.2.7. *$\mu \in \mathcal{R}_\theta^{(1)}$ for all $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$.*

Proof. The proof relies on *Fourier analysis* and the existence of spectral measures. As in Section 3.2.2, for ease of exposition we focus on the special case where $\mathbb{G} = \mathbb{Z}^d$, but the argument below extends to *any* countable Abelian group endowed with the discrete topology.

By translation invariance and the Herglotz theorem, there exist spectral measures λ_A and λ_D such that, for all $j, k \in \mathbb{Z}^d$,

$$\begin{aligned}
\mathbb{E}_\mu\left[(x_j-\theta_x)(x_k-\theta_x)\right] &= \int_{(-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}(j-k,\phi)}\mathrm{d}\lambda_A(\phi),\\
\mathbb{E}_\mu\left[(y_j-\theta_y)(y_k-\theta_y)\right] &= \int_{(-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}(j-k,\phi)}\mathrm{d}\lambda_D(\phi).
\end{aligned} \tag{3.112}$$

Let $a(\phi) = \sum_{k\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}(\phi,j)} a(0,k)$ be the characteristic function of the kernel $a(\cdot,\cdot)$ (recall (3.66)), and $T(t)$ the activity time of the random walk up to time t (recall (3.58)). Then

$$\begin{aligned}
\sum_{k\in\mathbb{Z}^d} a_{T(t)}(0,k)\,\mathrm{e}^{\mathrm{i}(k,\phi)} &= \sum_{n\in\mathbb{N}_0} \frac{\mathrm{e}^{-T(t)}[T(t)]^n}{n!} \sum_{k\in\mathbb{Z}^d} a^n(0,k)\,\mathrm{e}^{\mathrm{i}(k,\phi)}\\
&= \sum_{n\in\mathbb{N}_0} \frac{\mathrm{e}^{-T(t)}[T(t)\,a(\phi)]^n}{n!}\\
&= \mathrm{e}^{-T(t)(1-a(\phi))}.
\end{aligned} \tag{3.113}$$

Let $\mathcal{E}(t)$ be defined as in (3.62). Then, for fixed $t > 0$,

$$\mathbb{P}_{(0,A)}(\mathcal{E}(t)) = \sum_{k\in\mathbb{Z}^d} b_t^{(1)}((0,A),(k,A)) > 0. \tag{3.114}$$

and hence

$$
\begin{aligned}
&\mathbb{E}_\mu\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\sum_{k\in\mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))x_k-\theta_x\right)^2\right]\\
&=\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\sum_{k,l\in\mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))\,b_t^{(1)}((0,A),(l,A))\,\mathbb{E}_\mu\left[(x_k-\theta_x)(x_l-\theta_x)\right]\\
&=\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\sum_{k,l\in\mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))\,b_t^{(1)}((0,A),(l,A))\\
&\qquad\qquad\times\int_{(-\pi,\pi]^d}\mathrm{e}^{\mathrm{i}(k-l,\phi)}\mathrm{d}\lambda_A(\phi)\\
&=\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\sum_{k,l\in\mathbb{Z}^d}\mathbb{E}_{(0,A),(0,A)}\left[a_{T(t)}(0,k)\,a_{T'(t)}(0,l)\,1_{\mathcal{E}(t)}\,1_{\mathcal{E}'(t)}\right]\\
&\qquad\qquad\times\int_{(-\pi,\pi]^d}\mathrm{e}^{\mathrm{i}(k-l,\phi)}\mathrm{d}\lambda_A(\phi)\\
&=\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\\
&\times\int_{(-\pi,\pi]^d}\mathbb{E}_{(0,A),(0,A)}\left[\sum_{k\in\mathbb{Z}^d}a_{T(t)}\,\mathrm{e}^{\mathrm{i}(k,\phi)}(0,k)1_{\mathcal{E}(t)}\sum_{l\in\mathbb{Z}^d}a_{T'(t)}\,\mathrm{e}^{-\mathrm{i}(l,\phi)}(0,l)1_{\mathcal{E}'(t)}\right]\mathrm{d}\lambda_A(\phi)\\
&=\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\int_{(-\pi,\pi]^d}\mathbb{E}_{(0,A),(0,A)}\left[\mathrm{e}^{-T(t)(1-a(\phi))}1_{\mathcal{E}(t)}\,\mathrm{e}^{-T'(t)(1-\bar{a}(\phi))}1_{\mathcal{E}'(t)}\right]\mathrm{d}\lambda_A(\phi).
\end{aligned}
\tag{3.115}
$$

Since $a(\cdot,\cdot)$ is irreducible, $a(\phi)\neq 1$ for all $\phi\in(-\pi,\pi]^d\backslash\{0\}$. Taking the limit $t\to\infty$, we find

$$
\lim_{t\to\infty}\mathbb{E}_\mu\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\sum_{k\in\mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))x_k-\theta_x\right)^2\right]=\lambda_A(\{0\}). \tag{3.116}
$$

Similarly,

$$
\lim_{t\to\infty}\mathbb{E}_\mu\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))}\sum_{k\in\mathbb{Z}^d} b_t^{(1)}((0,A),(k,D))y_k-\theta_y\right)^2\right]=\lambda_D(\{0\}). \tag{3.117}
$$

Hence

$$
\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}_\mu \left[\left(\sum_{(k,R_k)\in\mathbb{Z}^d\times\{A,D\}} b_t^{(1)}((0,A),(k,R_k)\, z_{(k,R_k)} - \theta \right)^2 \right] \\
&= \lim_{t\to\infty} \mathbb{E}_\mu \Bigg[\Bigg(\mathbb{P}_{(0,A)}(\mathcal{E}(t)) \sum_{k\in\mathbb{Z}^d} \frac{b_t^{(1)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\, x_k - \frac{\theta_x}{1+K} \\
&\qquad + \mathbb{P}_{(0,A)}(\mathcal{E}^c(t)) \sum_{k\in\mathbb{Z}^d} \frac{b_t^{(1)}((0,A),(k,D))}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))}\, y_k - \frac{K\theta_y}{1+K} \Bigg)^2 \Bigg] \\
&\le \lim_{t\to\infty} \mathbb{P}_{(0,A)}(\mathcal{E}(t))\, \mathbb{E}_\mu \left[\left(\sum_{k\in\mathbb{Z}^d} \frac{b_t^{(1)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\, x_k - \frac{\theta_x}{(1+K)} \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))} \right)^2 \right] \\
&\qquad + \mathbb{P}_{(0,A)}(\mathcal{E}^c(t))\, \mathbb{E}_\mu \left[\left(\sum_{k\in\mathbb{Z}^d} \frac{b_t^{(1)}((0,A),(k,D))}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))}\, y_k - \frac{K\theta_y}{1+K} \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))} \right)^2 \right] \\
&= \frac{1}{1+K}\lambda_A(\{0\}) + \frac{K}{1+K}\lambda_D(\{0\}).
\end{aligned}
\tag{3.118}
$$

Hence, if $\lambda_A(\{0\}) = 0$ and $\lambda_D(\{0\}) = 0$, then $\mu \in \mathcal{R}_\theta^{(1)}$. We will show that $\lambda_A(\{0\}) = 0$ and $\lambda_D(\{0\}) = 0$ for $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$.

Let $\Lambda_N = [0,N)^d \cap \mathbb{Z}^d$. By the L^1-ergodic theorem, we have, for $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$,

$$
\lim_{N\to\infty} \mathbb{E}_\mu \left[\left(\frac{1}{\Lambda_N} \sum_{j\in\Lambda_N} x_j - \theta_x \right)^2 \right] = 0. \tag{3.119}
$$

(For general $\mathbb{G}$ not that countable groups endowed with the discrete topology are amenable. For amenable groups $\mathbb{G}$, $(\Lambda_N)_{N\in\mathbb{N}}$ must be replaced by a so-called Følner sequence, i.e., a sequence of finite subsets of $\mathbb{G}$ that exhaust $\mathbb{G}$ and satisfy

$$
\lim_{N\to\infty} |\mathfrak{g}\Lambda_N \triangle \Lambda_N|/|\Lambda_N| = 0 \tag{3.120}
$$

for any $\mathfrak{g} \in \mathbb{G}$ [57].) Using the spectral measure, we can write

$$
\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}_\mu \left[\left(\frac{1}{\Lambda_N} \sum_{j\in\Lambda_N} x_j - \theta_x \right)^2 \right] \\
&= \lim_{N\to\infty} \frac{1}{\Lambda_N^2} \sum_{j,k\in\Lambda_N} \int_{(-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}(j-k,\phi)} \mathrm{d}\lambda_A \\
&= \lim_{N\to\infty} \int_{(-\pi,\pi]^d} \left(\frac{1}{\Lambda_N} \sum_{j\in\Lambda_N} \mathrm{e}^{\mathrm{i}(j,\phi)} \right) \left(\frac{1}{\Lambda_N} \sum_{k\in\Lambda_N} \mathrm{e}^{-\mathrm{i}(k,\phi)} \right) \mathrm{d}\lambda_A \\
&= \lambda_A\{0\}.
\end{aligned}
\tag{3.121}
$$

In the last equality we use dominated convergence and

(a) For all $\phi \in (-\pi, \pi]^d$,

$$\lim_{N\to\infty} \frac{1}{\Lambda_N} \sum_{j,k\in\Lambda_N} \mathrm{e}^{-\mathrm{i}(k,\phi)} = 1_{\{0\}}(\phi). \tag{3.122}$$

(b) For all $\delta > 0$ there exist $\epsilon(N,\delta) > 0$ such that if $J_\delta = (-\delta, \delta)$, then

$$\left| \frac{1}{\Lambda_N} \sum_{j,k\in\Lambda_N} \mathrm{e}^{-\mathrm{i}(k,\phi)} - 1_{\{0\}}(\phi) \right| \leq 1_{J_\delta}(\phi) + \epsilon(N,\delta), \tag{3.123}$$

where $\epsilon(N,\delta) \downarrow 0$ as $N \to \infty$.

We conclude that $\lambda_A(\{0\}) = 0$. Similarly we can show that $\lambda_D(\{0\}) = 0$, and hence $\mu \in \mathcal{R}^{(1)}_\theta$. □

Recall that $(S_t)_{t\geq 0}$ is the semigroup associated with (2.4)–(2.5).

Lemma 3.2.8 (Preservation). *If $b(\cdot,\cdot)$ is transient and $\mu \in \mathcal{R}^{(1)}_\theta$, then the following hold:*

(a) $\mu S_t \in \mathcal{R}^{(1)}_\theta$ *for each* $t \geq 0$.

(b) *If* $t_n \to \infty$ *and* $\mu S_{t_n} \to \mu(\infty)$, *then* $\mu(\infty) \in \mathcal{R}^{(1)}_\theta$.

Proof. Our dynamics preserve translation invariance. To check property (1) of $\mathcal{R}^{(1)}_\theta$ (see (3.106)), set $f(z) = z_{(i,R_i)}$. Since $\mu \in \mathcal{R}^{(1)}_\theta$, applying Lemma 3.2.1 multiple times, we obtain

$$\begin{aligned}
\lim_{s\to\infty} \mathbb{E}_{\mu S_t}[z_{(i,R_i)}(s)] &= \lim_{s\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}_s\big((i,R_i),(k,R_k)\big)\, \mathbb{E}_{\mu S_t}[z_{(k,R_k)}] \\
&= \lim_{s\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}_s\big((i,R_i),(k,R_k)\big)\, \mathbb{E}_{\mu}[z_{(k,R_k)}(t)] \\
&= \lim_{s\to\infty} \sum_{(k',R'_{k'})\in\mathbb{G}\times\{A,D\}} b^{(1)}_{s+t}\big((i,R_i),(k',R_{k'})\big)\, \mathbb{E}_{\mu}[z_{(k',R_{k'})}] \\
&= \lim_{s\to\infty} \mathbb{E}_{\mu}[z_{(i,R_i)}(t+s)] = \theta.
\end{aligned} \tag{3.124}$$

To check property (2) of $\mathcal{R}^{(1)}_\theta$ (see (3.107)), we set $f(z) = z_{(i,R_i)} z_{(j,R_j)}$. Then, again by applying Lemma 3.2.1, we find

$$
\begin{aligned}
&\lim_{s\to\infty} \sum_{\substack{(k,R_k),(l,R_l)\\ \in\mathbb{G}\times\{A,D\}}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, b_s^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_{\mu S_t}[z_{(k,R_k)} z_{(l,R_l)}]\\
&= \lim_{s\to\infty} \sum_{\substack{(k,R_k),(l,R_l)\\ \in\mathbb{G}\times\{A,D\}}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, b_s^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_{\mu}[z_{(k,R_k)}(t) z_{(l,R_l)}(t)]\\
&= \lim_{s\to\infty} \Bigg[\sum_{\substack{(k',R_{k'}),(l',R_{l'})\\ \in\mathbb{G}\times\{A,D\}}} b_{t+s}^{(1)}\big((i,R_i),(k',R_{k'})\big)\, b_{t+s}^{(1)}\big((j,R_j),(l',R_{l'})\big)\\
&\qquad\qquad \times \mathbb{E}_{\mu}[z_{(k',R_{k'})} z_{(l',R_{l'})}]\\
&\quad + \int_0^t \mathrm{d}r \sum_{k'\in\mathbb{G}} b_{t-r+s}^{(1)}\big((i,R_i),(k',A)\big)\, b_{t-r+s}^{(1)}\big((j,R_j),(k',A)\big)\, \mathbb{E}_{\mu}[g(x_{k'}(r))] \Bigg].
\end{aligned}
\tag{3.125}
$$

Since $\mu \in \mathcal{R}_\theta^{(1)}$, we are left to show that

$$
\lim_{s\to\infty} \int_s^{t+s} \mathrm{d}u \sum_{k'\in\mathbb{G}} b_u^{(1)}\big((i,R_i),(k',A)\big)\, b_u^{(1)}\big((j,R_j),(k',A)\big)\, \mathbb{E}_{\mu}[g(x_{k'}(t+s-u))] = 0.
\tag{3.126}
$$

Using the notation of Section 3.2.2, we get

$$
\begin{aligned}
&\lim_{s\to\infty} \int_s^{t+s} \mathrm{d}u \sum_{k'\in\mathbb{G}} b_u^{(1)}\big((i,R_i),(k',A)\big)\, b_u^{(1)}\big((j,R_j),(k',A)\big)\, \mathbb{E}_{\mu}[g(x_{k'}(t+s-u))]\\
&\le \|g\| \lim_{s\to\infty} \int_s^{t+s} \mathrm{d}u \sum_{k'\in\mathbb{G}} b_u^{(1)}\big((i,R_i),(k',A)\big)\, b_u^{(1)}\big((j,R_j),(k',A)\big)\\
&= \|g\| \lim_{s\to\infty} \int_s^{t+s} \mathrm{d}u\, \mathbb{E}_{(i,R_i),(j,R_j)} \left[\sum_{k'\in\mathbb{G}} a_{T(u)}(i,k')\, 1_{\mathcal{E}(u)}\, a_{T'(u)}(j,k')\, 1_{\mathcal{E}'(u)} \right]\\
&\le \|g\| \lim_{s\to\infty} \int_s^{t+s} \mathrm{d}u\, \mathbb{E}_{(0,A),(0,A)} \left[\sum_{k'\in\mathbb{G}} a_{T(u)}(i,k')\, 1_{\mathcal{E}(u)}\, a_{T'(u)}(j,k')\, 1_{\mathcal{E}'(u)} \right] = 0,
\end{aligned}
\tag{3.127}
$$

where the last equality follows from the assumption $I_{\hat{a}} < \infty$ in Theorem 2.3.1, (3.60) and (3.79). The last inequality follows from the Markov property and the observation that, in order to get a contribution to the integral, the two random walks first have to meet at the same site and both be active. We conclude that $\mu S_t \in \mathcal{R}_\theta^{(1)}$ for all $t \ge 0$.

To show that $\mu(\infty) \in \mathcal{R}_\theta^{(1)}$, we proceed like in (3.124), to obtain

$$
\lim_{s\to\infty} \mathbb{E}_{\mu(\infty)}[z_{(i,R_i)}(s)] = \lim_{s\to\infty} \lim_{n\to\infty} \mathbb{E}_{\mu S_{t_n}}[z_{(i,R_i)}(s)] = \lim_{s\to\infty} \lim_{n\to\infty} \mathbb{E}_{\mu}[z_{(i,R_i)}(t_n+s)] = \theta,
\tag{3.128}
$$

and so (3.106) is satisfied. To get (3.107), we note that, by Lemma 3.2.1,

$$\begin{aligned}
&\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b^{(1)}_{t_n}\big((i,R_i),(k,R_k)\big)\, b^{(1)}_{t_n}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}] \\
&\le E_\mu[z_{(i,R_i)}(t_n)z_{(j,R_j)}(t_n)] \\
&\le \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b^{(1)}_{t_n}\big((i,R_i),(k,R_k)\big)\, b^{(1)}_{t_n}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}] \\
&\qquad + \|g\|\int_0^{t_n} \mathrm{d}s \sum_{k\in\mathbb{G}} b^{(1)}_{t_n-s}\big((i,R_i),(k,A)\big)\, b^{(1)}_{t_n-s}\big((j,R_j),(k,A)\big).
\end{aligned} \tag{3.129}$$

Letting $n\to\infty$, we see that, since $\mu\in\mathcal{R}^{(1)}_\theta$,

$$\begin{aligned}
\theta^2 &\le \mathbb{E}_{\mu(\infty)}[z_{(i,R_i)}z_{(j,R_j)}] \\
&\le \theta^2 + \|g\|\int_0^\infty \mathrm{d}s \sum_{k\in\mathbb{G}} b^{(1)}_r\big((i,R_i),(k,A)\big)\, b^{(1)}_r\big((j,R_j),(k,A)\big).
\end{aligned} \tag{3.130}$$

Inserting (3.130) into (3.107), we see that it is enough to show that

$$\begin{aligned}
&\lim_{s\to\infty} \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b^{(1)}_s\big((i,R_i),(k,R_k)\big)\, b^{(1)}_s\big((j,R_j),(l,R_l)\big) \\
&\qquad \times\ \|g\|\int_0^\infty \mathrm{d}r \sum_{k'\in\mathbb{G}} b^{(1)}_r((k,R_k),(k',A))\, b^{(1)}_r((l,R_l),(k',A)) \\
&= \lim_{s\to\infty}\|g\|\int_0^\infty \mathrm{d}r \sum_{k'\in\mathbb{G}} b^{(1)}_{r+s}\big((i,R_i),(k',A)\big)\, b^{(1)}_{r+s}\big((j,R_j),(k',A)\big) = 0.
\end{aligned} \tag{3.131}$$

However, from the assumption $I_{\hat{a}}<\infty$ in Theorem 2.3.1, (3.60) and (3.79), we have

$$\begin{aligned}
&\lim_{s\to\infty}\|g\|\int_0^\infty \mathrm{d}r \sum_{k'\in\mathbb{G}} b^{(1)}_{r+s}\big((i,R_i),(k',A)\big)\, b^{(1)}_{r+s}\big((j,R_j),(k',A)\big) \\
&= \lim_{s\to\infty}\|g\|\int_s^\infty \mathrm{d}r\ \mathbb{E}_{(i,R_i),(j,R_j)}\left[\sum_{k'\in\mathbb{G}} a_{T(r)}(i,k')1_{\mathcal{E}(r)}a_{T'(r)}(j,k')1_{\mathcal{E}'(r)}\right] = 0.
\end{aligned} \tag{3.132}$$

□

2. Uniqueness of the equilibrium. In this section we show that, for given θ, the equilibrium when it exists is unique. To prove this we extend the coupling argument in [14]. Consider two copies of the system (2.4)–(2.5) *coupled via their Brownian motions*:

$$\begin{aligned}
\mathrm{d}x^k_i(t) &= \sum_{j\in\mathbb{G}} a(i,j)\,[x^k_j(t)-x^k_i(t)]\,\mathrm{d}t + \sqrt{g(x^k_i(t))}\,\mathrm{d}w_i(t) + Ke\,[y^k_i(t)-x^k_i(t)]\,\mathrm{d}t, \\
\mathrm{d}y^k_i(t) &= e\,[x^k_i(t)-y^k_i(t)]\,\mathrm{d}t, \qquad k\in\{1,2\}.
\end{aligned} \tag{3.133}$$

Here, k labels the copy, and the two copies are driven by the same set of Brownian motions $(w_i(t))_{t\geq 0}$, $i \in \mathbb{G}$. As initial probability distributions we choose $\mu^1(0)$ and $\mu^2(0)$ that are both invariant and ergodic under translations.

Let

$$\bar{z}_i(t) = (z_i^1(t), z_i^2(t)), \qquad z_i^k(t) = (x_i^k(t), y_i^k(t)), \quad k \in \{1,2\}. \tag{3.134}$$

The coupled system $(\bar{z}_i(t))_{i\in\mathbb{G}}$ has a unique strong solution [67, Theorem 3.2] whose marginals are the single-component systems. Write $\hat{\mathbb{P}}$ to denote the law of the coupled system, and let $\Delta_i(t) = x_i^1(t) - x_i^2(t)$ and $\delta_i(t) = y_i^1(t) - y_i^2(t)$.

Lemma 3.2.9 (Coupling dynamics). *For every $t \geq 0$,*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big] = &-2\sum_{j\in\mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\left[|\Delta_j(t)|\,1_{\{\operatorname{sgn}\Delta_i(t)\neq \operatorname{sgn}\Delta_j(t)\}}\right]\\ &- 2Ke\,\hat{\mathbb{E}}\left[\big(|\Delta_i(t)| + |\delta_i(t)|\big)\,1_{\{\operatorname{sgn}\Delta_i(t)\neq \operatorname{sgn}\delta_i(t)\}}\right].\end{aligned} \tag{3.135}$$

Proof. Let $f(x) = |x|$, $x \in \mathbb{R}$. Then $f'(x) = \operatorname{sgn} x$ and $f''(x) = 0$ for $x \neq 0$, but f is not differentiable at $x = 0$, a point the path hits. Therefore, by a generalization of Itô's formula, we have

$$\begin{aligned}\mathrm{d}|\Delta_i(t)| &= \operatorname{sgn}\Delta_i(t)\,\mathrm{d}\Delta_i(t) + \mathrm{d}L_t^0,\\ \mathrm{d}\Delta_i(t) &= \sum_{j\in\mathbb{G}} a(i,j)[\Delta_j(t) - \Delta_i(t)]\,\mathrm{d}t + \left[\sqrt{g(x_i^1(t))} - \sqrt{g(x_i^2(t))}\right]\,\mathrm{d}w_i(t)\\ &\qquad + Ke\left[\delta_i(t) - \Delta_i(t)\right]\mathrm{d}t,\end{aligned} \tag{3.136}$$

where L_t^0 is the local time of $\Delta_i(t)$ at 0 (see [63, Section IV.43]). Next, we use that $\Delta_i(t)$ has zero local time at $x = 0$ because g is Lipschitz (see [63, Proposition V.39.3]). Taking expectation, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\mathbb{E}}\big[|\Delta_i(t)|\big] = \sum_{j\in\mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\Big[\operatorname{sgn}\Delta_i(t)\,[\Delta_j(t) - \Delta_i(t)]\Big] + Ke\,\hat{\mathbb{E}}\Big[\operatorname{sgn}\Delta_i(t)\,[\delta_i(t) - \Delta_i(t)]\Big]. \tag{3.137}$$

Similarly, we have

$$\begin{aligned}\mathrm{d}|\delta_i(t)| &= \operatorname{sgn}\delta_i(t)\,\mathrm{d}\delta_i(t),\\ \mathrm{d}\delta_i(t) &= e\left[\Delta_i(t) - \delta_i(t)\right]\mathrm{d}t.\end{aligned} \tag{3.138}$$

Taking expectation, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\mathbb{E}}\big[|\delta_i(t)|\big] = e\,\hat{\mathbb{E}}\Big[\operatorname{sgn}\delta_i(t)\,[\Delta_i(t) - \delta_i(t)]\Big]. \tag{3.139}$$

Combining (3.137) and (3.139), we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big] = &\sum_{j\in\mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\Big[\operatorname{sgn}\Delta_i(t)\,[\Delta_j(t) - \Delta_i(t)]\Big]\\ &+ K\,e\,\hat{\mathbb{E}}\Big[[\operatorname{sgn}\Delta_i(t) - \operatorname{sgn}\delta_i(t)]\,[\delta_i(t) - \Delta_i(t)]\Big].\end{aligned} \tag{3.140}$$

Note that

$$\operatorname{sgn} \Delta_i(t)\,[\Delta_j(t) - \Delta_i(t)] = |\Delta_j(t)| - |\Delta_i(t)| - 2\,|\Delta_j(t)|\,1_{\{\operatorname{sgn} \Delta_i(t) \neq \operatorname{sgn} \Delta_j(t)\}}. \quad (3.141)$$

By translation invariance, $\mathbb{E}[|\Delta_i(t)|]$ is independent of i. Hence the first sum in the right-hand side can be rewritten as

$$-2 \sum_{j \in \mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\Big[|\Delta_j(t)|\,1_{\{\operatorname{sgn} \Delta_i(t) \neq \operatorname{sgn} \Delta_j(t)\}}\Big]. \quad (3.142)$$

Similarly, the second sum in the right-hand side can be rewritten as

$$-2Ke\,\hat{\mathbb{E}}\left[\big(|\Delta_i(t)| + |\delta_i(t)|\big)\,1_{\{\operatorname{sgn} \Delta_i(t) \,\neq\, \operatorname{sgn} \delta_i(t)\}}\right]. \quad (3.143)$$

Combining (3.140) and (3.142)–(3.143), we get the claim. □

Lemma 3.2.9 tells us that $t \mapsto \hat{\mathbb{E}}[|\Delta_i(t)| + K|\delta_i(t)|]$ is a non-increasing Lyapunov function. Therefore $\lim_{t\to\infty} \hat{\mathbb{E}}[|\Delta_i(t)| + K|\delta_i(t)|] = c_i \in [0, 1+K]$ exists. To show that the coupling is successful we need the following lemma.

Lemma 3.2.10 (Uniqueness of equilibrium). *If $a(\cdot,\cdot)$ is transient, then $c_i = 0$ for all $i \in \mathbb{G}$, and so the coupling is successful, i.e.,*

$$\lim_{t\to\infty} \hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big] = 0. \quad (3.144)$$

Proof. Write $-h_i(t)$ to denote the right-hand side of (3.135). We begin with the observation that $t \mapsto h_i(t)$ has the following properties:

(a) $h_i \geq 0$.
(b) $0 \leq \int_0^\infty \mathrm{d}t\, h_i(t) \leq 1 + K$.
(c) h_i is differentiable with h_i' bounded.

Property (a) is evident. Property (b) follows from integration of (3.135):

$$\int_0^t \mathrm{d}s\, h_i(s) = \hat{\mathbb{E}}\big[|\Delta_i(0)| + K|\delta_i(0)|\big] - \hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big]. \quad (3.145)$$

The proof of Property (c) is given in Appendix A.4. It follows from (a)–(c) that $\lim_{t\to\infty} h(t) = 0$. Hence, for every $\epsilon > 0$,

$$\begin{aligned}
&\forall\, i,j \in \mathbb{G} \text{ with } a(i,j) > 0: \\
&\qquad \lim_{t\to\infty} \hat{\mathbb{P}}\Big(\{\Delta_i(t) < -\epsilon,\, \Delta_j(t) > \epsilon\} \cup \{\Delta_i(t) > \epsilon,\, \Delta_j(t) < -\epsilon\}\Big) = 0, \\
&\forall\, i \in \mathbb{G}: \\
&\qquad \lim_{t\to\infty} \hat{\mathbb{P}}\Big(\{\Delta_i(t) < -\epsilon,\, \delta_i(t) > \epsilon\} \cup \{\Delta_i(t) > \epsilon,\, \delta_i(t) < -\epsilon\}\Big) = 0.
\end{aligned} \quad (3.146)$$

In Appendix A.3 we will prove the following lemma:

Lemma 3.2.11 (Successful coupling). *For all $i,j \in \mathbb{G}$ and $\epsilon > 0$,*

$$\lim_{t\to\infty} \hat{\mathbb{P}}\Big(\{\Delta_i(t) < -\epsilon,\, \Delta_j(t) > \epsilon\} \cup \{\Delta_i(t) > \epsilon,\, \Delta_j(t) < -\epsilon\}\Big) = 0. \quad (3.147)$$

The proof of this lemma relies on the fact that $\hat{a}(\cdot,\cdot)$ is irreducible. Let

$$\begin{aligned} E_0 \times E_0 = &\Big\{\bar{z} \in E \times E\colon\ z^1_{(i,R_i)}(t) \geq z^2_{(i,R_i)}(t)\ \forall (i,R_i) \in \mathbb{G} \times \{A,D\}\Big\} \\ &\cup \Big\{\bar{z} \in E \times E\colon\ z^2_{(i,R_i)}(t) \geq z^1_{(i,R_i)}(t)\ \forall (i,R_i) \in \mathbb{G} \times \{A,D\}\Big\}. \end{aligned} \tag{3.148}$$

Then Lemma 3.2.11 together with (3.146) imply that $\lim_{t\to\infty} \hat{\mathbb{P}}\,(E_0 \times E_0) = 1$, which we express by saying that "one diffusion lies on top of the other".

Using Lemma 3.2.11 we can complete the proof of the successful coupling. Let $t_n \to \infty$ as $n \to \infty$ and suppose, by possibly going to further subsequences, that $\lim_{n\to\infty} \mu^1(t_n) = \nu^1_\theta$ and $\lim_{n\to\infty} \mu^2(t_n) = \nu^2_\theta$. Let $\bar{\nu}_\theta$ be the measure on $E \times E$ given by $\bar{\nu}_\theta = \nu^1_\theta \times \nu^2_\theta$. Using dominated convergence, invoking the preservation of translation invariance, and using the limiting distribution of $b^{(1)}_t(\cdot,\cdot)$ on $\{A,D\}$, we find

$$\begin{aligned} &\int_{E\times E} \mathrm{d}\bar{\nu}_\theta\, |\Delta_i| + K|\delta_i| \\ &\quad= (1+K) \int_{E_0\times E_0} \mathrm{d}\bar{\nu}_\theta \lim_{n\to\infty} \sum_{j\in\mathbb{G}} \Big[b^{(1)}_{t_n}\big((i,R_i),(j,A)\big)\, |x^1_i - x^2_i| \\ &\qquad\qquad\qquad\qquad\qquad\qquad + b^{(1)}_{t_n}\big((i,R_i),(j,D)\big)\, |y^1_i - y^2_i| \Big] \\ &\quad= \lim_{n\to\infty} (1+K) \int_{E_0\times E_0} \mathrm{d}\bar{\nu}_\theta \left| \sum_{j\in\mathbb{G}\times\{A,D\}} b^{(1)}_{t_n}\big((i,R_i),(j,R_j)\big)\, (z^1_{(j,R_j)} - z^2_{(j,R_j)}) \right| \\ &\quad\leq \lim_{n\to\infty} (1+K) \int_{E} \mathrm{d}\nu^1_\theta \left| \sum_{j\in\mathbb{G}\times\{A,D\}} b^{(1)}_{t_n}\big((i,R_i),(j,R_j)\big)\, z^1_{(j,R_j)} - \theta \right| \\ &\qquad + \lim_{n\to\infty} (1+K) \int_{E} \mathrm{d}\nu^2_\theta \left| \sum_{i\in\mathbb{G}\times\{A,D\}} b^{(1)}_{t_n}\big((i,R_i),(j,R_j)\big)\, z^2_{(j,R_j)} - \theta \right| = 0. \end{aligned} \tag{3.149}$$

Here, the last equality follows because both ν^1_θ and ν^2_θ are in $\mathcal{R}^{(1)}_\theta$ by Lemma 3.2.8. Thus, we see that $\bar{\nu}_\theta$ concentrates on the diagonal. Suppose now that there exists a sequence $(t_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \mathbb{E}[|\Delta_i(t_n)| + K|\delta_i(t_n)|] = \delta > 0$. Since $\{\mathcal{L}(\bar{Z}(t_n))\}_{n\in\mathbb{N}}$ is tight (recall (3.134)), by Prokhorov's theorem there exists a converging subsequence $\{\mathcal{L}(\bar{Z}(t_{n_k}))\}_{k\in\mathbb{N}}$. Let $\bar{\nu}_\theta$ denote the limiting measure. Then, by Lemma 3.2.8 and (3.149),

$$\delta = \lim_{k\to\infty} \mathbb{E}[|\Delta_i(t_{n_k})| + K|\delta_i(t_{n_k})|] = \int_{E\times E} \mathrm{d}\bar{\nu}_\theta\, [|\Delta_i| + K|\delta_i|] = 0. \tag{3.150}$$

Thus, $\lim_{t\to\infty} \mathbb{E}[|\Delta_i(t)| + K|\delta_i(t)|] = 0$, and we conclude that the coupling is successful. Hence, given the initial average density θ in (2.62), the equilibrium measure is unique if it exists. □

3. Stationarity of ν_θ and convergence to ν_θ.

Lemma 3.2.12 (Existence of equilibrium). *Let $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$. Then $\lim_{t\to\infty} \mu(t) = \nu_\theta$ for some invariant measure ν_θ.*

Proof. To prove that the limit is an invariant measure, suppose that $\mu(0) = \mu = \delta_\theta$. Since the state space of $(Z(t))_{t\geq 0}$ is compact, each sequence $\{\mathcal{L}(Z(t_n))\}_{n\in\mathbb{N}}$ is tight. Hence, by Prokhorov's theorem, there exists a converging subsequence such that $\lim_{n\to\infty} \delta_\theta S_{t_n} = \nu_\theta$. Since $\delta_\theta \in \mathcal{R}_\theta^{(1)}$, Lemma 3.2.8 tells us that $\lim_{n\to\infty} \delta_\theta S_{t_n} \in \mathcal{R}_\theta^{(1)}$. To prove that ν_θ is invariant, fix any $s_0 \geq 0$. Let $\mu = \delta_\theta S_{s_0}$. Then, by Lemma 3.2.8, $\mu \in \mathcal{R}_\theta^{(1)}$ and, by Lemma 3.2.11, we can find a further subsequence such that $\lim_{k\to\infty} \mu(t_{n_k}) = \nu_\theta$. By the Feller property of the SSDE in (2.4)–(2.5), we obtain

$$\nu_\theta S_{s_0} = \lim_{n\to\infty} \delta_\theta S_{t_n} S_{s_0} = \lim_{k\to\infty} \delta_\theta S_{s_0} S_{t_{n_k}} = \lim_{k\to\infty} \mu S_{t_{n_k}} = \nu_\theta. \tag{3.151}$$

Hence, ν_θ is indeed an invariant measure.

To prove the convergence of $\mu(t)$ to ν_θ, note that $\nu_\theta \in \mathcal{R}_\theta^{(1)}$ by Lemma 3.2.8. Let $\nu = \nu_\theta$. Then, by the invariance of ν_θ, we have $\lim_{t\to\infty} \nu S_t = \nu_\theta$. By Lemma 3.2.10, we have $\lim_{t\to\infty} \mu S_t = \lim_{t\to\infty} \nu S_t = \nu_\theta$ for all $\mu \in \mathcal{R}_\theta^{(1)}$. □

4. Ergodicity, mixing and associatedness.

Lemma 3.2.13 (Properties of equilibrium). *Let $\mu(0) \in \mathcal{R}_\theta^{(1)}$ be ergodic under translations. Then $\nu_\theta = \lim_{t\to\infty} \mu(t)$ is ergodic and mixing under translations, and is associated.*

Proof. After a standard approximation argument, [46, Corollary 1.5 and subsequent discussion] implies that associatedness is preserved over time. Note that δ_θ is an associated measure and lies in $\mathcal{R}_\theta^{(1)}$. Hence, by Lemma 3.2.12, $\nu_\theta = \lim_{t\to\infty} \delta_\theta S_t$ and therefore ν_θ is associated.

We prove the ergodicity of ν_θ by showing that the random field of components is mixing. To prove that ν_θ is mixing, we use associatedness and decay of correlations. Let $B, B' \subset \mathbb{G}$ be finite, and let c_j, d_i be positive constants for $j \in B$, $i \in B'$. For $k \in \mathbb{G}$, define the random variables

$$Y_0 = \sum_{j\in B} c_j z_{(j,R_j)}, \qquad Y_k = \sum_{i\in B'} d_i z_{(i+k,R_{i+k})}. \tag{3.152}$$

Note that Y_0 and Y_k are associated under ν_θ because $(z_{(i,R_i)})_{(i,R_i)\in\mathbb{G}\times\{A,D\}}$ are associated. Therefore, by [61, Eq.(2.2)], it follows that for $s,t \in \mathbb{R}$,

$$\left|\mathbb{E}_{\nu_\theta}[e^{\mathrm{i}(sY_0+tY_n)}] - \mathbb{E}_{\nu_\theta}[e^{\mathrm{i}sY_0}]\,\mathbb{E}_{\nu_\theta}[e^{\mathrm{i}tY_n}]\right| \leq |st| \operatorname{Cov}_{\nu_\theta}(Y_0, Y_n). \tag{3.153}$$

Since $\mu \in \mathcal{R}_\theta^{(1)}$ by Lemma 3.2.1,

$$\begin{aligned}\mathrm{Cov}_{\nu_\theta}(Y_0,Y_k) =&\sum_{j\in B}\sum_{i\in B'} c_j d_i \lim_{t\to\infty}\mathrm{Cov}_\mu(z_{(j,R_j)}(t), z_{(i+k,R_{i+k})}(t))\\ \leq&\|g\|\sum_{j\in B}\sum_{i\in B'} c_j d_i\\ &\times\int_0^\infty \mathrm{d}r \sum_{(l,R_l)\in\mathbb{G}\times\{A\}} b_r^{(1)}\big((j,R_j),(l,A)\big)\, b_r^{(1)}\big((i+k,R_{i+k}),(l,A)\big).\end{aligned} \tag{3.154}$$

The last integral gives the expected total time for two partition elements in the dual, starting in (j,R_j) and $(i+k,R_{i+k})$, to be active at the same site. To show that this integral converges to 0 as $\|k\|\to\infty$, we rewrite the sum as (recall (3.64)–(3.65))

$$\begin{aligned}&\mathbb{E}_{(i+k,R_{i+k}),(j,R_j)}\left[\left(\sum_{l\in\mathbb{G}} a_{T(r)}(j,l)\, a_{T'(r)}(i+k,l)\right) 1_{\mathcal{E}(r)}\, 1_{\mathcal{E}'(r)}\right]\\ &= \mathbb{E}_{(i+k,R_{i+k}),(j,R_j)}\left[\left(\sum_{l'\in\mathbb{G}} \hat{a}_{2M(r)}(i+k-j,l')\, a_{\Delta(r)}(l',0)\right) 1_{\mathcal{E}(r)}\, 1_{\mathcal{E}'(r)}\right]\\ &\leq \mathbb{E}_{(i+k,R_{i+k}),(j,R_j)}\left[\left(\sum_{l'\in\mathbb{G}} \hat{a}_{2M(r)}(i+k-j,l')\left[a_{\Delta(r)}(l',0)+a_{\Delta(r)}(0,l')\right]\right) 1_{\mathcal{E}(r)}\, 1_{\mathcal{E}'(r)}\right]\\ &= \mathbb{E}_{(i+k,R_{i+k}),(j,R_j)}\left[\hat{a}_{2M(r)+2\Delta(r)}(i+k-j,0)1_{\mathcal{E}(r)}\, 1_{\mathcal{E}'(r)}\right].\end{aligned} \tag{3.155}$$

Because $\hat{a}(\cdot,\cdot)$ is symmetric, we have $\hat{a}_{2M(r)+2\Delta(r)}(i+k-j,0)\leq \hat{a}_{2M(r)+2\Delta(r)}(0,0)$. Since

$$T(t)+T'(t)\leq 2M(r)+2\Delta(r)\leq 2\,(T(t)+T'(t))\,, \tag{3.156}$$

and the Fourier transform in (3.74)–(3.75) implies that

$$\int_0^\infty \mathrm{d}r\, \mathbb{E}_{(i+k,R_{i+k}),(j,R_j)}\left[\hat{a}_{2M(r)+2\Delta(r)}(0,0)1_{\mathcal{E}(r)}\, 1_{\mathcal{E}'(r)}\right]<\infty. \tag{3.157}$$

if and only if $I_{\hat{a}}<\infty$. Since we are in the transient regime, i.e., $I_{\hat{a}}<\infty$, we can use dominated convergence, in combination with the fact that $\lim_{\|k\|\to\infty}\hat{a}_t(i+k-j,0)=0$ for all i,j,t, to conclude that $\lim_{\|k\|\to\infty}\mathrm{Cov}_{\nu_\theta}(Y_0,Y_k)=0$. □

§3.2.4 Proof of the dichotomy

Theorem 2.3.1(a) follows from Lemmas 3.2.7, 3.2.12 and 3.2.13. The equality $\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_0]=\theta$ follows from the evolution equations in (2.4)–(2.5), the fact that ν_θ is an equilibrium measure, and the preservation of θ (see (2.63)). Theorem 2.3.1(b) follows from Lemma 3.2.5.

§3.3 Proofs: Long-time behaviour for Model 2

In Sections 3.3.1–3.3.4 we show that the results proved in Sections 3.2.1–3.2.4 carry over from model 1 to model 2. In Section 3.3.5 we show that symmetry of $a(\cdot,\cdot)$

is needed. In Section 3.3.6 we show what happens when for infinite seed-bank the fat-tailed wake-up time is modulated by a slowly varying function.

§3.3.1 Moment relations

Like in model 1, we start by relating the first and second moments of the system in (2.12)–(2.13) to the random walk that evolves according to the transition kernel $b^{(2)}(\cdot,\cdot)$ on $\mathbb{G} \times \{A,(D_m)_{m\in\mathbb{N}_0}\}$ given by (2.41). Also here these moment relations hold for all $g \in \mathcal{G}$. Moreover these moment relations holds for $\rho < \infty$ as well as for $\rho = \infty$. Below we write $\mathbb{E}_z$ for $\mathbb{E}_{\delta_z}$, the expectation when the process starts from the initial measure δ_z, $z \in E$.

Lemma 3.3.1 (First and second moment). *For $z \in E$, $t \geq 0$ and $(i,R_i),(j,R_j) \in \mathbb{G} \times \{A,(D_m)_{m\in\mathbb{N}_0}\}$,*

$$\mathbb{E}_z[z_{(i,R_i)}(t)] = \sum_{\substack{(k,R_k)\\ \in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\, z_{(k,R_k)} \tag{3.158}$$

and

$$\begin{aligned} &\mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\ &= \sum_{\substack{(k,R_k),(l,R_l)\\ \in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\, b_t^{(2)}\big((j,R_j),(l,R_l)\big)\, z_{(k,R_k)} z_{(l,R_l)} \\ &\quad + \int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{(t-s)}^{(2)}((i,R_i),(k,A))\, b_{(t-s)}^{(2)}((j,R_j),(k,A))\, \mathbb{E}_z[g(x_k(s))]. \end{aligned} \tag{3.159}$$

Proof. The proof follows from that of Lemma 3.2.1 after we replace $b^{(1)}(\cdot,\cdot)$ by $b^{(2)}(\cdot,\cdot)$ and use (2.12)–(2.13) instead of (2.4)–(2.5). □

Remark 3.3.2 (Density). From Lemma 3.3.1 we obtain that if μ is invariant under translations with $\mathbb{E}_\mu[x_0(0)] = \theta_x$ and $\mathbb{E}_\mu[y_{0,m}(0)] = \theta_{y,m}$ for all $m \in \mathbb{N}_0$, then

$$\begin{aligned} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = {}&\theta_x \sum_{(k,R_k)\in\mathbb{G}\times\{A\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big) \\ &+ \sum_{m\in\mathbb{N}_0} \theta_{y,m} \sum_{(k,R_k)\in\mathbb{G}\times\{D_m\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big) \end{aligned} \tag{3.160}$$

and

$$\begin{aligned} &\mathbb{E}_\mu[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\ &= \sum_{\substack{(k,R_k),(l,R_l)\\ \in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\, b_t^{(2)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_\mu[z_{(k,R_k)} z_{(l,R_l)}] \\ &\quad + \int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{t-s}^{(2)}\big((i,R_i),(k,A)\big)\, b_{t-s}^{(2)}\big((j,R_j),(k,A)\big)\, \mathbb{E}_\mu[g(x_i(s))]. \end{aligned} \tag{3.161}$$

- For $\rho < \infty$, the kernel $b^{(2)}(\cdot,\cdot)$ projected on the second component (= the seed-bank) corresponds to recurrent Markov chain. Therefore, by translation invariance in the first component, we have

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \frac{\theta_x + \sum_{m\in\mathbb{N}_0} K_m\theta_{y,m}}{1+\sum_{m\in\mathbb{N}_0} K_m} = \theta. \tag{3.162}$$

- For $\rho = \infty$, the kernel $b^{(2)}(\cdot,\cdot)$ viewed as a kernel on $\{A,(D_m)_{m\in\mathbb{N}_0}\}$ corresponds to a null-recurrent Markov chain. Hence, for all (i,R_i) and all $D_m,\ m\in\mathbb{N}_0$,

$$\lim_{t\to\infty}\sum_{k\in\mathbb{G}} b_t^{(2)}((i,R_i),(k,D_m)) = 0. \tag{3.163}$$

Since for $\rho=\infty$ we assume not only that $\mu\in\mathcal{T}_\theta^{\mathrm{erg}}$ but also that μ is colour regular, it follows that, for all $M\in\mathbb{N}_0$,

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}_\mu[z_{(i,R_i)}(t)] &= \lim_{t\to\infty}\theta_x\sum_{k\in\mathbb{G}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\\ &\quad+\sum_{m\in\mathbb{N}_0}\theta_{y,m}\sum_{(k,R_k)\in\mathbb{G}\times\{D_m\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\\ &=\lim_{t\to\infty}\sum_{m=M}^{\infty}\theta_{y,m}\sum_{(k,R_k)\in\mathbb{G}\times\{D_m\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big).\end{aligned} \tag{3.164}$$

Therefore

$$\lim_{t\to\infty}\mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta. \tag{3.165}$$

□

§3.3.2 The clustering case

In this section we prove convergence to a trivial equilibrium when $\rho<\infty$ and $I_{\hat a}=\infty$ and when $\rho=\infty$ and $I_{\hat a,\gamma}=\infty$. The proof follows along the same lines as in Section 3.2.2. Therefore we again first consider $g = dg_{\mathrm{FW}}$, and subsequently use a duality comparison argument to show that the results hold for $g\neq dg_{\mathrm{FW}}$ as well.

Case $g = dg_{\mathrm{FW}}$. We start by proving the equivalent of Lemma 3.2.4, which is Lemma 3.3.3 below.

Lemma 3.3.3 (Clustering). *Suppose that $\mu(0)\in\mathcal{T}_\theta^{\mathrm{erg}}$ and $g = dg_{\mathrm{FW}}$. Let $\mu(t)$ be the law at time t of the system defined in* (2.12)–(2.13). *Then the following two statements hold:*

- *If $\rho<\infty$ and $I_{\hat a}=\infty$, i.e., $\hat a(\cdot,\cdot)$ is recurrent, then*

$$\lim_{t\to\infty}\mu(t) = \theta\,[\delta_{(1,1^{\mathbb{N}_0})}]^{\otimes\mathbb{G}} + (1-\theta)\,[\delta_{(0,0^{\mathbb{N}_0})}]^{\otimes\mathbb{G}}. \tag{3.166}$$

- *If $\rho = \infty$ and $I_{\hat{a},\gamma} = \infty$ then*

$$\lim_{t\to\infty} \mu(t) = \theta\,[\delta_{(1,1^{\mathbb{N}_0})}]^{\otimes\mathbb{G}} + (1-\theta)\,[\delta_{(0,0^{\mathbb{N}_0})}]^{\otimes\mathbb{G}}. \tag{3.167}$$

Proof. We distinguish between $\rho < \infty$ and $\rho = \infty$, which exhibit different behaviour.

Case $\rho < \infty$. The same dichotomy as for model 1 holds when the average wake-up time is finite (recall (2.20)–(2.21), (2.50)). Indeed, the argument in (3.69)–(3.79) can be copied with Ke, e replaced by $\chi, \chi/\rho$ and A, B by $\chi/(1+\rho), 1/(1+\rho)$. Under the *symmetry assumption* in (2.73) we have $\tilde{a}(\phi) = 0$. Hence only the law of large numbers in (3.70) is needed, not the central limit theorem in (3.72), which may fail (see Section 3.3.5).

Case $\rho = \infty$. When the average wake-up time is infinite, we need the assumptions in (2.60) and (2.76). By the standard law of large numbers for stable random variables (see e.g. [34, Section XIII.6]), we have

$$\lim_{k\to\infty} \frac{1}{k}\sum_{\ell=1}^{k} \sigma_\ell = \frac{1}{\chi} \quad \mathbb{P}\text{-a.s.}, \qquad \lim_{k\to\infty} \frac{1}{k^{1/\gamma}}\sum_{\ell=1}^{k} \tau_\ell = W \quad \text{in } \mathbb{P}\text{-probability}, \tag{3.168}$$

with W a stable law random variable on $(0,\infty)$ with exponent γ. Therefore

$$\begin{aligned}
&\lim_{t\to\infty} \frac{1}{t^\gamma} N(t) = \lim_{t\to\infty} \frac{1}{t^\gamma} N'(t) = W^{-\gamma} \quad \text{in } \mathbb{P}\text{-probability},\\
&\lim_{t\to\infty} \frac{1}{t^\gamma} T(t) = \lim_{t\to\infty} \frac{1}{t^\gamma} T'(t) = \frac{1}{\chi} W^{-\gamma} \quad \text{in } \mathbb{P}\text{-probability},\\
&\lim_{t\to\infty} t^{1-\gamma}\,\mathbb{P}\big(\mathcal{E}(t)\big) = \lim_{t\to\infty} t^{1-\gamma}\,\mathbb{P}\big(\mathcal{E}'(t)\big) = \frac{1}{\chi}\,\mathbb{E}[W^{-\gamma}], \qquad t\to\infty.
\end{aligned} \tag{3.169}$$

For the last statement to make sense, we must check the following.

Lemma 3.3.4 (Finite limits). $\mathbb{E}[W^{-\gamma}] < \infty$.

Proof. Let $W_k = k^{-1/\gamma}\sum_{l=1}^{k}\tau_l$. Then $W_k^{-\gamma} \le k(\max_{1\le i\le k}\tau_i^\gamma)^{-1}$ and, since τ_i are i.i.d. random variables,

$$\mathbb{E}[W_k^{-\gamma}] \le \int_0^\infty \mathrm{d}x\,\mathbb{P}\left(k\left(\max_{1\le i\le k}\tau_i^\gamma\right)^{-1} > x\right) = \int_0^\infty \mathrm{d}x\,\mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k. \tag{3.170}$$

To estimate the integral in the right-hand side of (3.170), we introduce three constants, T, C_1, C_2. Let $\epsilon \in (0,1)$ and choose $T \in \mathbb{R}_+$ such that, for all $t > T$,

$$|[\mathbb{P}(\tau > t)/(Ct^{-\gamma})] - 1| < \epsilon \tag{3.171}$$

Since $\mathbb{P}(\tau \le t) = 1 - \chi^{-1}\sum_{m\in\mathbb{N}_0} K_m e_m\,\mathrm{e}^{-e_m t}$, we note that, under assumption (2.76), τ admits a continuous bounded density. Hence there exists a $C_1 \in \mathbb{R}_+$ such that $\mathbb{P}(\tau \le t) < C_1 t$. Finally, choose $C_2 \in \mathbb{R}_+$ such that $C_2 > \max(1, C_1^\gamma)$. Split

$$\begin{aligned}\int_0^\infty \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k = &\int_0^{k/T} \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k \\ &+ \int_{k/T}^{kC_2} \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k + \int_{kC_2}^\infty \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k.\end{aligned} \tag{3.172}$$

We estimate each of the three integrals separately. For the first integral, we use the estimate $(1-\mathbb{P}(\tau_1^\gamma \ge \frac{k}{x}))^k \le \exp[-k\mathbb{P}(\tau_1^\gamma \ge \frac{k}{x})]$ to obtain

$$\begin{aligned}\int_0^{k/T} \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k &= \int_0^{k/T} \mathrm{d}x \exp\left[-k\, \mathbb{P}\left(\tau_1 \ge \left(\tfrac{k}{x}\right)^{1/\gamma}\right)\right] \\ &\le \int_0^{k/T} \mathrm{d}x\, \mathrm{e}^{-(1-\epsilon)Cx} \\ &\le \frac{1}{(1-\epsilon)C}.\end{aligned} \tag{3.173}$$

For the second integral, we note that $t \mapsto t\, \mathbb{P}(\tau_1^\gamma > t)$ is a continuous function on $[\frac{1}{C_2}, T]$, and hence attains a minimum value $C_3 \in \mathbb{R}_+$ on $[\frac{1}{C_2}, T]$. Therefore

$$\begin{aligned}\int_{k/T}^{kC_2} \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k &= \int_{k/T}^{kC_2} \mathrm{d}x\, \left[1-\mathbb{P}\left(\tau_1^\gamma \ge \tfrac{k}{x}\right)\right]^k \\ &\le \int_{k/T}^{kC_2} \mathrm{d}x \exp\left[-x\left(\frac{k}{x}\mathbb{P}(\tau_1^\gamma \ge \tfrac{k}{x})\right)\right] \\ &\le \frac{1}{C_3}.\end{aligned} \tag{3.174}$$

For the third integral, we compute

$$\int_{kC_2}^\infty \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k \le \int_{kC_2}^\infty \mathrm{d}x\, \left(C_1^\gamma \tfrac{k}{x}\right)^{\frac{k}{\gamma}} = \int_0^{1/C_2} \mathrm{d}v\, \tfrac{k}{v^2}(C_1^\gamma v)^{\frac{k}{\gamma}} = \frac{C_1^\gamma k}{\frac{k}{\gamma}-1}\left(\frac{C_1^\gamma}{C_2}\right)^{\frac{k}{\gamma}-1}, \tag{3.175}$$

where in the first equality we substitute $v = \frac{k}{x}$. Since $C_2 > C_1^\gamma$, we see that the right-hand side tends to zero as $k \to \infty$. Hence

$$\mathbb{E}[W_k^{-\gamma}] \le \frac{1}{(1-\epsilon)(C/\gamma)} + \frac{1}{C_3} + \frac{C_1^\gamma k}{\frac{k}{\gamma}-1}\left(\frac{C_1^\gamma}{C_2}\right)^{k-1}, \tag{3.176}$$

and by dominated convergence it follows that $\mathbb{E}[W^{-\gamma}] = \lim_{k\to\infty} \mathbb{E}[W_k^{-\gamma}] < \infty$. □

By (2.73), we have $\hat{a}(\phi) = a(\phi)$ and $\tilde{a}(\phi) = 0$ in (3.66), and so (3.74) becomes, with the help of (3.169),

$$\mathbb{E}_{(0,A),(0,A)}\left[\left(\sum_{i\in\mathbb{G}} a_{T(t)}(0,i)\, a_{T'(t)}(0,i)\right) 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)}\right] \asymp t^{-2(1-\gamma)} f(t), \qquad t \to \infty, \tag{3.177}$$

with (recall (3.68))

$$f(t) = \hat{a}_{ct^\gamma}(0,0) \tag{3.178}$$

for some $c \in (0,\infty)$. Here we use that deviations of $T(t)/t^\gamma$ and $T'(t)/t^\gamma$ away from order 1 are stretched exponentially costly in t [31], and therefore are negligible. Since $t \mapsto \hat{a}_t(0,0)$ is regularly varying at infinity (recall (2.60)), it follows that

$$\hat{a}_{ct^\gamma}(0,0) \asymp \hat{a}_{t^\gamma}(0,0), \qquad t \to \infty. \tag{3.179}$$

Combining (3.63) and (3.177)–(3.179), we get

$$I = \infty \quad \Longleftrightarrow \quad I_{\hat{a},\gamma} = \infty \tag{3.180}$$

with $I_{\hat{a},\gamma} = \int_1^\infty \mathrm{d}t\, t^{-2(1-\gamma)}\, \hat{a}_{t^\gamma}(0,0)$. Putting $s = t^\gamma$, we have

$$I_{\hat{a},\gamma} = \int_1^\infty \mathrm{d}s\, s^{-(1-\gamma)/\gamma} \hat{a}_s(0,0), \tag{3.181}$$

which is precisely the integral defined in (2.80). □

Case $g \neq dg_{\mathrm{FW}}$. To prove that the dichotomy criterion of Lemma 3.3.3 holds for general $g \in \mathcal{G}$ we need the equivalent of Lemma 3.2.5. Replacing (2.4)–(2.5) by (2.12)–(2.13), replacing $b^{(1)}$ by $b^{(2)}$ in the proof of Lemma 3.2.5, and using the moment relations in Lemma 3.3.1 instead of the moment relations in Lemma 3.2.1, we see that Lemma 3.3.3 also holds for $g \in \mathcal{G}$.

§3.3.3 The coexistence case

In this section we prove the coexistence results stated in Theorem 2.3.3. Like for model 1 the proofs hold for general $g \in \mathcal{G}$ and we need not distinguish between $g = dg_{\mathrm{FW}}$ and $g \neq dg_{\mathrm{FW}}$. For $\rho < \infty$, the argument is given in Section 3.3.3 and proceeds as in Section 3.2.3. It is organised along the same 4 Steps as the argument for model 1, plus an extra Step 5 that settles the statement in (2.87). For $\rho = \infty$, the argument is given in Section 3.3.3 and is also organised along 5 Steps, but structured differently. In Step 1 we define a set of measures that is preserved under the evolution. In Step 2 we use a coupling argument to show the existence of invariant measures. In Step 3 we show that these invariant measures have vanishing covariances in the seed-bank direction. In Step 4 we use the vanishing covariances to show uniqueness of the invariant measure by coupling. Finally, in Step 5 we show that the unique equilibrium measure is invariant, ergodic and mixing under translations, and is associated.

- **Proof of coexistence for finite seed-bank**

1. Properties of measures preserved under the evolution. For model 2 with $\rho < \infty$, the class of preserved measures is equivalent to $\mathcal{R}_\theta^{(1)}$ for model 1 and is now defined as follows.

Definition 3.3.5 (Preserved class of measures). Let $\mathcal{R}_\theta^{(2)}$ denote the set of measures $\mu \in \mathcal{T}$ satisfying, for all $(i,R_i),(j,R_j) \in \mathbb{G} \times \{A,(D_m)_{m\in\mathbb{N}_0}\}$,

(1)

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta, \tag{3.182}$$

(2)

$$\begin{aligned}\lim_{t\to\infty} \sum_{\substack{(k,R_k),(l,R_l)\\ \in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}}} & b_t^{(2)}\big((i,R_i),(k,R_k)\big)\, b_t^{(2)}\big((j,R_j),(l,R_l)\big)\\ & \times \mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}] = \theta^2.\end{aligned} \tag{3.183}$$

□

Like for model 1, properties (1) and (2) of Definition 3.3.5 hold if and only if

$$\begin{aligned}&\lim_{t\to\infty} \mathbb{E}_\mu\left[\left(\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}} b_t^{(2)}((0,A),(k,R_k))\, z_{(k,R_k)} - \theta\right)^2\right] = 0\\ &\text{for some } (i,R_i)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}.\end{aligned} \tag{3.184}$$

Also for model 2 with $\rho<\infty$ we have $\mathcal{T}_\theta^{\text{erg}} \subset \mathcal{R}_\theta^{(1)}$. To see why, note for all $t>0$ and $m\in\mathbb{N}_0$, $(x_i(t))_{i\in\mathbb{G}}$ and $(y_{i,m}(t))_{i\in\mathbb{G}}$ still are stationary time series. Hence with the help of the Herglotz theorem we can define spectral measures λ_A, λ_{D_m} for $m\in\mathbb{N}_0$ as in (3.112). Let $(RW_t)_{t\geq 0}$ be the random walk evolving according to $b^{(2)}(\cdot,\cdot)$. Introduce the sets

$$\begin{aligned}\mathcal{E}(t) &= \big\{\text{at time } t \text{ the random walk is active}\big\},\\ \mathcal{E}_m(t) &= \big\{\text{at time } t \text{ the random walk is dormant with colour } m\big\}.\end{aligned} \tag{3.185}$$

Note that

$$\begin{aligned}&\lim_{t\to\infty} \mathbb{E}_\mu\left[\left(\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}} b_t^{(2)}((0,A),(k,R_k))\, z_{(k,R_k)} - \theta\right)^2\right]\\ &\leq \lim_{t\to\infty} \mathbb{P}_{(0,A)}(\mathcal{E}(t))\, \mathbb{E}_\mu\left[\left(\sum_{k\in Z^d} \frac{b_t^{(2)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\, x_k - \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\frac{\theta_x}{1+\rho}\right)^2\right]\\ &+ \sum_{m\in\mathbb{N}_0} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t))\, \mathbb{E}_\mu\left[\left(\sum_{k\in Z^d} \frac{b_t^{(2)}((0,A),(k,(D_m))}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))}\, y_{k,m} - \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))}\frac{K_m\theta_{y,m}}{1+\rho}\right)^2\right].\end{aligned} \tag{3.186}$$

Hence we can use the same argument as in the proof of Lemma 3.2.7 to show that $\mathcal{T}_\theta^{\text{erg}} \subset \mathcal{R}_\theta^{(2)}$.

Also Lemma 3.2.8 carries over after we replace $b^{(1)}(\cdot,\cdot)$ by $b^{(2)}(\cdot,\cdot)$ and $\mathcal{R}_\theta^{(1)}$ by $\mathcal{R}_\theta^{(2)}$, as defined in (3.3.5).

2. Uniqueness of the equilibrium. To prove uniqueness of the equilibrium for given θ, we use a similar coupling as for model 1 in Section 3.2.3 in Step 3. Consider two copies of the system in (2.12)–(2.13) *coupled via their Brownian motions*:

$$\begin{aligned} \mathrm{d}x_i^k(t) &= \sum_{j\in\mathbb{G}} a(i,j)\left[x_j^k(t)-x_i^k(t)\right]\mathrm{d}t+\sqrt{g(x_i^k(t))}\,\mathrm{d}w_i(t) && (3.187)\\ &\qquad + \sum_{m\in\mathbb{N}_0} K_m e_m\left[y_{i,m}^k(t)-x_i^k(t)\right]\mathrm{d}t, && (3.188)\\ \mathrm{d}y_{i,m}^k(t) &= e_m\left[x_i^k(t)-y_{i,m}^k(t)\right]\mathrm{d}t, \qquad m\in\mathbb{N}_0, \qquad k\in\{1,2\}. && (3.189)\end{aligned}$$

Here, k labels the copy, and the two copies are driven by the same Brownian motions $(w_i(t))_{t\geq 0}$, $i\in\mathbb{G}$. As initial measures we choose $\mu^1(0),\mu^2(0)\in\mathcal{T}_\theta^{\mathrm{erg}}$.

Let

$$\bar{z}_i(t)=\left(z_i^1(t),z_i^2(t)\right), \qquad z_i^k(t)=\left(x_i^k(t),(y_{i,m}^k(t))_{m\in\mathbb{N}_0}\right), \quad k\in\{1,2\}. \tag{3.190}$$

By [67, Theorem 3.2], the coupled system $(\bar{z}_i(t))_{i\in\mathbb{G}}$ has a unique strong solution whose marginals are the single-component systems. Write $\hat{\mathbb{P}}$ to denote the law of the coupled system, and let $\Delta_i(t)=x_i^1(t)-x_i^2(t)$ and $\delta_{i,m}(t)=y_{i,m}^1(t)-y_{i,m}^2(t)$, $m\in\mathbb{N}_0$. The analogue of Lemma 3.2.9 reads:

Lemma 3.3.6 (Coupling dynamics $\rho<\infty$). *For every $t\geq 0$,*

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\mathbb{E}}\left[|\Delta_i(t)|+\sum_{m\in\mathbb{N}_0}K_m|\delta_i(t)|\right]\\ &\quad = -2\sum_{j\in\mathbb{G}}a(i,j)\,\hat{\mathbb{E}}\left[|\Delta_j(t)|\,1_{\{\mathrm{sgn}\,\Delta_i(t)\neq\mathrm{sgn}\,\Delta_j(t)\}}\right]\\ &\qquad\qquad - 2\sum_{m\in\mathbb{N}_0}K_m e_m\,\hat{\mathbb{E}}\left[\left(|\Delta_i(t)|+|\delta_{i,m}(t)|\right)1_{\{\mathrm{sgn}\,\Delta_i(t)\neq\mathrm{sgn}\,\delta_{i,m}(t)\}}\right]. \end{aligned} \tag{3.191}$$

Proof. Note that the left-hand side of (3.191) is well defined because $\rho<\infty$. The proof of Lemma 3.3.6 carries over from that of Lemma 3.2.9 after replacing (2.4)–(2.5) by (2.12)–(2.13). □

The analogue of Lemma 3.2.10 reads as follows.

Lemma 3.3.7 (Successful coupling $\rho<\infty$). *If $a(\cdot,\cdot)$ is transient, then the coupling is successful, i.e.,*

$$\lim_{t\to\infty}\hat{\mathbb{E}}\big[|\Delta_i(t)|+\sum_{m\in\mathbb{N}_0}K_m|\delta_{i,m}(t)|\big]=0, \qquad \forall i\in\mathbb{G}. \tag{3.192}$$

Proof. This follows in the same way as in the proof of Lemma 3.2.10, by defining $-h_i(t)$ as in the right-hand side of (3.191). Using that the second line of (3.146) now holds for $\delta_{i,m}(t)$ and all $m\in\mathbb{N}_0$, we can finish the proof after replacing $b_t^{(1)}(\cdot,\cdot)$ in (3.149) by $b_t^{(2)}(\cdot,\cdot)$ and summing over the seed-banks D_m, $m\in\mathbb{N}_0$. □

3. Stationarity of the equilibrium ν_θ and convergence to ν_θ. Lemma 3.2.12 holds also for $\mu \in \mathcal{R}_\theta^{(2)}$. This follows after replacing $\mu \in \mathcal{R}_\theta^{(1)}$ by $\mu \in \mathcal{R}_\theta^{(2)}$ in the proof of Lemma 3.2.12, using the equivalent of Lemma 3.2.8 and invoking Lemma 3.3.7 instead of Lemma 3.2.10.

4. Ergodicity, mixing and associatedness. Also Lemma 3.2.13 holds, after replacing $b^{(1)}(\cdot,\cdot)$ by $b^{(2)}(\cdot,\cdot)$. The proof even simplifies, since we can invoke the symmetry of $a(\cdot,\cdot)$ in (3.155).

5. Variances under the equilibrium measure ν_θ. If $\limsup_{m\to\infty} e_m = 0$, then the claim in (2.87) is a direct consequence of the proof of Lemma 3.3.10 for $\rho = \infty$. If $\liminf_{m\to\infty} e_m > 0$, then the claim follows from the fact that $\mu \in \mathbb{R}_\theta^{(2)}$ and

$$\mathrm{Var}_{\nu_\theta}(y_{0,m}) = \int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{t-s}^{(2)}\big((0,D_m),(k,A)\big)\, b_{t-s}^{(2)}\big((0,D_m),(k,A)\big)\, \mathbb{E}_\mu[g(x_i(s))]. \tag{3.193}$$

Since $e_m > 0$ for all $m \in \mathbb{N}_0$ and $\liminf_{m\to\infty} e_m > 0$, there is a positive probability that after the first steps the two random walks are both active at 0, i.e., are both in state $(0,A)$. Hence, for all $m \in \mathbb{N}_0$ there exists a constant $c > 0$ such that

$$\mathrm{Var}_{\nu_\theta}(y_{0,m}) \geq c\mathrm{Var}_{\nu_\theta}(x_0). \tag{3.194}$$

Since ν_θ is a non-trivial equilibrium, we have $\mathrm{Var}_{\nu_\theta}(x_0) > 0$.

- **Proof of coexistence for infinite seed-bank**

1. Properties of measures preserved under the evolution. For $\rho = \infty$, the class of preserved measures is also given by $\mathcal{R}_\theta^{(2)}$ (recall Definition 3.3.5). We show that if $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$ is colour regular, then $\mu \in \mathcal{R}_\theta^{(2)}$. Let the sets $\mathcal{E}_m(t)$, $t > 0$, $m \in \mathbb{N}_0$, be defined as in (3.185), and define λ_A and λ_{D_m} analogously to (3.112), like for $\rho < \infty$. The equivalent of (3.115) is

$$\begin{aligned}\mathbb{E}_\mu &\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\sum_{k\in\mathbb{G}} b_t^{(2)}((0,A),(k,A))x_k - \theta_x\right)^2\right]\\ &= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\int_{[-\pi,\pi]^d} \mathbb{E}_{(0,A),(0,A)}\left[\mathrm{e}^{-T(t)(1-a(\phi))}1_{\mathcal{E}(t)}\mathrm{e}^{-T'(t)(1-\bar{a}(\phi))}1_{\mathcal{E}'(t)}\right]\mathrm{d}\lambda_A.\end{aligned} \tag{3.195}$$

Using that $T(t), T'(t) \to \infty$ as $t \to \infty$ (see (3.169)), that $T(t)$, $T'(t)$, $\mathcal{E}(t)$, $\mathcal{E}'(t)$ are asymptotically independent and that $a(\cdot,\cdot)$ is irreducible, we still find

$$\lim_{t\to\infty} \mathbb{E}_\mu\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\sum_{k\in\mathbb{G}} b_t^{(2)}((0,A),(k,A))x_k - \theta_x\right)^2\right] = \lambda_A(\{0\}) \tag{3.196}$$

and, similarly,

$$\lim_{t\to\infty} \mathbb{E}_\mu \left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))} \sum_{k\in\mathbb{G}} b_t^{(2)}((0,A),(k,A)) y_{k,m} - \theta_{y,m} \right)^2 \right] = \lambda_{D_m}(\{0\}). \tag{3.197}$$

Since μ is ergodic, we have $\lambda_A(\{0\}) = 0$ and $\lambda_{D_m}(\{0\}) = 0$ for all $m \in \mathbb{N}_0$ (recall (3.121)). By the colour regularity,

$$\lim_{t\to\infty} \theta_x \mathbb{P}_{(0,A)}(\mathcal{E}(t)) + \sum_{m\in\mathbb{N}_0} \theta_{y,m} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t)) = \theta. \tag{3.198}$$

Therefore we can rewrite (3.186) as

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}_\mu \left[\left(\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}} b_t^{(2)}((0,A),(k,R_k))\, z_{(k,R_k)} - \theta \right)^2 \right] \\
&\quad \leq \lim_{t\to\infty} \mathbb{P}_{(0,A)}(\mathcal{E}(t))\, \mathbb{E}_\mu \left[\left(\sum_{k\in Z^d} \frac{b_t^{(2)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))} (x_k - \theta_x) \right)^2 \right] \\
&\qquad + \sum_{m\in\mathbb{N}_0} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t))\, \mathbb{E}_\mu \left[\left(\sum_{k\in Z^d} \frac{b_t^{(2)}((0,A),(k,(D_m))}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))} (y_{k,m} - \theta_{y,m}) \right)^2 \right] \\
&\quad = \lim_{t\to\infty} \mathbb{P}_{(0,A)}(\mathcal{E}(t))\, \lambda_A(\{0\}) \; + \sum_{m\in\mathbb{N}_0} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t))\, \lambda_{D_m}(\{0\}) = 0.
\end{aligned} \tag{3.199}$$

We conclude that indeed $\mu \in \mathcal{R}_\theta^{(2)}$.

Like for $\rho < \infty$, Lemma 3.2.8 carries over after we replace $b^{(1)}(\cdot,\cdot)$ by $b^{(2)}(\cdot,\cdot)$ and $\mathcal{R}_\theta^{(1)}$ by $\mathcal{R}_\theta^{(2)}$.

2. Existence of invariant measures ν_θ for $\rho = \infty$. Since the dynamics for $\rho = \infty$ and $\rho < \infty$ are the same, we can still use the coupling in (3.187)–(3.189). Also Lemma 3.3.6 holds for $\rho = \infty$, but if $\rho = \infty$, then the left-hand side of (3.191) can become infinite. Therefore we cannot use the line of argument used for model 1 to show that the coupling is successful for arbitrary colour regular initial measures $\mu_1,\ \mu_2 \in \mathcal{T}_\theta^{\mathrm{erg}}$. However, we can prove the following lemma.

Lemma 3.3.8 (Successful coupling). *If $\mu_1,\ \mu_2 \in \mathcal{T}_\theta^{\mathrm{erg}}$ are both colour regular and satisfy*

$$\hat{\mathbb{E}} \left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m |\delta_i(0)| \right] < \infty, \tag{3.200}$$

then the coupling in (3.187)–(3.189) *is successful.*

Proof. We proceed similarly as in Step 3 for $\rho < \infty$. Note, in particular, that $h_i(t)$ (recall (3.191)) is bounded from above by $\hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_i(0)|\right]$ (compare with (3.145)). Also for $\rho = \infty$ we obtain Lemma 3.2.11. Like for model 1, if we define

$$\begin{aligned} E_0 \times E_0 = &\Big\{\bar{z} \in E \times E\colon\ z^1_{(i,R_i)}(t) \geq z^2_{(i,R_i)}(t)\ \forall (i,R_i) \in \mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}\Big\} \\ &\cup \Big\{\bar{z} \in E \times E\colon\ z^2_{(i,R_i)}(t) \geq z^1_{(i,R_i)}(t)\ \forall (i,R_i) \in \mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}\Big\}, \end{aligned} \tag{3.201}$$

then we find $\lim_{t\to\infty} \mathbb{P}(E_0 \times E_0) = 1$ and hence the coupled diffusions $(Z^1(t))_{t\geq 0}$ and $(Z^2(t))_{t\geq 0}$ lay on top of each other as $t \to \infty$. However, in (3.149) the limiting distribution of $b^{(1)}_{t_n}(\cdot,\cdot)$ was used "to compensate" the factors K_m in $|\Delta_i| + \sum_{m\in\mathbb{N}_0} K_m |\delta_{i,m}|$. Since, for $\rho = \infty$, $b^{(1)}_{t_n}(\cdot,\cdot)$ does not have a well-defined limiting distribution for the projection on the colour components, we need a different strategy.

To obtain a successful coupling, as before, let $(t_n)_{n\in\mathbb{N}}$ be a subsequence such that $\nu^1_\theta = \lim_{n\to\infty} \mathcal{L}(Z^1(t_n))$ with $\mathcal{L}(Z^1(0)) = \mu^1$ and $\nu^2_\theta = \lim_{n\to\infty} \mathcal{L}(Z^2(t_n))$ with $\mathcal{L}(Z^2(0)) = \mu^2$. For $\mathbb{G} = \mathbb{Z}^d$, let $\Lambda_N = [0,N)^d \cap \mathbb{Z}^d$, $N \in \mathbb{N}$. (As noted before, for amenable groups $\mathbb{G}$, $(\Lambda_N)_{N\in\mathbb{N}}$ must be replaced by a so-called Fϕlner sequence.) Note that

$$\mathbb{E}_{\nu^1_\theta}\left[\left(\frac{1}{|\Lambda_N|}\sum_{j\in\Lambda_N} x_j - \theta\right)^2\right] = \frac{1}{|\Lambda_N|^2}\sum_{i,j\in\Lambda_N} \mathrm{Cov}_{\nu^1_\theta}(x_i,x_j). \tag{3.202}$$

Since μ^1 is colour regular and $\mu^1 \in \mathcal{T}^{\mathrm{erg}}_\theta$, we have $\mu^1 \in \mathcal{R}^{(2)}_\theta$. Hence, by Lemma 3.3.1,

$$\begin{aligned} \mathrm{Cov}_{\nu^1_\theta}(x_i,x_j) &= \lim_{n\to\infty} \mathrm{Cov}_{\mu^1}(x_i(t_n), x_j(t_n)) \\ &\leq \lim_{n\to\infty} \|g\| \int_0^{t_n} \mathrm{d}s \sum_{k\in\mathbb{G}} b^{(2)}_{(t_n-s)}((i,A),(k,A))\, b^{(2)}_{(t-s)}((j,A),(k,A)) \\ &\leq \|g\| \int_0^\infty \mathrm{d}s \sum_{k\in\mathbb{G}} \mathbb{E}_{(i,A),(j,A)}\left[a_{T(s)}(i,k)\, 1_{\mathcal{E}(s)}\, a_{T'(s)}(j,k)\, 1_{\mathcal{E}'(s)}\right] \\ &\leq \|g\| \int_0^\infty \mathrm{d}s\, \mathbb{E}_{(i,A),(j,A)}\left[\hat{a}_{T(s)+T'(s)}(i-j,0)\, 1_{\mathcal{E}(s)}\, 1_{\mathcal{E}'(s)}\right]. \end{aligned} \tag{3.203}$$

Since $I_{\alpha,\gamma} < \infty$, we see that the last integral is finite. Since $\lim_{||i-j||\to\infty} \hat{a}_t(i-j,0) = 0$ for all $t > 0$, it follows by transience and dominated convergence that $\lim_{||i-j||\to\infty} \mathrm{Cov}_{\nu^1_\theta}(x_i,x_j) = 0$. Since $\mathrm{Cov}_{\nu^1_\theta}(x_i,x_j) \leq 1$ for all $i,j \in \mathbb{G}$, for all $\epsilon > 0$

there exists an $L \in \mathbb{N}$ such that

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}_{\nu_\theta^1}\left[\left(\frac{1}{|\Lambda_N|}\sum_{j\in\Lambda_N} x_j - \theta\right)^2\right] &= \lim_{N\to\infty}\frac{1}{|\Lambda_N|^2}\sum_{i,j\in\Lambda_N}\mathrm{Cov}_{\nu_\theta^1}(x_i,x_j)\\
&= \lim_{N\to\infty}\frac{1}{|\Lambda_N|^2}\sum_{\substack{i,j\in\Lambda_N\\ \|i-j\|\le L}}\mathrm{Cov}_{\nu_\theta^1}(x_i,x_j) + \frac{1}{|\Lambda_N|^2}\sum_{\substack{i,j\in\Lambda_N\\ \|i-j\|> L}}\mathrm{Cov}_{\nu_\theta^1}(x_i,x_j)\\
&\le \lim_{N\to\infty}\frac{|\{i,j\in\Lambda_N\colon\ \|i-j\|\le L\}|}{|\Lambda_N|^2} + \epsilon\lim_{N\to\infty}\frac{|\{i,j\in\Lambda_N\colon\ \|i-j\|> L\}|}{|\Lambda_N|^2} < \epsilon.
\end{aligned} \tag{3.204}$$

We conclude that

$$\lim_{N\to\infty}\mathbb{E}_{\nu_\theta^1}\left[\left(\frac{1}{\Lambda_N}\sum_{j\in\Lambda_N} x_j-\theta\right)^2\right]=0, \tag{3.205}$$

and the same holds for ν_2^θ. Let $\lim_{n\to\infty}\mathcal{L}(\bar{Z}(t_n)) = \bar\nu_\theta$ such that $\lim_{n\to\infty}\mathcal{L}(Z^1(t_n)) = \nu_\theta^1$ and $\lim_{n\to\infty}\mathcal{L}(Z^2(t_n)) = \nu_\theta^2$. Then by translation invariance of $\bar\nu_\theta$ and the fact that $\bar\nu_\theta(E_0\times E_0)=1$, we find

$$\begin{aligned}
\int_{E\times E}\mathrm{d}\bar\nu_\theta|\Delta_i| &= \int_{E_0\times E_0}\mathrm{d}\bar\nu_\theta\frac{1}{|\Lambda_N|}\sum_{j\in\Lambda_N}|x_j^1-x_j^2|\\
&\le \int_{E_0}\mathrm{d}\nu_\theta^1\left|\frac{1}{|\Lambda_N|}\sum_{j\in\Lambda_N}x_j^1-\theta\right| + \int_{E_0}\mathrm{d}\nu_\theta^2\left|\frac{1}{|\Lambda_N|}\sum_{j\in\Lambda_N}x_j^2-\theta\right|.
\end{aligned} \tag{3.206}$$

Letting $N\to\infty$, we see by translation invariance of $\bar\nu_\theta$ that $\mathbb{E}_{\bar\nu_\theta}[|\Delta_i|]=0$ for all $i\in\mathbb{G}$.

The result in (3.205) holds also for x_i replaced by $y_{i,m}$, $m\in\mathbb{N}_0$, since the integral in (3.203) can only become smaller when we start from a dormant site. Replacing $|\Delta_i|$ in (3.206) by $|\delta_{i,m}|$, we obtain, for all $m\in\mathbb{N}_0$,

$$\mathbb{E}_{\bar\nu_\theta}[|\delta_{i,m}|]=0, \qquad \forall\, m\in\mathbb{N}_0. \tag{3.207}$$

We conclude that the coupling is successful. □

Let $(S_t)_{t\ge0}$ denote the semigroup associated with (2.12)–(2.13). To prove the existence of an invariant measure, note that $E\times E$ is a compact space. Hence, if $t_n\to\infty$, then the sequence μS_{t_n} has a convergent subsequence. In Lemma 3.3.9 below we show that each weak limit point of the sequence μS_{t_n} is invariant under the evolution of (2.12)–(2.13).

Lemma 3.3.9 (Invariant measure). *Suppose that $\mu\in\mathcal{R}_\theta^{(2)}$ and that μ is colour regular. If $t_n\to\infty$ and $\mu S_{t_n}\to\nu_\theta$, then ν_θ is an invariant measure under the evolution in* (2.12)–(2.13).

Proof. Fix $s > 0$. Let $\mu_1 = \mu$ and $\mu_2 = \mu S_s$. We couple μ_1 and μ_2 via their Brownian motions (see (3.187)–(3.189)). Note that, by the SSDE in (2.12)–(2.13),

$$\begin{aligned}
&\hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_{i,m}(0)|\right]\\
&= \mathbb{E}\left[|x_i(0)-x_i(s)| + \sum_{m\in\mathbb{N}_0} K_m|y_{i,m}(0)-y_{i,m}(s)|\right]\\
&= \mathbb{E}\Bigg[\Bigg|\int_0^s \sum_{j\in\mathbb{G}} a(i,j)\,[x_j(r)-x_i(r)]\,\mathrm{d}r + \int_0^s \sqrt{g(x_i(r))}\,\mathrm{d}w_i(r)\\
&\qquad + \int_0^s \sum_{m\in\mathbb{N}_0} K_m e_m\,[y_{i,m}(r)-x_i(r)]\,\mathrm{d}r\Bigg|\\
&\qquad + \sum_{m\in\mathbb{N}_0} K_m \int_0^s |e_m[y_{i,m}(r)-x_i(r)]|\,\mathrm{d}r\Bigg].
\end{aligned} \tag{3.208}$$

Using that all rates are finite and that, by Knight's theorem (see [62, Theorem V.1.9 p.183]), we can write the Brownian integral as a time-transformed Brownian motion, we see that $\hat{\mathbb{E}}[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_{i,m}(0)|] < \infty$. Hence, by Lemma 3.200, we can successfully couple μ^1 and μ^2, and $\lim_{n\to\infty}\mu^2 S_{t_n} = \lim_{n\to\infty}\mu S_s S_{t_n} = \nu_\theta$. By the Feller property of the SSDE in (2.12)–(2.13), it follows that

$$\nu_\theta S_s = \lim_{n\to\infty}\mu(t_n)S_s = \lim_{n\to\infty}\mu S_{t_n} S_s = \lim_{n\to\infty}\mu S_s S_{t_n} = \nu_\theta. \tag{3.209}$$

We conclude that ν_θ is indeed an invariant measure for the SSDE in (2.12)–(2.13). □

3. Invariant measures have vanishing covariances in the seed-bank direction for $\rho = \infty$. In this step we prove that an invariant measure ν_θ has vanishing variances in the seed-bank direction. In Step 5 we use this property to successfully couple any two invariant measures.

Lemma 3.3.10 (Deterministic deep seed-banks). *If $\nu_\theta = \lim_{n\to\infty}\mu S_{t_n}$ for some colour regular $\mu \in \mathcal{R}_\theta^{(2)}$ and $t_n \to \infty$, then*

$$\lim_{m\to\infty} \mathrm{Var}_{\nu_\theta}[y_{i,m}] = 0 \qquad \forall i \in \mathbb{G}. \tag{3.210}$$

Proof. Since ν_θ is translation invariant, it is enough to show that $\lim_{m\to\infty}\mathrm{Var}_{\nu_\theta}[y_{0,m}] = 0$. Since $\mu(0) \in \mathcal{R}_\theta^{(2)}$, it follows from Lemma 3.3.1 that

$$\begin{aligned}
&\lim_{m\to\infty}\mathrm{Var}_{\nu_\theta}[y_{0,m}]\\
&= \lim_{m\to\infty}\lim_{n\to\infty}\mathbb{E}_\mu\left[(y_{0,m}(t_n) - \mathbb{E}_\mu[y_{0,m}(t_n)])^2\right]\\
&= \lim_{m\to\infty}\lim_{n\to\infty}\int_0^{t_n}\mathrm{d}s\sum_{k\in\mathbb{G}} b^{(2)}_{t_n-s}((0,D_m),(k,A))\, b^{(2)}_{t_{n_k}-s}((0,D_m),(k,A))\mathbb{E}_z[g(x_k(s))].
\end{aligned} \tag{3.211}$$

Since g is positive and bounded, it is therefore enough to prove that

$$\lim_{m\to\infty}\lim_{n\to\infty}\int_0^{t_n} \mathrm{d}u \sum_{k\in\mathbb{G}} b_u^{(2)}((0,D_m),(k,A))\, b_u^{(2)}((0,D_m),(k,A)) = 0. \tag{3.212}$$

Recall (see e.g. (3.203)) that $b_u^{(2)}((0,D_m),(k,A))\, b_u^{(2)}((0,D_m),(k,A))$ is the probability that two random walks, denoted by RW and RW' and moving according to $b^{(2)}(\cdot,\cdot)$, are at time u at the same site k and both active. Define

$$\tau = \big\{t \geq 0 : RW(t) = RW'(t) = (i,A) \text{ for some } i \in \mathbb{G}\big\}. \tag{3.213}$$

Then we can rewrite the left-hand side of (3.212) as

$$\begin{aligned}
&\lim_{m\to\infty}\lim_{n\to\infty}\int_0^{t_n} \mathrm{d}u \sum_{k\in\mathbb{G}} b_u^{(2)}((0,D_m),(k,A))\, b_u^{(2)}((0,D_m),(k,A))\\
&= \lim_{m\to\infty}\lim_{n\to\infty}\int_0^{t_n} \mathrm{d}u\, \mathbb{E}_{(0,D_m),(0,D_m)}\left[\sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}}\, 1_{\{RW'(u)=k\}}\, 1_{\mathcal{E}(u)}\, 1_{\mathcal{E}'(u)}\right]\\
&= \lim_{m\to\infty}\lim_{n\to\infty}\int_0^{t_n} \mathrm{d}u\, \mathbb{E}_{(0,D_m),(0,D_m)}\Bigg[\sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}}\, 1_{\{RW'(u)=k\}}\, 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)}\\
&\qquad\qquad \times \big(1_{\{\tau<\infty\}} + 1_{\{\tau=\infty\}}\big)\Bigg]\\
&= \lim_{m\to\infty}\lim_{n\to\infty} \mathbb{E}_{(0,D_m),(0,D_m)}\Bigg[1_{\{\tau<\infty\}}\\
&\qquad \times \mathbb{E}_{(0,D_m),(0,D_m)}\left[\int_0^{t_n} \mathrm{d}u \sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}} 1_{\{RW'(u)=k\}} 1_{\mathcal{E}(u)} 1_{\mathcal{E}'(u)} \mid \mathcal{F}_\tau\right]\Bigg]\\
&= \lim_{m\to\infty}\lim_{n\to\infty} \mathbb{E}_{(0,D_m),(0,D_m)}\Bigg[1_{\{\tau<\infty\}}\\
&\qquad \times \mathbb{E}_{(0,A),(0,A)}\left[\int_0^{t_n-\tau} \mathrm{d}u \sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}} 1_{\{RW'(u)=k\}} 1_{\mathcal{E}(u)} 1_{\mathcal{E}'(u)}\right]\Bigg]\\
&\leq \lim_{m\to\infty}\lim_{n\to\infty} \mathbb{E}_{(0,D_m),(0,D_m)}\Bigg[1_{\{\tau<\infty\}}\\
&\qquad \times \mathbb{E}_{(0,A),(0,A)}\left[\int_0^{\infty} \mathrm{d}u \sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}} 1_{\{RW'(u)=k\}} 1_{\mathcal{E}(u)} 1_{\mathcal{E}'(u)}\right]\Bigg]\\
&= \lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,D_m)}(\tau<\infty)\, I_{\hat{a},\gamma},
\end{aligned} \tag{3.214}$$

where we use that $I_{\hat{a},\gamma} < \infty$, the strong Markov property, and the fact that for $\tau = \infty$ the product of the indicators equals 0 for all $u \in \mathbb{R}_{\geq 0}$. Therefore (3.210) holds if

$$\lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,D_m)}(\tau<\infty) = 0. \tag{3.215}$$

Define

$$\tau^* = \inf\big\{t \geq 0 : \text{ both } RW \text{ and } RW' \text{ are active at time } t\big\}. \tag{3.216}$$

Note that $\tau^* \leq \tau$. Theorefore we can write (recall that in model 2 the random walk kernel $a(\cdot,\cdot)$ is assumed to be symmetric),

$$\begin{aligned}
&\lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,D_m)}\left(\tau<\infty\right)\\
&= \lim_{m\to\infty} \mathbb{E}_{(0,D_m),(0,D_m)}[1_{\{\tau<\infty\}}]\\
&= \lim_{m\to\infty} \mathbb{E}_{(0,D_m),(0,D_m)}\left[1_{\{\tau^*<\infty\}}\,\mathbb{E}_{(0,D_m)^2}\left[1_{\{\tau<\infty\}}\mid \mathcal{F}_{\tau^*}\right]\right]\\
&= \lim_{m\to\infty} \mathbb{E}_{(0,D_m),(0,D_m)}\left[\mathbb{E}^{RW(\tau^*),RW'(\tau^*)}\left[1_{\{\tau<\infty\}}\right]\right]\\
&= \lim_{m\to\infty} \sum_{k,l\in\mathbb{G}} \mathbb{P}_{(0,D_m),(0,D_m)}\left(RW(\tau^*)=(k,A),RW'(\tau^*)=(l,A)\right)\\
&\qquad\qquad \times \mathbb{E}_{(k,A),(l,A)}\left[1_{\{\tau<\infty\}}\right]\\
&= \lim_{m\to\infty} \sum_{k,l\in\mathbb{G}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)}(0,k)\,\hat{a}_{T'(\tau^*)}(0,l)]\,\mathbb{E}_{(0,A),(l-k,A)}\left[1_{\{\tau<\infty\}}\right]\\
&= \lim_{m\to\infty} \sum_{k,l\in\mathbb{G}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)}(0,k)\,\hat{a}_{T'(\tau^*)}(-k,l-k)]\,\mathbb{E}_{(0,A),(l-k,A)}\left[1_{\{\tau<\infty\}}\right]\\
&= \lim_{m\to\infty} \sum_{k,j\in\mathbb{G}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)}(0,-k)\,\hat{a}_{T'(\tau^*)}(-k,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]\\
&= \lim_{m\to\infty} \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]\\
&= \lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\ \|j\|\leq L}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]\\
&\qquad + \lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\ \|j\|> L}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right].
\end{aligned} \tag{3.217}$$

To prove that the expression in the right-hand side tends to zero, we fix $\epsilon>0$ and prove that there exists an $L\in\mathbb{N}$ such that both sums are smaller that $\frac{\epsilon}{2}$.

Claim 1: There exists an L such that

$$\lim_{m\to\infty} \sum_{j\in\mathbb{G},\|j\|>L} \mathbb{E}_{(0,D_m)^2}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}[1_{\{\tau<\infty\}}] < \frac{\epsilon}{2}. \tag{3.218}$$

Using the symmetry of the kernel $a(\cdot,\cdot)$ in model 2, we find

$$\begin{aligned}
\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right] &= \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s\, 1_{\{\tau\in\mathrm{d}s\}}\right] \\
&\leq \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s \sum_{k\in\mathbb{G}} 1_{\mathcal{E}(s)} 1_{\mathcal{E}'(s)} 1_{\{RW=k\}} 1_{\{RW'=k\}}\right] \\
&\leq \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s \sum_{k\in\mathbb{G}} \hat{a}_{T(s)}(0,k)\, \hat{a}_{T'(s)}(j,k)\, 1_{\mathcal{E}(s)}\, 1_{\mathcal{E}'(s)}\right] \\
&\leq \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s\, \hat{a}_{T(s)+T'(s)}(j,0) 1_{\mathcal{E}(s)} 1_{\mathcal{E}'(s)}\right].
\end{aligned} \tag{3.219}$$

The last integral in the right-hand side is dominated by $I_{\hat{a},\gamma}$ (recall (2.100)). Since, for all $t\in\mathbb{R}_{\geq 0}$,

$$\lim_{\|j\|\to\infty} \hat{a}_t(0,j) = 0, \tag{3.220}$$

it follows by dominated convergence that for each $\epsilon>0$ we can find an L such that, for all $\|j\|>L$,

$$\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right] < \tfrac{\epsilon}{2}. \tag{3.221}$$

Hence, for L sufficiently large, we find

$$\begin{aligned}
&\lim_{m\to\infty} \sum_{j\in\mathbb{G},|\|j\||>L} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\left[\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]\right] \\
&\leq \lim_{m\to\infty} \tfrac{\epsilon}{2} \sum_{j\in\mathbb{G},\|j\|>L} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)] \leq \tfrac{\epsilon}{2}.
\end{aligned} \tag{3.222}$$

Claim 2: For L given as in Claim 1,

$$\lim_{m\to\infty} \sum_{j\in\mathbb{G},\|j\|\leq L} \mathbb{E}_{(0,D_m)^2}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\, \mathbb{E}^{(0,A),(j,A)}[1_{\{\tau<\infty\}}] < \frac{\epsilon}{2}. \tag{3.223}$$

For the first sum, note that

$$\begin{aligned}
&\lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\ \|j\|\leq L}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{(T(\tau^*)+T'(\tau^*))}(0,j)]\, \mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right] \\
&\quad\leq \lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\ \|j\|\leq L}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)] \\
&\quad= \lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\ \|j\|\leq L}} \mathbb{E}_{(0,A),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)],
\end{aligned} \tag{3.224}$$

where in the last equality we condition on the first time one of the two random walks wakes up, and use the strong Markov property. We will show that the right-hand side tends to zero as $m\to\infty$. Recall that we assumed (2.76): $e_m \sim Bm^{-\beta}$ for $\beta>0$. Note that, in order for the random walks to be both active at the same time, the random walk starting in $(0, D_m)$ has to become active at least once. Hence, for all $t \geq 0$, we have

$$\lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,A)}(\tau^* \leq t) \leq \lim_{m\to\infty} 1 - e^{-e_m t} = 0. \tag{3.225}$$

By (3.169) and [31], we also have for the random walk starting in $(0, A)$ that

$$\lim_{t\to\infty} T(t) \sim ct^\gamma. \tag{3.226}$$

Fix $\epsilon > 0$. Since $\lim_{t\to\infty} \hat{a}_t(0, j) = 0$ for all $j \in \mathbb{G}$, we can find a $T^\star$ such that, for all $t > T^\star$,

$$\sum_{\substack{j\in\mathbb{G}\\ \|j\|\leq L}} \hat{a}_t(0, j) < \tfrac{\epsilon}{6}. \tag{3.227}$$

By (3.226), we can find a $\tilde{t} \in \mathbb{R}_{\geq 0}$ such that $\mathbb{P}_{(0,A)}(T(\tilde{t}) > T^\star) \geq 1 - \frac{\epsilon}{6}$. By (3.225), we can find an $M \in \mathbb{N}_0$ such that for all $m > M$,

$$\lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,A)}(\tau^* \leq \tilde{t}) < \tfrac{\epsilon}{6}, \tag{3.228}$$

and hence

$$\lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\ \|j\|\leq L}} \mathbb{E}_{(0,A),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0, j)] < \tfrac{\epsilon}{6} + \tfrac{\epsilon}{6} + \tfrac{\epsilon}{6} = \tfrac{\epsilon}{2}. \tag{3.229}$$

□

4. Uniqueness of the invariant measure ν_θ when $\rho = \infty$.

Lemma 3.3.11 (Uniqueness of and convergence to ν_θ.). *For all $\theta \in (0, 1)$ there exists a unique invariant measure ν_θ such that $\lim_{t\to\infty} \mu(t) = \nu_\theta$ for all colour regular $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$.*

Proof. Suppose that ν_θ^1 and ν_θ^2 and are two different weak limit points of $\mu(t_n)$ as $t_n \to \infty$, and that $\mu \in \mathcal{R}_\theta^{(2)}$ is colour regular. Let $(\bar{Z}(t))_{t\geq 0} = (Z^1(t), Z^2(t))_{t\geq 0}$ be the coupled process from (3.133) with $\mathcal{L}(\bar{Z}(0)) = \bar{\nu}_\theta$, $\mathcal{L}(Z^1(0)) = \nu_\theta^1$ and $\mathcal{L}(Z^2(0)) = \nu_\theta^2$. Define the process Y^1 by

$$\begin{aligned} &Y^1 = \big(Y^1(m)\big)_{m\in\{-1\}\cup\mathbb{N}_0},\\ &Y^1(-1) = (x_i^1(0))_{i\in\mathbb{G}}, \qquad Y^1(m) = (y_{i,m}^1(0))_{i\in\mathbb{G}} \text{ for } m \in \mathbb{N}_0. \end{aligned} \tag{3.230}$$

Thus, Y^1 has state space $[0, 1]^{\mathbb{G}}$ and $\mathcal{L}(Y^1) = \mathcal{L}(Z^1(0)) = \nu_\theta^1$. We can interpret Y^1 as a process that describes the states of the population *in the seed-bank direction.* Similarly, define the process Y^2 by

$$\begin{aligned} &Y^2 = \big(Y^2(m)\big)_{m\in\{-1\}\cup\mathbb{N}_0},\\ &Y^2(-1) = (x_i^2(0))_{i\in\mathbb{G}}, \qquad Y^2(m) = (y_{i,m}^2(0))_{i\in\mathbb{G}} \text{ for } m \in \mathbb{N}_0. \end{aligned} \tag{3.231}$$

Thus, Y^2 has state space $[0, 1]^{\mathbb{G}}$ and $\mathcal{L}(Y^2) = \mathcal{L}(Z^2(0)) = \nu_\theta^2$.

Define the σ-algebra's $\mathcal{B}^1_M$ and $\mathcal{B}^1$, respectively, $\mathcal{B}^2_M$ and $\mathcal{B}^2$ by

$$\mathcal{B}^k = \cap_{M\in\mathbb{N}_0}\mathcal{B}^k_M, \qquad \mathcal{B}^k_M = \sigma\big(\, y^k_{i,m} \colon\, i\in\mathbb{G},\, m\geq M\big), \quad k\in\{1,2\}. \tag{3.232}$$

Here, $\mathcal{B}^1$ and $\mathcal{B}^2$ are *the tail-σ-algebras in the seed-bank direction.* By Lemma 3.3.10, we have

$$\lim_{m\to\infty}\mathcal{L}_{\nu^1_\theta}(y_{i,m}) = \lim_{m\to\infty}\mathcal{L}_{\nu^2_\theta}(y_{i,m}) = \delta_\theta. \tag{3.233}$$

Hence, $\mathcal{B}^1 = \mathcal{B}^2$, both are trivial, and ν^1_θ and ν^2_θ agree on $\mathcal{B}$. Therefore Goldstein's Theorem [39] implies that there exists a successful coupling of Y^1 and Y^2. Consequently, there exists a random variable $T^{\text{coup}}\in\{-1\}\cup\mathbb{N}_0$ such that, for all $m\geq T^{\text{coup}}$, $Y^1(m)=Y^2(m)$, i.e., $|\delta_{i,m}(0)|=0$ for all $i\in\mathbb{G}$ and $\mathbb{P}(T^{\text{coup}}<\infty)=1$. Hence

$$\hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_i(0)|\right] = \hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m=0}^{T^{\text{coup}}} K_m|\delta_i(0)|\right]. \tag{3.234}$$

However, we cannot conclude that the left-hand side of (3.234) is finite. Therefore, let $\bar{\nu}_\theta|_{\{T^{\text{coup}}<T\}}$ denote the restriction of the measure $\bar{\nu}_\theta$ to the set $\{T^{\text{coup}}<T\}$. Since $\{T^{\text{coup}}<T\}$ is a translation-invariant event in the spatial direction, the measure $\bar{\nu}_\theta|_{\{T^{\text{coup}}<T\}}$ is translation invariant. Moreover,

$$\begin{aligned}&\hat{\mathbb{E}}_{\bar{\nu}_\theta|_{\{T^{\text{coup}}<T\}}}\left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_i(0)|\right]\\ &= \hat{\mathbb{E}}_{\bar{\nu}_\theta|_{\{T^{\text{coup}}<T\}}}\left[|\Delta_i(0)| + \sum_{m=0}^{T} K_m|\delta_i(0)|\right] < \infty.\end{aligned} \tag{3.235}$$

Therefore we can use the dynamics in (3.191) and conclude that, for all $T\in\mathbb{N}$, $\hat{\mathbb{P}}_{\bar{\nu}_\theta|_{\{T^{\text{coup}}<T\}}}(E_0\times E_0) = 1$ (recall (3.201)). Since $\lim_{T\to\infty}\bar{\nu}_\theta|_{\{T^{\text{coup}}<T\}} = \bar{\nu}_\theta$, it follows that

$$\hat{\mathbb{P}}_{\bar{\nu}_\theta}(E_0\times E_0) = 1. \tag{3.236}$$

By (3.206) and (3.207), we conclude that $\nu^1_\theta = \nu^2_\theta$ and hence that all weak limit points of $(\mu(t))_{t\geq 0}$ are the same. Suppose now that $\mu^1(0)\in\mathcal{T}^{\text{erg}}_\theta$ and $\mu^2(0)\in\mathcal{T}^{\text{erg}}_\theta$ are two different colour regular initial measures. By the above argument, we know that $\lim_{t\to\infty}\mu^1(t)=\nu^1_\theta$ and $\lim_{t\to\infty}\mu^2(t)=\nu^2_\theta$. By Lemma 3.3.10, we know that ν^1_θ and ν^2_θ have the same trivial tail-σ-algebras in the seed-bank direction. Hence, repeating the above argument, we find that $\nu^1_\theta=\nu^2_\theta$. We conclude that for each colour regular initial measure $\mu\in\mathcal{T}^{\text{erg}}_\theta$ the SSDE in (2.12)–(2.13) converges to a unique non-trivial equilibrium measure ν_θ. □

5. Ergodicity, mixing and associatedness. The equivalent of Lemma 3.2.13 for $\rho=\infty$ follows in the same way as for $\rho<\infty$.

§3.3.4 Proof of the dichotomy

Theorem 2.3.3(I)(a) follows from Lemma (3.3.7) and Steps 3-5 in Section 3.3.3. The equality $\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta$, $m \in \mathbb{N}_0$, follows from (2.12)–(2.13), the fact that ν_θ is an equilibrium measure, and the preservation of θ (see Section 2.3.2). Theorem 2.3.3(I)(b) follows by combining Lemma 3.3.3 with the analogue of Lemma 3.2.5. Theorem 2.3.3(II) follows from Lemmas 3.3.3, 3.3.10, 3.3.11, the analogue of Lemma 3.2.5, and Step 6 in Section 3.3.3. The equality $\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta$, $m \in \mathbb{N}_0$, follows from (3.165) in Step 1 of Section 3.3.3.

Corollary 2.3.4(1) corresponds to $\gamma \in (1,\infty)$ and $\rho < \infty$, and migration dominates. Corollary 2.3.4(2) corresponds to $\gamma \in [\frac{1}{2}, 1]$ and $\rho = \infty$, and $I_{\hat{a},\gamma}$ shows in interplay between migration and seed-bank. Corollary 2.3.4(3) corresponds to $\gamma \in (0, \frac{1}{2}, 1)$ and $\rho = \infty$, and the seed-bank dominates: $I_{\hat{a},\gamma} < \infty$ because $\hat{a}_t(0,0) \leq 1$.

§3.3.5 Different dichotomy for asymmetric migration

It remains to explain how the counterexample below Theorem 2.3.3 arises. We focus on the case when $\rho < \infty$, which implies $\mathbb{E}(\tau) < \infty$, but we assume $\mathbb{E}(\tau^2) = \infty$. Therefore the central limit theorem does not hold for $T(t)$, $T'(t)$, and $\Delta(t) \gg \sqrt{M(t)}$. Hence (3.75) must be replaced by

$$f(t) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-[1+o(1)]\,2Bt\,[1-\hat{a}(\phi)]}\, \mathbb{E}\left[\cos\Big(\Delta(t)\tilde{a}(\phi)\Big)\right]. \tag{3.237}$$

The key observation is that if $\tilde{a}(\phi) \neq 0$ (due to the asymmetry of $a(\cdot,\cdot)$; recall (3.66)), then the expectation in (3.237) can change the integrability properties of $f(t)$.

Under the assumption that τ has a *one-sided stable distribution* with parameter $\gamma \in (1,2)$, we have (3.70) with $A = \chi/(1+\rho)$ and $B = 1/(1+\rho)$, while there exists a constant $C \in (0,\infty)$ such that (see [34, Chapter XVII])

$$\mathbb{E}[\cos(\Delta(t)\tilde{a}(\phi))] = \mathrm{e}^{-[1+o(t)]\,At|C\tilde{a}(\phi)|^\gamma}. \tag{3.238}$$

Substituting (3.238) into (3.237), we see that for large t the contribution to $f(t)$ comes from ϕ such that $\hat{a}(\phi) \to 1$ and $\tilde{a}(\phi) \to 0$. By our choice of the migration kernel in (2.90), this holds as $\phi = (\phi_1,\phi_2) \to (0,0)$. Using that $1 - \hat{a}(\phi) \sim \frac{1}{2}(\phi_1^2 + \phi_2^2)$ and $\tilde{a}(\phi) \sim \frac{1}{2}\eta(\phi_1+\phi_2)$ for $(\phi_1,\phi_2) \to (0,0)$, we find that (3.237) equals

$$f(t) = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} \mathrm{d}\phi \, \mathrm{e}^{-[1+o(1)]\,\{Bt(\phi_1^2+\phi_2^2)+At[|\frac{1}{2}C\eta(\phi_1+\phi_2)|]^\gamma\}}, \qquad t \to \infty. \tag{3.239}$$

Hence the integral in (3.239) is determined by ϕ such that

$$B(\phi_1^2+\phi_2^2) + A\big[|\tfrac{1}{2}C\eta(\phi_1+\phi_2)|\big]^\gamma \leq \frac{c}{t}. \tag{3.240}$$

for c a positive constant, and we find that $f(t) \asymp t^{-(\frac{1}{\gamma}+\frac{1}{2})}$. Since $\gamma \in (1,2)$, $f(t)$ is much smaller than $\hat{a}_t(0,0) \asymp 1/t$, valid for two-dimensional simple random walk. Thus we see that $t \mapsto f(t)$ is integrable, while $t \mapsto \hat{a}_t(0,0)$ is not.

§3.3.6 Modulation of the law of the wake-up times by a slowly varying function

The integral in (2.96) is the *total hazard of coalescence* of two dual lineages:

- If $\gamma \in (0,1)$, then the probability for each of the lineages to be active at time s decays like $\asymp \varphi(s)^{-1}s^{-(1-\gamma)}$ [1]. Hence the expected total time they are active up to time s is $\asymp \varphi(s)^{-1}s^{\gamma}$. Because the lineages only move when they are active, the probability that the two lineages meet at time s is $\asymp a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. Hence the total hazard is $\asymp \int_1^\infty \mathrm{d}s\, \varphi(s)^{-2}s^{-2(1-\gamma)}\, a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. After the transformation $t = t(s) = \varphi(s)^{-1}s^{\gamma}$, we get the integral in (2.96), modulo a constant. (When carrying out this transformation, we need that $\lim_{s\to\infty} s\varphi'(s)/\varphi(s) = 0$, which is immediate from (2.95), and $\varphi(t(s))/\varphi(s) \asymp 1$ as $s \to \infty$, which is immediate from the bound we imposed on ψ together with the fact that $\lim_{s\to\infty} \log\varphi(s)/\log s = 0$.)

- If $\gamma = 1$, then the probability for each of the lineages to be active at time s decays like $\hat{\varphi}(s)^{-1}$ [1]. Hence the expected total time they are active up to time s is $\asymp s\hat{\varphi}(s)^{-1}$. Hence the total hazard is $\asymp \int_1^\infty \mathrm{d}s\, \hat{\varphi}(s)^{-2}\, a^{(N)}_{\hat{\varphi}(s)^{-1}s}(0,0)$. After the transformation $t = t(s) = \hat{\varphi}(s)^{-1}s$, we get the integral in (2.96), modulo a constant.

§3.4 Proofs: Long-time behaviour for Model 3

The arguments for model 2 in Section 3.3 all carry over with minor adaptations. The only difference is that for $\rho = \infty$ the clustering criterion changes. In this section we prove the new clustering criterion and comment on the modifications needed in the corresponding proofs for model 2 in Section 3.3.

§3.4.1 Moment relations

Like in model 1 and 2, we can relate the first and second moments of the system in (2.18)–(2.19) to the random walk that evolves according to the transition kernel $b^{(3)}(\cdot,\cdot)$ on $\mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}$ given by (2.53). Replacing in Lemma 3.3.1 the kernel $b^{(2)}(\cdot,\cdot)$ by $b^{(3)}(\cdot,\cdot)$, we find the moment relation for model 3. Also here these moment relations hold for all $g \in \mathcal{G}$. Moreover these moment relations holds for $\rho < \infty$ as well as for $\rho = \infty$.

§3.4.2 The clustering case

To obtain the equivalent of Lemma 3.3.3, we need to replace the kernel $\hat{a}(\cdot,\cdot)$ by the convoluted kernel $(\hat{a} * \hat{a}^{\dagger})(\cdot,\cdot)$. Each time one of the two copies of the random walk with migration kernel $a(\cdot,\cdot)$ moves from the active state to the dormant state, it makes a transition according to the displacement kernel $a^{\dagger}(\cdot,\cdot)$ (recall (2.97)). Therefore the expression in (3.60) needs to be replaced by

$$I = \int_0^\infty \mathrm{d}t \sum_{k,k'\in\mathbb{N}} \sum_{i,i'\in\mathbb{G}} \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,A)}\Big[\hat{a}_{T(k,t)}(0,i)\,\hat{a}_{T'(k',t)}(0,i')\,\hat{a}^\dagger_k(i,j)\,\hat{a}^\dagger_{k'}(i',j)\,1_{\mathcal{E}(k,t)}\,1_{\mathcal{E}'(k',t)}\Big], \tag{3.241}$$

where $\hat{a}^\dagger_k(\cdot,\cdot)$ is the step-k transition kernel of the random walk with displacement kernel $\hat{a}^\dagger(\cdot,\cdot)$. Using the *symmetry* of both kernels, we can carry out the sum over j, i' and write

$$\begin{aligned} I &= \int_0^\infty \mathrm{d}t \sum_{k,k'\in\mathbb{N}} \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,A)}\Big[\hat{a}_{T(k,t)+T'(k',t)}(0,j)\ \hat{a}^\dagger_{k+k'}(0,j)\,1_{\mathcal{E}(k,t)}\,1_{\mathcal{E}'(k',t)}\Big] \\ &= \int_0^\infty \mathrm{d}t \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,A)}\Big[\hat{a}_{T(t)+T'(t)}(0,j)\ \hat{a}^\dagger_{N(t)+N'(t)}(0,j)\,1_{\mathcal{E}(t)}\,1_{\mathcal{E}'(t)}\Big] \\ &= \int_0^\infty \mathrm{d}t\,\mathbb{E}_{(0,A)}\Big[\big(\hat{a}_{T(t)+T'(t)} * \hat{a}^\dagger_{N(t)+N'(t)}\big)(0,0)\,1_{\mathcal{E}(t)}\,1_{\mathcal{E}'(t)}\Big]. \end{aligned} \tag{3.242}$$

The last expression is the analogue of (3.63).

For $\rho < \infty$, following the same line of argument as for model 2, we find with the help of (2.98) that

$$I \asymp \int_1^\infty \mathrm{d}t\,(\hat{a}_t * \hat{a}^\dagger_t)(0,0). \tag{3.243}$$

For $\rho = \infty$, with the help of the Fourier transform we compute

$$\begin{aligned} &\mathbb{E}_{(0,A)}\big[\big(a_{T(t)+T'(t)} * a^\dagger_{N(t)+N'(t)}\big)(0,0)\,\big] \\ &= \mathbb{E}_{(0,A)}\left[\frac{1}{(2\pi)^d}\int_{(-\pi,\pi]^d} \mathrm{d}\phi\,\mathrm{e}^{-(T(t)+T'(t))[1-\hat{a}(\phi)]}\,\hat{a}^\dagger(\phi)^{N(t)+N'(t)}\right] \\ &= \frac{1}{(2\pi)^d}\int_{(-\pi,\pi]^d} \mathrm{d}\phi\,\mathrm{e}^{-[1+o(1)]\,2ct^{-\gamma}\,[1-\hat{a}(\phi)]}\,\mathrm{e}^{-[1+o(1)]\,2t^{-\gamma}[1-\hat{a}^\dagger(\phi)]} \\ &\asymp (\hat{a}_{ct^{-\gamma}} * \hat{a}^\dagger_{t^{-\gamma}})(0,0) \asymp (\hat{a}_{t^{-\gamma}} * \hat{a}^\dagger_{t^{-\gamma}})(0,0), \end{aligned} \tag{3.244}$$

where we use (2.98), (3.169) and the fact that deviations of $T(t)/t^\gamma$ and $T'(t)/t^\gamma$ away from order 1 are stretched exponentially costly in t [31]. Hence

$$I \asymp \int_1^\infty \mathrm{d}t\,t^{-2(1-\gamma)}(\hat{a}_{t^\gamma} * \hat{a}^\dagger_{t^\gamma})(0,0). \tag{3.245}$$

Putting $s = t^\gamma$ we obtain, instead of (3.180),

$$I = \infty \quad \Longleftrightarrow \quad I_{\hat{a}*\hat{a}^\dagger,\gamma} = \infty \tag{3.246}$$

with

$$I_{\hat{a}*\hat{a}^\dagger,\gamma} = \int_1^\infty \mathrm{d}s\,s^{-(1-\gamma)/\gamma}\,(\hat{a}_s * \hat{a}^\dagger_s)(0,0), \tag{3.247}$$

which is precisely the integral in (2.100).

§3.4.3 The coexistence case

The coexistence results in Theorem 2.3.6 follow for both $\rho < \infty$ and $\rho = \infty$ by the same type of argument as the one we used for model 2 in Section 3.3.3. We replace (2.12)–(2.13) by (2.18)–(2.19), replace $b^{(2)}(\cdot,\cdot)$ (see 2.41) by $b^{(3)}(\cdot,\cdot)$ (see 2.53), and use the Fourier transform of $\hat{a} * \hat{a}^{\dagger}(\cdot,\cdot)$ instead of $\hat{a}(\cdot,\cdot)$. The key of the argument is that, in the coexistence case, for $\rho < \infty$ we have $I_{\hat{a}*\hat{a}^{\dagger}} < \infty$, while for $\rho = \infty$ we have $I_{\hat{a}*\hat{a}^{\dagger},\gamma} < \infty$.

§3.4.4 Proof of the dichotomy

This follows in exactly the same way as for model 2.

APPENDIX A

Appendix Part I

§A.1 Derivation of continuum frequency equations

Model 1. We give the derivation of (2.4)–(2.5) as the continuum limit of an individual-based model when the size of the colonies tends to infinity. We start with the continuum limit of the Fisher-Wright model with (strong) seed-bank for a *single-colony* model as defined in [12]. Subsequently we show how the limit extends to a *multi-colony* model with seed-bank.

Single-colony model. The Fisher-Wright model with (strong) seed-bank defined in [12] consists of a *single colony* with $N \in \mathbb{N}$ active individuals and $M \in \mathbb{N}$ dormant individuals. Each individual can carry one of two types: $\heartsuit$ or $\diamondsuit$. Let $\epsilon \in [0,1]$ be such that ϵN is integer and $\epsilon N \leq M$. Put $\delta = \frac{\epsilon N}{M}$. The evolution of the population is described by a discrete-time Markov chain that undergoes four transitions per step:

(1) From the N active individuals, $(1-\epsilon)N$ are selected uniformly at random without replacement. Each of these individuals resamples, i.e. it adopts the type of an active individual selected uniformly at random with replacement, and remains active.

(2) Each of the ϵN active individuals not selected first resamples, it adopts the type of an active individual selected uniformly at random with replacement, and subsequently becomes dormant.

(3) From the M dormant individuals, $\delta M = \epsilon N$ are selected uniformly at random without replacement, and each of these becomes active. Since these individuals come from the dormant population they do not resample.

(4) Each of $(1-\delta)M$ dormant individuals not selected remains dormant and retains its type.

Note that the total sizes of the active and the dormant population remain fixed. During the evolution the dormant and active population *exchange* individuals. We are interested in the fractions of individuals of type $\heartsuit$ in the active and the dormant population. For an example of the evolution see Fig. 1.3.

Let $c = \epsilon N = \delta M$, i.e., c is the number of pairs of individuals that change state. Label the N active individuals from 1 to N and the M dormant individuals from 1 up to M. We denote by $[N] = \{1, \dots, N\}$ and by $[M] = \{1, \dots, M\}$. Let

$\xi(k) = (\xi_j(k))_{j\in[N]} \in \{0,1\}^{[N]}$ be the random vector where $\xi_j(k) = 1$ if the j'th individual is of type $\heartsuit$ at time k and $\xi_j(k) = 0$ if the j'th individual is of type $\diamondsuit$ at time k. Similarly, we let $\eta(k) = (\eta_j(k))_{j\in[M]} \in \{0,1\}^{[M]}$ be the random vector where $\eta_j(k) = 1$ if the j'th individual is of type $\heartsuit$ at time k and $\eta_j(k) = 0$ if the j'th individual is of type $\diamondsuit$ at time k. Let $I^N = \{0, \frac{1}{N}, \frac{2}{N}, \frac{3}{N}\dots, 1\}$ and $I^M = \{0, \frac{1}{M}, \frac{2}{M}, \frac{3}{M}\dots, 1\}$. Define the variables

$$\begin{aligned} X^N(k) &= \frac{1}{N}\sum_{j\in[N]} \mathbf{1}_{\{\xi_j(k)=\heartsuit\}} \quad \text{on } I^N,\\ Y^N(k) &= \frac{1}{N}\sum_{j\in[N]} \mathbf{1}_{\{\eta_j(k)=\heartsuit\}} \quad \text{on } I^M. \end{aligned} \tag{A.1}$$

Let $\mathbb{P}_{x,y}$ denote the law of

$$(X^N, Y^N) = (X^N(k), Y^N(k))_{k\in\mathbb{N}_0} \tag{A.2}$$

given that $(X^N(0), Y^N(0)) = (x,y) \in I^N \times I^M$. Then, as shown in [12],

$$\begin{aligned} p_{x,y}(\bar{x},\bar{y}) &= \mathbb{P}_{x,y}(X_1^N = \bar{x}, Y_1^N = \bar{y})\\ &= \sum_{c'=0}^{c} \mathbb{P}_{x,y}(Z = c')\,\mathbb{P}_{x,y}(U = \bar{x}N - c')\,\mathbb{P}_{x,y}(V = \bar{y}M - yM + c'). \end{aligned} \tag{A.3}$$

Here, Z denotes the number of dormant $\heartsuit$-individuals in generation 0 that become active in generation 1 ($\mathcal{L}_{x,y}(Z) = \mathrm{Hyp}_{M,c,yM}$), U denotes the number of active individuals in generation 1 that are offspring of active $\heartsuit$-individuals in generation 0 ($\mathcal{L}_{x,y}(U) = \mathrm{Bin}_{N-c,x}$), and V denotes the number of active individuals in generation 0 that become dormant $\heartsuit$-individuals in generation 1 ($\mathcal{L}_{x,y}(V) = \mathrm{Bin}_{c,x}$).

Speed up time by a factor N. The generator G^N for the process

$$((X^N(\lfloor Nk\rfloor), Y^N(\lfloor Nk\rfloor))_{k\in\mathbb{N}_0} \tag{A.4}$$

equals

$$\begin{aligned} &(G^N f)(x,y) = N\,\mathbb{E}_{x,y}\big[f(X^N(1), Y^N(1)) - f(x,y)\big],\\ &(x,y) \in I^N \times I^M, \end{aligned} \tag{A.5}$$

where the prefactor N appears because one step of the Markov chain takes time $\frac{1}{N}$. Inserting the Taylor expansion for f (which we assume to be smooth), using that $X^N(1) = \frac{U+Z}{N}$ and $Y^N(1) = \frac{yM+V-U}{M}$ and letting $N \to \infty$, we end up with the limiting generator G given by

$$\begin{aligned} &(Gf)(x,y) = c(y-x)\frac{\partial f}{\partial x}(x,y) + \frac{c}{K}(x-y)\frac{\partial f}{\partial y}(x,y) + \tfrac{1}{2}x(1-x)\frac{\partial^2 f}{\partial x^2}(x,y),\\ &(x,y) \in [0,1]\times[0,1], \end{aligned} \tag{A.6}$$

where $K = \frac{M}{N}$ is the relative size of the dormant population compared to the active population. This is the generator of the Markov process in the continuum limit [32, Section 7.8]. It follows from the form of G that this limit is described by the system of coupled stochastic differential equations

$$\begin{aligned} \mathrm{d}x(t) &= c\,[y(t) - x(t)]\,\mathrm{d}t + \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t), \\ \mathrm{d}y(t) &= \frac{c}{K}\,[x(t) - y(t)]\,\mathrm{d}t. \end{aligned} \tag{A.7}$$

This is the version of (2.4)–(2.5) for a single colony (no migration) and exchange rate

$$e = \frac{c}{K}. \tag{A.8}$$

Multi-colony model. First fix a number $L \in \mathbb{N}$ and consider $|\mathbb{G}| = L$ colonies. The *multi-colony* version with migration is obtained by allowing the $(1-\epsilon)N$ selected active individuals to undergo a migration in step (1):

(1) Each active individual at colony $i \in \mathbb{G}$ chooses colony $j \in \mathbb{G}$ with probability $\frac{1}{N}a(i,j)$ and adopts the type of a parent chosen from colony j. If an active individual does not migrate, it adopts the type of a parent chosen from its own population.

Using the same strategy as in the single-colony model, this results in (2.4)–(2.5), for $|\mathbb{G}| = L$. Subsequently we can let $L \to \infty$ and use convergence of generators to obtain (2.4)–(2.5) for countable $\mathbb{G}$.

Model 2. The same argument works for (2.12)–(2.13). Steps (1)-(4) are extended by considering a seed-bank with colours labelled by $\mathbb{N}_0$. First we consider the truncation where only finitely many colours are allowed, for which the argument carries through with minor adaptations. Afterwards, we pass to the limit of infinitely many colours, which is straightforward for a finite time horizon because large colours are only seen after large times. See also [60].

Model 3. To get (2.18)–(2.19), also extend Step (3) by adding a displacement via the kernel $a^{\dagger}(\cdot,\cdot)$ for each transition into the seed-bank.

§A.2 Alternative models

In this appendix we consider the Moran versions of models 1 and 2. What is written below is based on [60]. In the Moran version each active individual resamples at rate 1 and becomes dormant at a certain rate, while each dormant individual does not resample and becomes active at a certain rate. Since switches between active and dormant are done independently, the sizes of the active and the dormant population are *no longer* fixed and individuals *change* state without the necessity to *exchange* state. In model 1 there are two Poisson clocks, in model 2 there are two sequences of Poisson clocks, namely, two for each colour. In Appendices A.2.1–A.2.2 we compute the scaling limit for the case where the number of colours is $\mathfrak{m} = 1$ and $\mathfrak{m} = 2$, respectively. The extension to $\mathfrak{m} \geq 3$ is given in Appendix A.2.3. Migration can be added in the same way as is done in Appendix A.1.

§A.2.1 Alternative for Model 1

To describe the Moran version of Model 1 we need the following variables.

- Total number of individuals: $N \in \mathbb{N}$.
- Two types: $\heartsuit$ and $\diamondsuit$.
- $X(t)$ is the number of $\heartsuit$-individuals in the active population at time t.
- $Y(t)$ is the number of $\heartsuit$-individuals in the dormant population at time t.
- $Z(t)$ is the number of individuals in the active population at time t (either $\heartsuit$ or $\diamondsuit$).

In the Moran model with seed-bank each active individual resamples at rate 1, each active individual becomes dormant at rate ϵ and each dormant individual becomes active at rate δ. Hence the transition rates for $(X(t), Y(t), Z(t))$ are:

- $(i, j, k) \to (i+1, j, k)$ at rate $(k-i)\frac{i}{k}$.
- $(i, j, k) \to (i-1, j, k)$ at rate $i\frac{(k-i)}{k}$.
- $(i, j, k) \to (i-1, j+1, k-1)$ at rate ϵi.
- $(i, j, k) \to (i+1, j-1, k+1)$ at rate δj.
- $(i, j, k) \to (i, j, k-1)$ at rate $\epsilon \frac{k-i}{N}$.
- $(i, j, k) \to (i, j, k+1)$ at rate $\delta \frac{N-k-j}{N}$.

For the scaling limit we consider the variables

$$\bar{X}(t) = \frac{1}{N} X(Nt), \quad \bar{Y}(t) = \frac{1}{N} Y(Nt), \quad \bar{Z}(t) = \frac{1}{N} Z(Nt). \tag{A.9}$$

Hence

$$(\bar{X}(t), \bar{Y}(t), \bar{Z}(t)) \in I^N \times I^N \times I^N, \qquad I^N = \left\{0, \tfrac{1}{N}, \tfrac{2}{N}, \dots, \tfrac{N-1}{N}, 1\right\}. \tag{A.10}$$

Since in (A.9) we speed up time by a factor N, we must also speed up the transition rates by a factor N. To get a meaningful scaling limit, we assume that there exist $c^A, c^D \in (0, \infty)$ such that (see [12, p. 8])

$$N\epsilon = c^A, \qquad N\delta = c^D, \qquad N \in \mathbb{N}. \tag{A.11}$$

We can then write down the generator G^N:

$$
\begin{aligned}
(G^N f)\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right) = {} & N(k-i)\frac{i}{k}\left[f\left(\frac{i+1}{N},\frac{j}{N},\frac{k}{N}\right)-f\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)\right]\\
& + Ni\frac{k-i}{k}\left[f\left(\frac{i-1}{N},\frac{j}{N},\frac{k}{N}\right)-f\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)\right]\\
& + c^A i\left[f\left(\frac{i-1}{N},\frac{j+1}{N},\frac{k-1}{N}\right)-f\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)\right]\\
& + c^D j\left[f\left(\frac{i+1}{N},\frac{j-1}{N},\frac{k+1}{N}\right)-f\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)\right]\\
& + c^A (k-i)\left[f\left(\frac{i}{N},\frac{j}{N},\frac{k-1}{N}\right)-f\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)\right]\\
& + c^D (N-k-j)\left[f\left(\frac{i}{N},\frac{j}{N},\frac{k+1}{N}\right)-f\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)\right]
\end{aligned}
\tag{A.12}
$$

Assuming that f is smooth and Taylor expanding f around $\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right)$, we get

$$
\begin{aligned}
(G^N f)\left(\frac{i}{N},\frac{j}{N},\frac{k}{N}\right) = {} & \frac{i(k-i)}{k}\left[\left(\frac{1}{N}\right)\frac{\partial^2 f}{\partial x^2}+\mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right]\\
& + c^A i\left[\left(\frac{-1}{N}\right)\frac{\partial f}{\partial x}+\left(\frac{1}{N}\right)\frac{\partial f}{\partial y}+\left(\frac{-1}{N}\right)\frac{\partial f}{\partial z}+\mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right]\\
& + c^D j\left[\left(\frac{1}{N}\right)\frac{\partial f}{\partial x}+\left(\frac{-1}{N}\right)\frac{\partial f}{\partial y}+\left(\frac{1}{N}\right)\frac{\partial f}{\partial z}+\mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right]\\
& + c^A (k-i)\left[\left(\frac{-1}{N}\right)\frac{\partial f}{\partial z}+\mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right]\\
& + c^D (N-k-j)\left[\left(\frac{1}{N}\right)\frac{\partial f}{\partial z}+\mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right].
\end{aligned}
\tag{A.13}
$$

Next, suppose that

$$
\lim_{N\to\infty}\frac{i}{N}=x,\qquad \lim_{N\to\infty}\frac{j}{N}=y,\qquad \lim_{N\to\infty}\frac{k}{N}=z.
\tag{A.14}
$$

Letting $N\to\infty$ in (A.13), we obtain the limiting generator G:

$$
\begin{aligned}
(Gf)(x,y,z) = {} & z\frac{x}{z}\left(1-\frac{x}{z}\right)\left(\frac{\partial^2 f}{\partial x^2}\right)+[c^D\,y-c^A\,x]\frac{\partial f}{\partial x}\\
& + [c^A\,x-c^D\,y]\frac{\partial f}{\partial y}+\left[c^D\,(1-z)-c^A\,z\right]\frac{\partial f}{\partial z}.
\end{aligned}
\tag{A.15}
$$

Therefore the continuum limit equals

$$
\begin{aligned}
\mathrm{d}x(t) &= \sqrt{z(t)\,\frac{x(t)}{z(t)}\left(1-\frac{x(t)}{z(t)}\right)}\,\mathrm{d}w(t)+\left[c^D\,y(t)-c^A\,x(t)\right]\mathrm{d}t,\\
\mathrm{d}y(t) &= \left[c^A\,x(t)-c^D\,y(t)\right]\mathrm{d}t,\\
\mathrm{d}z(t) &= \left[c^D\,(1-z(t))-c^A\,z(t)\right]\mathrm{d}t.
\end{aligned}
\tag{A.16}
$$

Since $z(t)$ is the fraction of active individuals in the population, $1-z(t)$ is the fraction of dormant individuals in the population. Therefore the equivalent of the parameter K in Appendix A.1 is $K(t) = (1-z(t))/z(t)$. Moreover, $x(t)/z(t)$ is the fraction of $\heartsuit$-individuals in the active population at time t and $y(t)/(1-z(t))$ is the fraction of $\heartsuit$-individuals in the dormant population at time t. The last line of (A.16) is an autonomous differential equation whose solution converges to

$$z^* = \frac{1}{1+\frac{c^A}{c^D}} \tag{A.17}$$

exponentially fast. After this transition period we can replace $z(t)$ by z^*, and we see that $K^* = c^A/c^D$.

Time is to be scaled by the total number of active *and* dormant individuals, instead of the total number of active individuals only:

$$\begin{aligned} x(t) &= \frac{\text{number of active individuals of type } \heartsuit}{\text{total number of individuals}},\\ y(t) &= \frac{\text{number of dormant individuals of type } \heartsuit}{\text{total number of individuals}}. \end{aligned} \tag{A.18}$$

To compare the Moran model with a 1-colour seed-bank with the Fisher-Wright model with a 1-colour seed-bank, we look at the variables

$$\bar{x}(t) = \left(1+\frac{c^A}{c^D}\right) x\left(\frac{t}{1+\frac{c^A}{c^D}}\right), \qquad \bar{y}(t) = \left(1+\frac{c^A}{c^D}\right)\left(\frac{c^D}{c^A}\right) y\left(\frac{t}{1+\frac{c^A}{c^D}}\right). \tag{A.19}$$

After a short transition period in which $z(t)$ tends to z^*, we see that by setting

$$K = K^* = \frac{c^A}{c^D}, \qquad e = \frac{c^D}{c^A}\frac{c^A c^D}{c^A+c^D}, \tag{A.20}$$

we obtain

$$\begin{aligned} \mathrm{d}\bar{x}(t) &= \sqrt{\bar{x}(t)(1-\bar{x}(t))}\,\mathrm{d}w(t) + Ke\,[\bar{y}(t)-\bar{x}(t)]\,\mathrm{d}t,\\ \mathrm{d}\bar{y}(t) &= e\,[\bar{x}(t)-\bar{y}(t)]\,\mathrm{d}t, \end{aligned} \tag{A.21}$$

which is the single-colony version of (2.4)–(2.5) but without migration. Migration can be added in the same way as was done in Appendix A.1.

§A.2.2 Alternative for Model 2: Two colours

We consider the following system:

- Total number of individuals: $N \in \mathbb{N}$.
- Two types: $\heartsuit$ and $\diamondsuit$.
- $X(t)$ is the number of $\heartsuit$-individuals in the active population at time t.
- $Y_1(t)$ is the number of $\heartsuit$-individuals of colour 1 in the dormant population at time t.

- $Y_2(t)$ is the number of $\heartsuit$-individuals of colour 2 in the dormant population at time t.

- $Z_{D_1}(t)$ is the number of dormant individuals of colour 1 at time t (either $\heartsuit$ or $\diamondsuit$).

- $Z_{D_2}(t)$ is the number of dormant individuals of colour 2 at time t. (either $\heartsuit$ or $\diamondsuit$).

Note that the number of active individuals at time t (either $\heartsuit$ or $\diamondsuit$) is given by $Z_A(t) = N - Z_{D_1}(t) - Z_{D_2}(t)$. Since the number of individuals N is constant during the evolution, $Z_A(t)$ can be derived from $Z_{D_1}(t)$ and $Z_{D_2}(t)$. Each active individual resamples at rate 1, and becomes dormant at rate ϵ. When an individual becomes dormant, it gets either colour 1 with probability p_1 or colour 2 with probability p_2, where $p_1, p_2 \in (0,1)$ and $p_1 + p_2 = 1$. For ease of notation, we denote the rate to become dormant with colour 1 by $\epsilon_1 = \epsilon \cdot p_1$ and the rate to become dormant with colour 2 by $\epsilon_2 = \epsilon \cdot p_2$. A dormant individual with colour 1 becomes active at rate δ_1, a dormant individual with colour 2 becomes active at rate δ_2. Thus, the transition rates for $(X(t), Y_1(t), Y_2(t), Z_{D_1}(t), Z_{D_2}(t))$ are:

- $(i,j,k,l,m) \to (i+1,j,k,l,m)$ at rate $(N-l-m-i)\frac{i}{N-l-m}$.
- $(i,j,k,l,m) \to (i-1,j,k,l,m)$ at rate $i\frac{(N-l-m-i)}{N-l-m}$.
- $(i,j,k,l,m) \to (i-1,j+1,k,l+1,m)$ at rate $\epsilon_1 i$.
- $(i,j,k,l,m) \to (i+1,j-1,k,l-1,m)$ at rate $\delta_1 j$.
- $(i,j,k,l,m) \to (i-1,j,k+1,l,m+1)$ at rate $\epsilon_2 i$.
- $(i,j,k,l,m) \to (i+1,j,k-1,l,m-1)$ at rate $\delta_2 k$.
- $(i,j,k,l,m) \to (i,j,k,l+1,m)$ at rate $\epsilon_1(N-l-m-i)$.
- $(i,j,k,l,m) \to (i,j,k,l,m+1)$ at rate $\epsilon_2(N-l-m-i)$.
- $(i,j,k,l,m) \to (i,j,k,l-1,m)$ at rate $\delta_1(l-j)$.
- $(i,j,k,l,m) \to (i,j,k,l,m-1)$ at rate $\delta_2(m-k)$.

Proceeding in the same way as for the 1-colour seed-bank, we define the scaled variables

$$\begin{aligned} \bar{X}(t) &= \frac{1}{N}X(Nt), \quad \bar{Y}_1(t) = \frac{1}{N}Y_1(Nt), \quad \bar{Y}_2(t) = \frac{1}{N}Y_2(Nt), \\ \bar{Z}_{D_1}(t) &= \frac{1}{N}Z_{D_1}(Nt), \quad \bar{Z}_{D_2}(t) = \frac{1}{N}Z_{D_1}(Nt). \end{aligned} \tag{A.22}$$

We assume that there exist $c_1^A, c_2^A, c_1^D, c_2^D \in (0,\infty)$ such that

$$N\epsilon_1 = c_1^A, \qquad N\epsilon_2 = c_2^A, \qquad N\delta_1 = c_1^D, \qquad N\delta_2 = c_2^D, \qquad N \in \mathbb{N}, \tag{A.23}$$

and further assume that

$$\begin{aligned}&\lim_{N\to\infty}\frac{i}{N}=x,\quad \lim_{N\to\infty}\frac{j}{N}=y_1,\quad \lim_{N\to\infty}\frac{k}{N}=y_2,\\&\lim_{N\to\infty}\frac{N-l-m}{N}=z_A\quad \lim_{N\to\infty}\frac{N-l-m}{N}=z_{D_2},\quad \lim_{N\to\infty}\frac{N-l-m}{N}=z_{D_1}.\end{aligned}\tag{A.24}$$

Using the same method of converging generators as for model 1, we obtain the following continuum limit:

$$\begin{aligned}\mathrm{d}x(t)&=\sqrt{z_A(t)\frac{z_A-x(t)}{z_A(t)}\frac{x(t)}{z_A(t)}}\,\mathrm{d}w(t)\\&\quad+\left[c_1^D y_1(t)-c_1^A x(t)\right]\mathrm{d}t+\left[c_2^D y_2(t)-c_2^A x(t)\right]\mathrm{d}t,\\\mathrm{d}y_1(t)&=\left[c_1^A x(t)-c_1^D y_1(t)\right]\mathrm{d}t,\\\mathrm{d}y_2(t)&=\left[c_2^A x(t)-c_2^D y_2(t)\right]\mathrm{d}t,\\\mathrm{d}z_A(t)&=\left[c_1^D z_{D_1}(t)-c_1^A z_A(t)+c_2^D z_{D_2}(t)-c_2^A z_A(t)\right]\mathrm{d}t,\\\mathrm{d}z_{D_1}(t)&=\left[c_1^A z_A(t)-c_1^D z_{D_1}(t)\right]\mathrm{d}t,\\\mathrm{d}z_{D_2}(t)&=\left[c_2^A z_A(t)-c_2^D z_{D_2}(t)\right]\mathrm{d}t.\end{aligned}\tag{A.25}$$

Note that the equation for $z_A(t)=1-z_{D_1}(t)-z_{D_2}(t)$ follows directly from the equations from $z_{D_1}(t)$ and $z_{D_2}(t)$. It is therefore redundant, but we use it for notational reasons. Again, we see that $z(t)=(z_A(t),z_{D_1}(t),z_{D_2}(t))$ is governed by an autonomous system of differential equations. Solving this system, we see that

$$\lim_{t\to\infty}z_A(t)=\frac{1}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}},\quad \lim_{t\to\infty}z_{D_1}(t)=\frac{\frac{c_1^A}{c_1^D}}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}},\quad \lim_{t\to\infty}z_{D_2}(t))=\frac{\frac{c_2^A}{c_2^D}}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}.\tag{A.26}$$

To compare the Moran model with a 2-colour seed-bank with the Fisher-Wright model with a 2-colour seed-bank, we look at the variables

$$\begin{aligned}\bar{x}(t)&=\left(1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}\right)x\left(\frac{t}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}\right),\\\bar{y}_1(t)&=\left(1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}\right)\left(\frac{c_1^D}{c_1^A}\right)y_1\left(\frac{t}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}\right),\\\bar{y}_2(t)&=\left(1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}\right)\left(\frac{c_2^D}{c_2^A}\right)y_2\left(\frac{t}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}\right).\end{aligned}\tag{A.27}$$

Defining

$$K_m=\frac{c_m^A}{c_m^D},\quad e_m=\frac{c_m^D}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}},\qquad m\in\{1,2\},\tag{A.28}$$

we see that, after a short transition period, the system becomes

$$\begin{aligned}
\mathrm{d}\bar{x}(t) &= \sqrt{\bar{x}(t)(1-\bar{x}(t))}\,\mathrm{d}w(t) + K_1 e_1\,[\bar{y}_2(t)-\bar{x}(t)]\,\mathrm{d}t + K_2 e_2\,[\bar{y}_1(t)-\bar{x}(t)]\,\mathrm{d}t,\\
\mathrm{d}\bar{y}_1(t) &= e_1\,[\bar{x}(t)-\bar{y}_1(t)]\,\mathrm{d}t,\\
\mathrm{d}\bar{y}_2(t) &= e_2\,[\bar{x}(t)-\bar{y}_2(t)]\,\mathrm{d}t,
\end{aligned} \tag{A.29}$$

which is the single-colony version of (2.12)–(2.13) with 2 colours and without migration. Note, in particular, that after $z(t)$ reaches the equilibrium point in (A.26), we have

$$K_m = \frac{\text{number of dormant individuals with colour } m}{\text{number of active individuals}}, \qquad m \in \{1,2\}. \tag{A.30}$$

It is instructive to show how the above result can also be derived with the help of *duality.* The argument that follows easily extends to an n-coloured seed-bank for any $n \in \mathbb{N}$ finite, to be considered in Appendix A.2.3. Recall from (A.25) that

$$\begin{aligned}
\mathrm{d}z_A(t) &= \left[c_1^D z_{D_1}(t) - c_1^A z_A(t) + c_2^D z_{D_2}(t) - c_2^A z_A(t)\right]\mathrm{d}t,\\
\mathrm{d}z_{D_1}(t) &= \left[c_1^A z_A(t) - c_1^D z_{D_1}(t)\right]\mathrm{d}t,\\
\mathrm{d}z_{D_2}(t) &= \left[c_2^A z_A(t) - c_2^D z_{D_2}(t)\right]\mathrm{d}t.
\end{aligned} \tag{A.31}$$

Let

$$\begin{aligned}
\bar{z}_A(t) &= \left(1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}\right) z_A(t),\\
\bar{z}_{D_1}(t) &= \left(1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}\right)\left(\frac{c_1^D}{c_1^A}\right) z_{D_1}(t),\\
\bar{z}_{D_2}(t) &= \left(1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}\right)\left(\frac{c_2^D}{c_2^A}\right) z_{D_2}(t).
\end{aligned} \tag{A.32}$$

Substitute (A.32) into (A.31), to obtain

$$\begin{aligned}
\mathrm{d}\bar{z}_A(t) &= c_1^A\,[\bar{z}_{D_1}(t)-\bar{z}_A(t)] + c_2^A\,[\bar{z}_{D_2}(t)-\bar{z}_A(t)]\,\mathrm{d}t,\\
\mathrm{d}\bar{z}_{D_1}(t) &= c_1^D\,[\bar{z}_A(t)-\bar{z}_{D_1}(t)]\,\mathrm{d}t,\\
\mathrm{d}\bar{z}_{D_2}(t) &= c_2^D\,[\bar{z}_A(t)-\bar{z}_{D_2}(t)]\,\mathrm{d}t.
\end{aligned} \tag{A.33}$$

To define a dual for the process $(\bar{z}_A(t), \bar{z}_{D_1}(t), \bar{z}_{D_2}(t))_{t\geq 0}$, let $(M(t))_{t\geq 0}$ be the continuous-time Markov chain on $\{A, D_1, D_2\}$ with transition rates

$$\begin{aligned}
A \to D_m &\text{ at rate } c_m^A, \quad m \in \{1,2\},\\
D_m \to A &\text{ at rate } c_m^D, \quad m \in \{1,2\}.
\end{aligned} \tag{A.34}$$

Consider l independent copies of $(M(t))_{t\geq 0}$, evolving on the same state space $\{A, D_1, D_2\}$. Let $(L(t))_{t\geq 0} = (L_A(t), L_{D_1}(t), L_{D_2}(t))_{t\geq 0}$ be the process that counts how many copies of $M(t)$ are on site $\{A\}$, $\{D_1\}$ and $\{D_2\}$ at time t. Let $l = m + n_1 + n_2$. Then $(L(t))_{t\geq 0}$ is the Markov process on $\mathbb{N}_0^3$ with transition rates

$$(m, n_1, n_2) \to \begin{cases} (m-1, n_1+1, n_2) & \text{at rate } mc_1^A,\\ (m-1, n_1, n_2+1) & \text{at rate } mc_2^A,\\ (m+1, n_1-1, n_2) & \text{at rate } n_1 c_1^D,\\ (m+1, n_1, n_2-1) & \text{at rate } n_2 c_2^D. \end{cases} \tag{A.35}$$

Note that $L_A(t) + L_{D_1}(t) + L_{D_2}(t) = L_A(0) + L_{D_1}(0) + L_{D_2}(0) = m + n_1 + n_2 = l$. Define $H\colon \mathbb{R}^3 \times \mathbb{N}_0^3 \to \mathbb{R}$ by

$$H((\bar{z}_A, \bar{z}_{D_1}, \bar{z}_{D_2}), (m, n_1, n_2)) := \bar{z}_A^m \bar{z}_{D_1}^{n_1} \bar{z}_{D_2}^{n_2} \tag{A.36}$$

Using the generator criterion [48, Proposition 1.2], we see that, for all $t \geq 0$,

$$\begin{aligned} &\mathbb{E}\left[H((\bar{z}_A(t), \bar{z}_{D_1}(t), \bar{z}_{D_2}(t)), (m(0), n_1(0), n_2(0)))\right] \\ &\quad = \mathbb{E}\left[H((\bar{z}_A(0), \bar{z}_{D_1}(0), \bar{z}_{D_2}(0)), (m(t), n_1(t), n_2(t)))\right]. \end{aligned} \tag{A.37}$$

Therefore $(L(t))_{t\geq 0}$ and $(\bar{z}(t))_{t\geq 0}$ are dual to each other with duality function H.

Since $(M(t))_{t\geq 0}$ is a irreducible and recurrent, we can define

$$\begin{aligned} \pi_A &= \lim_{t\to\infty} \mathbb{P}(M(t) = A) = \frac{1}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}, \\ \pi_{D_1} &= \lim_{t\to\infty} \mathbb{P}(M(t) = D_1) = \frac{\frac{c_1^A}{c_1^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}, \\ \pi_{D_2} &= \lim_{t\to\infty} \mathbb{P}(M(t) = D_2) = \frac{\frac{c_2^A}{c_2^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}. \end{aligned} \tag{A.38}$$

Using the duality relation in (A.37) together with (A.38) and (A.32), we find

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}[\bar{z}_A(t)] &= \pi_A \bar{z}_A(0) + \pi_{D_1} \bar{z}_{D_1}(0) + \pi_{D_2} \bar{z}_{D_2}(0) \\ &= \frac{1}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}\, \bar{z}_A(0) + \frac{\frac{c_1^A}{c_1^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}\, \bar{z}_{D_1}(0) + \frac{\frac{c_2^A}{c_2^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}\, \bar{z}_{D_2}(0) \\ &= z_A(0) + z_{D_1}(0) + z_{D_2}(0) = 1. \end{aligned} \tag{A.39}$$

Using the duality relation in(A.37) once more, we get

$$\lim_{t\to\infty} \mathbb{E}[\bar{z}_A(t)] = \lim_{t\to\infty} \mathbb{E}[\bar{z}_{D_1}(t)] = \lim_{t\to\infty} \mathbb{E}[\bar{z}_{D_2}(t)] = 1. \tag{A.40}$$

Computing the limiting second moment $\lim_{t\to\infty} \mathbb{E}[\bar{z}_A(t)^2]$ by duality, we obtain

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}[\bar{z}_A(t)^2] &= \lim_{t\to\infty} \sum_{\substack{i,j\in\\ \{A,D_1,D_2\}}} \mathbb{P}(M_t^1 = i)\, \bar{z}_i(0)\, \mathbb{P}(M_t^2 = j)\, \bar{z}_j(0) \\ &= \sum_{i\in\{A,D_1,D_2\}} \pi_i \bar{z}_i(0) \sum_{j\in\{A,D_1,D_2\}} \pi_j \bar{z}_j(0) = 1. \end{aligned} \tag{A.41}$$

Similarly, we find $\lim_{t\to\infty} \mathbb{E}[\bar{z}_{D_1}(t)^2] = 1$ and $\lim_{t\to\infty} \mathbb{E}[\bar{z}_{D_2}(t)^2] = 1$. Combining (A.40) and (A.41), we find

$$\lim_{t\to\infty} \bar{z}_A(t) = \lim_{t\to\infty} \bar{z}_{D_1}(t) = \lim_{t\to\infty} \bar{z}_{D_2}(t) = 1. \tag{A.42}$$

Hence we conclude that

$$\lim_{t\to\infty} z_A(t) = \frac{1}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}, \lim_{t\to\infty} z_{D_1}(t) = \frac{\frac{c_1^A}{c_1^D}}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}, \lim_{t\to\infty} z_{D_2}(t) = \frac{\frac{c_2^A}{c_2^D}}{1+\frac{c_1^A}{c_1^D}+\frac{c_2^A}{c_2^D}}. \tag{A.43}$$

Continuing as in (A.27), we again find the single-colony version of (2.12)-(2.13) with 2 colours and no migration.

§A.2.3 Alternative for Model 2: Three or more colours

The argument in Appendix A.2.2 can be extended to an $\mathfrak{m} \in \mathbb{N}$-colour seed-bank, by introducing sequences of variables $(Y_m(t))_{m=0}^{\mathfrak{m}}$ and $(Z_m(t))_{m=0}^{\mathfrak{m}}$ that count the number of $\heartsuit$-individuals in the colour-m seed-bank at time t, respectively, the total number of individuals in the colour-m seed-bank at time t. Let $\epsilon > 0$ be the total rate at which an active individual becomes dormant, and define a probability vector $(p_m)_{m=0}^{\mathfrak{m}}$ such that $\epsilon_m = \epsilon p_m$ is the rate at which an active individual becomes dormant with colour m. Let δ_m be the rate at which m-dormant individuals become active. Via the same line of argument as in Appendix A.2.2, we see that the equivalent of (A.25) reads

$$\begin{aligned}
\mathrm{d}x(t) &= \sqrt{z_A(t)\frac{z_A - x(t)}{z_A(t)}\frac{x(t)}{z_A(t)}}\,\mathrm{d}w(t) + \sum_{m=0}^{\mathfrak{m}} \left[c_m^D y_m(t) - c_m^A x(t)\right] \mathrm{d}t,\\
\mathrm{d}y_m(t) &= \left[c_m^A x(t) - c_m^D y_m(t)\right] \mathrm{d}t,\\
\mathrm{d}z_A(t) &= \sum_{m=0}^{\mathfrak{m}} \left[c_m^D z_{D_m}(t) - c_m^A z_A(t)\right] \mathrm{d}t,\\
\mathrm{d}z_{D_m}(t) &= \left[c_m^A z_A(t) - c_m^D z_{D_m}(t)\right] \mathrm{d}t, \qquad 0 \le m \le N.
\end{aligned} \tag{A.44}$$

Solving the autonomous system describing $z(t) = (z_A(t), (z_{D_m}(t))_{m=0}^N)$ via duality, and subsequently substituting into (A.44) the variables

$$\begin{aligned}
\bar{x}(t) &= \left(1+\sum_{n=0}^{\mathfrak{m}}\frac{c_n^A}{c_n^D}\right) x\left(\frac{t}{1+\sum_{n=0}^{\mathfrak{m}}\frac{c_n^A}{c_n^D}}\right),\\
\bar{y}_m(t) &= \left(1+\sum_{n=0}^{\mathfrak{m}}\frac{c_n^A}{c_n^D}\right)\left(\frac{c_m^D}{c_m^A}\right) y_m\left(\frac{t}{1+\sum_{n=0}^{\mathfrak{m}}\frac{c_n^A}{c_n^D}}\right), \qquad 0 \le m \le N,
\end{aligned} \tag{A.45}$$

we find the single-colony version of (2.12)–(2.13) with N-colours and no migration. Migration can be added as in Appendix A.1.

It is straightforward to derive the version (2.12)–(2.13) with N-colours and M colonies. Afterwards we can let $N, M \to \infty$ and use convergence of generators, to find (2.12)–(2.13). The limit is unproblematic because we are interested in finite time horizons only.

§A.3 Successful coupling

To prove Lemma 3.2.11 we proceed as in [14], with minor adaptations. The notation used in this appendix is the same as in Section 3.2.3. For model 1 we write down the full proof. The proof holds works for model 2 and 3 by invoking the colours $m \in \mathbb{N}_0$ and the SSDE in (2.12)–(2.13), respectively, (2.18)–(2.19).

Proof of Lemma 3.2.11. The proof consists of 5 steps.

Step 1. If $z \in E$ with $x_i = 0$ and $x_k > 0$ for some $k \neq i$, then

$$\mathbb{P}_z\left(\exists\, t^* > 0 \text{ such that } x_i(t) = 0 \;\forall\, t \in [0, t^*]\right) = 0. \tag{A.46}$$

Proof. Suppose that z is such that $x_i = 0$, but $x_k > 0$ for some $i, k \in \mathbb{G}$. By (2.4),

$$x_i(t) = \int_0^t \sum_{j\in\mathbb{G}} a(i,j)[x_j(s) - x_i(s)]\,\mathrm{d}s + \int_0^t Ke[y_i(s) - x_i(s)]\,\mathrm{d}s + \int_0^t \sqrt{g(x_i(s))}\,\mathrm{d}w_i(s). \tag{A.47}$$

Suppose that there exists a $T > 0$ such that $x_i(t) = 0$ for all $t \in [0, T]$, and therefore $g(x_i(t)) = 0$. Then we obtain for all $t \in [0, T]$ that

$$\int_0^t \sum_{j\in\mathbb{G}} a(i,j) x_j(s)\,\mathrm{d}s + \int_0^t Key_i(s)\,\mathrm{d}s = 0. \tag{A.48}$$

Hence, by path continuity of $(Z(t))_{t\geq 0}$, we see that $y_i(t) = 0$ for all $t \in [0, T]$, as well as $x_j(t) = 0$ for all $j \in \mathbb{G}$ such that $a(i,j) > 0$. Repeating this argument, we obtain by irreducibility of $a(\cdot,\cdot)$ that $x_k(t) = 0$ for all $k \in \mathbb{G}$ and hence $y_k(t) = 0$ for all $k \in \mathbb{G}$. By path continuity, this contradicts the assumption that $x_k(0) > 0$. We conclude that (A.46) holds. □

Step 2. If $\bar{z} \in E \times E$ and $g(x_i^1) \neq g(x_i^2)$, then for all j,

$$\hat{\mathbb{P}}_{\bar{z}}\left(\exists\, t^* > 0 \text{ such that } \Delta_j(t) = 0 \;\forall\, t \in [0, t^*]\right) = 0. \tag{A.49}$$

Proof. Note that the SSDE in (2.4)–(2.5) can be rewritten as

$$\begin{aligned} \mathrm{d}z_{(i,R_i)}(t) = &\sum_{(j,R_j)\in\mathbb{G}\times\{A,D\}} b^{(1)}((i,R_i),(j,R_j))[z_{(j,R_j)}(t) - z_{(i,R_i)}(t)]\,\mathrm{d}t \\ &\qquad + \sqrt{g(z_{(i,R_i)}(t))}\; 1_{\{R_i = A\}}\,\mathrm{d}w_i(t), \\ &\forall\,(i,R_i) \in \mathbb{G}\times\{A,D\}, \end{aligned} \tag{A.50}$$

with $b^{(1)}(\cdot,\cdot)$ defined as in (2.31).

Suppose that $\bar{z}$ is such that $g(x_i^1) \neq g(x_i^2)$. Suppose there exist a $T > 0$ such that $\Delta_j(t) = 0$ for all $t \in [0, T]$. Then also $\sqrt{g(x_j^1(t))} - \sqrt{g(x_j^2(t))} = 0$ for all $t \in [0, T]$.

Using (A.50) on $\Delta_j(t) = z^1_{(j,A)}(t) - z^2_{(j,A)}(t)$, we obtain

$$0 = \int_0^t \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \times \Big[\Big(z^1_{(k,R_k)}(s) - z^2_{(k,R_k)}(s)\Big) - \Big(z^1_{(j,R_j)}(s) - z^2_{(j,R_j)}(s)\Big)\Big]\,\mathrm{d}s. \tag{A.51}$$

Hence

$$\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \Big[\Big(z^1_{(k,R_k)}(t) - z^2_{(k,R_k)}(t)\Big) - \Big(z^1_{(j,R_j)}(t) - z^2_{(j,R_j)}(t)\Big)\Big] = 0 \qquad \forall\, t\in[0,T]. \tag{A.52}$$

Using (A.50), we can write the SDE for

$$\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \Big[\Big(z^1_{(j,R_j)}(t) - z^2_{(j,R_j)}(t)\Big) - \Big(z^1_{(i,R_i)}(t) - z^2_{(i,R_i)}(t)\Big)\Big], \tag{A.53}$$

which yields that, for all $t\in[0,T]$,

$$\begin{aligned} &-\int_0^t \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \Big(\sqrt{g(z^1_{k,R_k}(s))} - \sqrt{g(z^2_{k,R_k}(s))}\Big) 1_{\{R_k=A\}}\,\mathrm{d}w_k(s) \\ &= \int_0^t \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1),2}((j,A),(l,R_l)) \\ &\qquad\times \Big[\Big(z^1_{(j,R_j)}(s) - z^2_{(j,R_j)}(s)\Big) - \Big(z^1_{(i,R_i)}(s) - z^2_{(i,R_i)}(s)\Big)\Big]\,\mathrm{d}s, \end{aligned} \tag{A.54}$$

where $b^{(1),2}(\cdot,\cdot)$ is the 2-step kernel of $b^{(1)}(\cdot,\cdot)$.

The two process in the right-hand side form a process of bounded variation, while the process in the left-hand side is a continuous square-integrable martingale, whose quadratic variation is given by

$$\int_0^t \sum_{k\in\mathbb{G}} a(j,k)^2 \Big(\sqrt{g(x^1_k(s))} - \sqrt{g(x^2_k(s))}\Big)^2\,\mathrm{d}s. \tag{A.55}$$

Since a square-integrable martingale of bounded variation is constant, it follows that (A.55) equals 0. Hence, for all k such that $a(j,k)>0$, it follows that $g(x^1_k(t)) = g(x^2_k(t))$ for all $t\in[0,T]$. Moreover, the right-hand side of (A.54) is equal to 0. Iterating the right-hand side of (A.54) further, we find by the irreducibility of $a(\cdot,\cdot)$ that $g(x^1_i(t)) = g(x^2_i(t))$ for all $t\in[0,T]$, which contradicts the assumption on $\bar z$ that $g(x^1_i(0)) \neq g(x^2_i(0))$. Hence we find that there does not exist a $T>0$ such that $\Delta_j(t) = 0$ for all $t\in[0,T]$. □

Step 3. If $\bar z\in E\times E,\ i,k\in\mathbb{G}$ and $g(x^1_i) = g(x^2_i)$ with $\Delta_i<0$ and $\Delta_k>0$ for some $k\neq i$, then

$$\hat{\mathbb{P}}_{\bar z}\left(\exists\, t^*\in[0,\tfrac12]\colon\ \Delta_i(t^*)<0,\ \Delta_k(t^*)>0,\ g(x^1_i(t^*))\neq g(x^2_i(t^*))\right) > 0. \tag{A.56}$$

Proof. Note that by assumption we have $x_i^1 < 1$ and $x_k^1 > 0$. Let $t_0 \in [0, \frac{1}{4}]$. If $x_i^1 > 0$, then set $t_0 = 0$. Otherwise, by Step 1 and path continuity, we find with probability 1 a $t_0 \in [0, \frac{1}{4}]$ such that $x_i^1(t_0) > 0$, $\Delta_i(t_0) < 0$ and $\Delta_k(t_0) > 0$. Let $\tilde{z} = \bar{z}(t_0)$. By the existence of t_0 and the Markov property, it is enough to prove that

$$\hat{\mathbb{P}}_{\tilde{z}}\left(\exists\, t^* \in [0, \tfrac{1}{4}] \colon\ \Delta_i(t^*) < 0,\ \Delta_k(t^*) > 0,\ g(x_i^1(t^*)) \neq g(x_i^2(t^*))\right) > 0 \tag{A.57}$$

in order to prove (A.56). Define the following two martingales:

$$M_i(t) = \int_0^t \sqrt{g(x_i^1(s))}\, \mathrm{d}w_i(s), \tag{A.58}$$

$$M_k(t) = \int_0^t \left(\sqrt{g(x_k^1(s))} - \sqrt{2g(x_k^2(s))}\right) \mathrm{d}w_k(s). \tag{A.59}$$

Their corresponding quadratic variation processes are given by

$$\langle M_i(t)\rangle = \int_0^t g(x_i(s))\, \mathrm{d}s, \tag{A.60}$$

$$\langle M_k(t)\rangle = \int_0^t \left(\sqrt{g(x_k^1(s))} - \sqrt{2g(x_k^2(s))}\right)^2 \mathrm{d}s. \tag{A.61}$$

By Knight's theorem (see [62, Theorem V.1.9 p.183]), we can write $M_i(t)$ and $M_k(t)$ as time-transformed Brownian motions:

$$M_i(t) = w_i\left(\langle M_i(t)\rangle\right), \tag{A.62}$$

$$M_k(t) = w_k\left(\langle M_k(t)\rangle\right). \tag{A.63}$$

We may assume that $g(\tilde{x}_i^1) = g(\tilde{x}_i^2)$, otherwise we can set $t^* = 0$. Recall that $0 < \tilde{x}_i^1 < 1$, $\tilde{\Delta}_i < 0$ and $\tilde{\Delta}_k > 0$, and, since $0 < g(\tilde{x}_i^1) = g(\tilde{x}_i^2)$, also $\tilde{x}_i^2 < 1$. Choose an $\epsilon \in (0, \frac{1}{15})$ such that $\tilde{x}_i^1, \tilde{x}_i^2 \in [5\epsilon, 1 - 5\epsilon]$, $-\tilde{\Delta}_i > 5\epsilon$ and $\tilde{\Delta}_k > 5\epsilon$. Let $\xi \in (0, \epsilon)$ be such that $g(\xi) < \min\{g(u) \colon \epsilon \le u \le 1 - \epsilon\}$, and set $c_1 = \min\{g(u) \colon \xi \le u \le 1 - \xi\}$ and $c_2 = \|g\|$. Then we can make the following estimates:

$$\langle M_i(t)\rangle \le c_2 t \qquad \langle M_k(t)\rangle \le c_2 t,\ t \ge 0, \tag{A.64}$$

$$\langle M_i(t)\rangle \ge c_1 t \qquad \text{for } t \ge 0 \text{ such that } x_i(s) \in [\xi, 1-\xi]\ \forall\, s \in [0, t]. \tag{A.65}$$

Define $c_3 = \min\{\frac{\xi}{2Ke}, \frac{\xi}{2}\}$. Fix $T \in [0, c_3]$ and define

$$\begin{aligned} \Omega_0 &= \left\{ \min_{t \in [0, c_1 T]} w_i(t) < -1,\ \max_{t \in [0, c_2 T]} w_i(t) < \epsilon,\ \max_{t \in [0, c_2 T]} |w_k(t)| < \epsilon \right\}, \\ \Omega_1 &= \left\{ \exists t^* \in [0,1] \text{ such that } \Delta_i(t^*) < 0,\ \Delta_k(t^*) > 0,\ g(x_i^1(t^*)) = g(x_i^2(t^*)) \right\}. \end{aligned} \tag{A.66}$$

Note that $\mathbb{P}(\Omega_0) > 0$. Therefore it suffices that $\Omega_0 \subset \Omega_1$.

We start by checking the conditions Δ_k. Using (2.4), we can write

$$\begin{aligned} \Delta_k(t) = \Delta_k(0) &+ \int_0^t \sum_{l \in \mathbb{G}} a(k,l)(\Delta_l(s) - \Delta_k(s))\, \mathrm{d}s + \int_0^t Ke\left[\delta_k(s) - \Delta_k(s)\, \mathrm{d}s\right] \\ &+ \int_0^t \left(\sqrt{g(x_k^1(s))} - \sqrt{2g(x_k^2(s))}\right)^2 \mathrm{d}w_k(s). \end{aligned} \tag{A.67}$$

Since $|\Delta_l(t)| \le 1$, $|\delta_k(t)| \le 1$ for all $t \ge 0$, and $M_k(t) = w_k(\langle M_k(t)\rangle)$ for $t \in [0,T]$, we may estimate

$$\Delta_k(t) > 5\epsilon - 2c_3 - 2Kec_3 - \epsilon = 2\epsilon. \tag{A.68}$$

So, on Ω_0, $\Delta_k(t) > 0$ for all $t \in [0,T]$. By expanding $x_i^1(t)$, we find

$$x_i^1(t) = x_i^1(0) + \int_0^t \sum_{l\in\mathbb{G}} a(i,l)(x_l^1(s) - x_i^1(s))\,\mathrm{d}s + \int_0^t Ke(y_i^1(s) - x_i^1(s))\,\mathrm{d}s + M_i(t), \tag{A.69}$$

so that on Ω_0 we have, for $t \in [0,T]$,

$$x_i^1(t) < 1 - 10\epsilon + c_3 + Kec_3 + \epsilon = 1 - 8\epsilon. \tag{A.70}$$

To check the conditions on $x_i^1(t)$ and $\Delta_i(t)$, we define the following random times:

$$\begin{aligned} \sigma &= \inf\{t \ge 0 : x_i^1(t) = \xi\}, \\ \tau &= \inf\left\{t > 0\colon\ g(x_i^1(t)) \neq g(x_i^2(t))\right\}. \end{aligned} \tag{A.71}$$

We will prove that, on Ω_0, we have $\sigma < \tau$ and $x_i^2(\tau) \ge x_i^1(\tau) + 3\epsilon$. To do so, we first prove that $\sigma < T$. Assume the contrary $\sigma \ge T$. Then by (A.70) we have $x_i^1(t) \in [\xi, 1-\xi]$ for all $t \in [0,T]$, which implies that $\min_{[0,T]} M_i(t) < -1$. Hence there exists a κ such that, by (A.69),

$$x_i^1(\kappa) < 1 - 10\epsilon + \epsilon - 1 < 0. \tag{A.72}$$

However, this contradicts the fact that $x_i^1 > 0$ for all $t \ge 0$. We conclude that $\sigma < T$. Now suppose that $\tau > \sigma$. Expanding Δ_i, we get, for $t < \tau$,

$$\Delta_i(t) = \Delta_i(0) + \int_0^t \sum_{l\in\mathbb{G}} a(i,l)(\Delta_l(s) - \Delta_i(s))\,\mathrm{d}s + \int_0^t Ke[\delta_i(s) - \Delta_i(s)]\,\mathrm{d}s, \tag{A.73}$$

which can be rewritten as

$$x_i^2(t) = x_i^1(t) - x_i^1(0) + x_i^2(0) - \int_0^t \sum_{l\in\mathbb{G}} a(i,l)[\Delta_l(s) - \Delta_i(s)]\,\mathrm{d}s - \int_0^t Ke[\delta_i(s) - \Delta_i(s)]\,\mathrm{d}s. \tag{A.74}$$

By (A.74), we obtain, for $t \in [0,\sigma]$,

$$\begin{aligned} x_i^2(t) &\le 1 - 5\epsilon + 2\epsilon + 2\epsilon = 1 - \epsilon, \\ x_i^2(t) &\ge x_i^1(t) + 5\epsilon - 2\epsilon \ge 3\epsilon, \end{aligned} \tag{A.75}$$

so $x_i^2(t) \in [\epsilon, 1-\epsilon]$ for $t \in [0,\sigma]$. But then $g(x_i^1(\sigma)) = g(\xi) < g(x_i^2(t))$ by the definition of ξ. Hence we obtain a contradiction and conclude that $\tau \le \sigma$. From (A.75) we obtain that $\Delta_i(t) < 0$ for all $t \in [0,\tau]$, which concludes the proof that $\Omega_0 \subset \Omega_1$. $\square$

Step 4. If $\bar{z} \in E \times E$ and $\Delta_i < 0, \Delta_j = 0,\ \Delta_k > 0$ for some i, j, k, then

$$\hat{\mathbb{P}}_{\bar{z}}\left(\exists\, t^* \in [0,1]\colon\ \Delta_i(t^*) < 0, \Delta_j(t^*) \neq 0, \Delta_k(t^*) > 0\right) > 0. \tag{A.76}$$

Proof. Suppose that $\bar{z}$ satisfies $\Delta_i < 0,\ \Delta_j = 0,\ \Delta_k > 0$. Define

$$\begin{aligned} \Gamma_0 &= \{\bar{z} \in E \times E : \Delta_i < 0, \Delta_j \neq 0, \Delta_k > 0\}, \\ \Gamma_1 &= \{\bar{z} \in E \times E : \Delta_i < 0,\ g(x_i^1) \neq g(x_i^2),\ \Delta_k > 0\}. \end{aligned} \tag{A.77}$$

By Step 3 and path continuity, there exists a $T \in [0, \frac{1}{2}]$ such that $\mathbb{P}^{\bar{z}}\left(\bar{z}(T) \in \Gamma_1\right) > 0$. By the Markov property,

$$\hat{\mathbb{P}}_{\bar{z}}\big(\exists\, t^* \in [0,1]\colon\ \bar{z}(t^*) \in \Gamma_0\big) \geq \int_{\Gamma_1} \hat{\mathbb{P}}_{\bar{z}}(\bar{z}(T) \in d\tilde{z})\, \hat{\mathbb{P}}_{\tilde{z}}\big(\exists t^* \in [0, \tfrac{1}{2}]\colon\ \bar{z}(t^*) \in \Gamma_0\big). \tag{A.78}$$

By path continuity, we can find for $\bar{z} \in \Gamma_1$ a t' such that, for all $t \leq t'$, $\Delta_i(t) < 0$, $\Delta_k(t) > 0$ and $g(x_i^1(t)) \neq g(x_i^2(t))$. By Step 2 there exists a $t^* < t'$ such that $\bar{z}(t^*) \in \Gamma_0$. Hence both probabilities in the integral on the right-hand side of (A.78) are positive. □

Step 5. Proof of Lemma 3.2.11.

Proof. Suppose that (3.147) holds for the pair i, j, and $a(j,k) > 0$, but (3.147) fails for the pair i, k. This implies that there exist $\epsilon_0 > 0$, $\delta_0 > 0$ and a positive increasing sequence $(t_n)_{n\in\mathbb{N}}$ of times with $t_n \to \infty$, such that

$$\lim_{t\to\infty} \hat{\mathbb{P}}_{\bar{z}}\left(\{\Delta_i(t) < \epsilon_0, \Delta_k(t) > \epsilon_0\} \cup \{\Delta_i(t) > \epsilon_0, \Delta_k(t) < \epsilon_0\}\right) > \delta_0. \tag{A.79}$$

By compactness of $E\times E$, there exists a subsequence t_{n_k} such that $\mathcal{L}(\bar{z}(t_{n_k}))$ converges and (A.79) holds. Let $\bar{\nu} = \lim_{k\to\infty} \mathcal{L}(\bar{z}(t_{n_k}))$. Then

$$\begin{aligned} \bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_j > \epsilon_0\} \cup \{\Delta_i > \epsilon_0, \Delta_j < \epsilon_0\}\right) &= 0, \\ \bar{\nu}\left(\{\Delta_j < \epsilon_0, \Delta_k > \epsilon_0\} \cup \{\Delta_j > \epsilon_0, \Delta_k < \epsilon_0\}\right) &= 0, \\ \bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_k > \epsilon_0\} \cup \{\Delta_i > \epsilon_0, \Delta_k < \epsilon_0\}\right) &> \delta_0. \end{aligned} \tag{A.80}$$

Assume without loss of generality that $\bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_k > \epsilon_0\}\right) > 0$. Hence, by (A.80),

$$\bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_k > \epsilon_0\}\right) = \bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_j \in (-\epsilon_0, \epsilon_0), \Delta_k > \epsilon_0\}\right) > 0. \tag{A.81}$$

For each $\bar{z} \in \{\Delta_i < \epsilon_0, \Delta_j \in (-\epsilon_0, \epsilon_0), \Delta_k > \epsilon_0\}$, Step 4 implies that

$$\hat{\mathbb{P}}_{\bar{z}}\left(\exists\, t^* \in [0,1]\colon\ \Delta_i(t^*) < 0, \Delta_j(t^*) \neq 0, \Delta_k(t^*) > 0\right) > 0, \tag{A.82}$$

and therefore, by (A.81),

$$\hat{\mathbb{P}}_{\bar{\nu}}\left(\exists\, t^* \in [0,1]\colon\ \Delta_i(t^*) < 0, \Delta_j(t^*) \neq 0, \Delta_k(t^*) > 0\right) > 0. \tag{A.83}$$

By path continuity, we can find $T \in [0,1]$ and $\epsilon > 0$ such that

$$\hat{\mathbb{P}}_{\bar{\nu}}\left(\Delta_i(T) < -\epsilon,\ |\Delta_j(T)|,\ \Delta_k(T) > \epsilon\right) > 0. \tag{A.84}$$

Let $\bar{\mu}(t_n) = \mathcal{L}(\bar{z}(t_n))$. Then, by the Markov property and (A.84),

$$\begin{aligned}&\liminf_{n\to\infty} \hat{\mathbb{P}}_{\bar{\mu}(t_n)}\left(\Delta_i(T) < -\epsilon,\ |\Delta_j(T)| > \epsilon,\ \Delta_k(T) > \epsilon\right)\\ &= \liminf_{n\to\infty} \hat{\mathbb{P}}_{\bar{\mu}(0)}\left(\Delta_i(T+t_n) < -\epsilon,\ |\Delta_j(T+t_n)| > \epsilon,\ \Delta_k(T+t_n) > \epsilon\right) > 0.\end{aligned} \tag{A.85}$$

However, this violates (3.147) for either i, j or j, k. We conclude that (A.79) fails and that (A.79) holds for i, k. By irreducibility, (A.79) holds for all $i, k \in \mathbb{G}$. □

§A.4 Bounded derivative of Lyapunov function

Recall from Section 3.2.3 that

$$\begin{aligned}h(t) = 2\sum_{j\in\mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\Delta_j(t)\}}\right]&\\ + 2Ke\,\hat{\mathbb{E}}\left[\left(|\Delta_i(t)| + |\delta_i(t)|\right) 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\delta_i(t)\}}\right].&\end{aligned} \tag{A.86}$$

In this section we show that $h'(t)$ exists for all $t > 0$ and is bounded. To do so, we need to get rid of the indicator in the expectations.

Let

$$h_{1,j}(t) = \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\Delta_j(t)\}}\right] \tag{A.87}$$

and

$$h_2(t) = 2Ke\,\hat{\mathbb{E}}\left[\left(|\Delta_i(t)| + |\delta_i(t)|\right) 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\delta_i(t)\}}\right]. \tag{A.88}$$

Then $h(t) = 2\sum_{j\in\mathbb{G}} a(i,j)h_{1,j}(t) + h_2(t)$. We show that $h_{1,j}(t)$ is differentiable with bounded derivative for $j \in \mathbb{G}$. The proof of the differentiability of $h_2(t)$ is similar. Fix $t \geq 0$. Note that

$$\begin{aligned}&\hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\Delta_j(t)\}}\right]\\ &= \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\Delta_j(t)\}} \mid |\Delta_i(t)| \neq 0, |\Delta_i(t)| \neq 0\right] \mathbb{P}\left(|\Delta_i(t)| \neq 0, |\Delta_j(t)| \neq 0\right)\\ &\quad + \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\mathrm{sgn}\,\Delta_i(t)\neq \mathrm{sgn}\,\Delta_j(t)\}} \mid |\Delta_i(t)| = 0 \text{ or } |\Delta_j(t)| = 0\right]\\ &\qquad \times \mathbb{P}\left(|\Delta_i(t)| = 0 \text{ or } |\Delta_j(t)| = 0\right).\end{aligned} \tag{A.89}$$

Since $\Delta_i(t)$ and $\Delta_j(t)$ have zero local time, the second term vanishes and $\mathbb{P}(|\Delta_i(t)| \neq 0, |\Delta_j(t)| \neq 0) = 1$. By continuity of $\Delta_i(t)$ and $\Delta_j(t)$, we can define sets

$$B_n = \left\{|\Delta_i(r)| > 0 \text{ and } |\Delta_j(r)| > 0, \forall r \in \mathcal{B}(t, \tfrac{1}{n})\right\}. \tag{A.90}$$

Then

$$\cdots \subset B_n \subset B_{n+1} \subset B_{n+2} \subset \cdots, \tag{A.91}$$

so

$$B_n = \bigcup_{i=0}^{n} B_i \tag{A.92}$$

and we define

$$B := \bigcup_{i=0}^{\infty} B_n = \lim_{n\to\infty} B_n. \tag{A.93}$$

Since $\mathbb{P}(|\Delta_i(t)| \neq 0, |\Delta_j(t)| \neq 0) = 1$, it follows that $\mathbb{P}(B) = 1$.

For each B_n, we have

$$B_n = C_n \cup C_n^c, \qquad C_n = \left\{\omega \in B_n \colon 1_{\{\operatorname{sgn}\Delta_i(r) \neq \operatorname{sgn}\Delta_j(r)\}} = 1,\ \forall r \in \mathcal{B}(t, \tfrac{1}{n})\right\}, \tag{A.94}$$

and, by the definition of B_n,

$$\cdots \subset C_n \subset C_{n+1} \subset C_{n+2} \subset \cdots \qquad\qquad \cdots \subset C_n^c \subset C_{n+1}^c \subset C_{n+2}^c \subset \cdots \tag{A.95}$$

Let $C = \bigcup_{i=0}^{\infty} C_i$ and $C^c = \bigcup_{i=0}^{\infty} C_i^c$ be such that $B = C \cup C^c$. Using (3.136), we obtain

$$\begin{aligned}
&\frac{1}{s}\left(h_{1,j}(t+s) - h_{1,j}(t)\right)\\
&= \frac{1}{s}\left[\hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t+s) \neq \operatorname{sgn}\Delta_j(t+s)\}}\right] - \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\Delta_j(t)\}}\right]\right]\\
&= \frac{1}{s}\left[\hat{\mathbb{E}}\left[|\Delta_j(t+s)|\, 1_{\{\operatorname{sgn}\Delta_i(t+s) \neq \operatorname{sgn}\Delta_j(t+s)\}} - |\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\Delta_j(t)\}}\middle| B\right]\right]\\
&= \frac{1}{s}\left[\hat{\mathbb{E}}\left[|\Delta_j(t+s)| - |\Delta_j(t)|\middle| C\right]\right]\mathbb{P}(C)\\
&= \frac{1}{s}\hat{\mathbb{E}}\left[\sum_{j\in\mathbb{G}} a(i,j)\int_t^{t+s} \operatorname{sgn}(\Delta_i(r))[\Delta_j(r) - \Delta_i(r)]\,\mathrm{d}r\middle| C\right]\mathbb{P}(C)\\
&\quad + \frac{1}{s}\hat{\mathbb{E}}\left[\int_t^{t+s} \operatorname{sgn}(\Delta_i(r))\left[\sqrt{g(x_i^1(r))} - \sqrt{g(x_i^2(r))}\right]\mathrm{d}w_i(r)\middle| C\right]\mathbb{P}(C)\\
&\quad + \frac{1}{s}\hat{\mathbb{E}}\left[Ke\int_t^{t+s} \operatorname{sgn}(\Delta_i(r))\left[\delta_i(r) - \Delta_i(r)\right]\mathrm{d}r\middle| C\right]\mathbb{P}(C)\\
&= \sum_{j\in\mathbb{G}} a(i,j)\hat{\mathbb{E}}\left[\frac{1}{s}\int_t^{t+s} \operatorname{sgn}(\Delta_i(r))[\Delta_j(r) - \Delta_i(r)]\,\mathrm{d}r\middle| C\right]\mathbb{P}(C)\\
&\quad + \frac{1}{s}\hat{\mathbb{E}}\left[\int_t^{t+s} \operatorname{sgn}(\Delta_i(r))\left[\sqrt{g(x_i^1(r))} - \sqrt{g(x_i^2(r))}\right]\mathrm{d}w_i(r)\middle| C\right]\mathbb{P}(C)\\
&\quad + \hat{\mathbb{E}}\left[Ke\,\frac{1}{s}\int_t^{t+s} \operatorname{sgn}(\Delta_i(r))\left[\delta_i(r) - \Delta_i(r)\right]\mathrm{d}r\middle| C\right]\mathbb{P}(C).
\end{aligned} \tag{A.96}$$

In the last equality, the first and third term are bounded, because $\Delta_i(t), \delta_i(t)$ and $\Delta_j(t)$ are continuous functions of t, and $\operatorname{sgn}(\Delta_i)$ is constant since we conditioned on the set C. Therefore, letting $s \to 0$, it follows from the fundamental theorem of calculus that these terms are bounded. The second term is more involved. Since, on the set C,

$$\operatorname{sgn}(\Delta_i(r))\left[\sqrt{g(x_i^1(r))} - \sqrt{g(x_i^2(r))}\right] \tag{A.97}$$

is a continuous function, we can rewrite the stochastic integral as a time-transformed

Brownian motion:

$$\begin{aligned}
&\frac{1}{s}\hat{\mathbb{E}}\left[\int_t^{t+s} \operatorname{sgn}\left(\Delta_i(r)\right)\left[\sqrt{g(x_i^1(r))}-\sqrt{g(x_i^2(r))}\right]\,\mathrm{d}w_i(r)\middle|C\right] \\
&\quad=\frac{1}{s}\hat{\mathbb{E}}\left[W\left(\int_0^{t+s}\left[\sqrt{g(x_i^1(r))}-\sqrt{g(x_i^2(r))}\right]^2\mathrm{d}r\right)\right. \\
&\qquad\left.-W\left(\int_0^{t}\left[\sqrt{g(x_i^1(r))}-\sqrt{g(x_i^2(r))}\right]^2\mathrm{d}r\right)\middle|C\right].
\end{aligned} \tag{A.98}$$

Since the normal distribution is differentiable with respect to its variance, we are done.

PART II

SPATIAL POPULATIONS WITH SEED-BANK ON THE HIERARCHICAL GROUP

This part is based on:
A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: renormalisation on the hierarchical group. *Preprint*, 2021

CHAPTER 4

Models and main results

§4.1 Background, goals and outline

§4.1.1 Background

Single colony with seed-bank. In populations with a *seed-bank*, individuals can temporarily become dormant and refrain from reproduction, until they can become active again. In [10] and [12] the evolution of a population evolving according to the two-type Fisher-Wright model with seed-bank was studied. Individuals move in and out of the seed-bank at prescribed rates. Outside the seed-bank individuals are subject to *resampling*, while inside the seed-bank their resampling is *suspended*. Both the long-time behaviour and the genealogy of the population were analysed in detail.

Seed-banks are observed in many taxa, including plants, bacteria and other micro-organisms. Typically, they arise as a response to unfavourable environmental conditions. The dormant state of an individual is characterised by low metabolic activity and interruption of phenotypic development (see e.g. [55]). After a varying and possibly large number of generations, a dormant individual can be resuscitated under more favourable conditions and reprise reproduction after having become active again. This strategy is known to have important implications for population persistence, maintenance of genetic variability and stability of ecosystems. It acts as a *buffer* against evolutionary forces such as genetic drift, selection and environmental variability.

Multiple colonies with seed-bank. In [43] we considered a *spatial* version of the two-type Fisher-Wright model with seed-bank in which individuals can *migrate* between colonies, organised into a *geographic space*, each having a seed-bank consisting of *multiple layers*, each with their own rate of moving in (becoming dormant) and moving out (waking up). We found that the presence of the seed-bank *enhances genetic diversity* compared to the spatial model without seed-bank. Interestingly, we found that the seed-bank can affect the longtime behaviour of the system both qualitatively and quantitatively.

In [43] we settled existence and uniqueness of the spatial model when the geographic space is $\mathbb{Z}^d$, $d \in \mathbb{N}$. We proved convergence to equilibrium, showed that there is a dichotomy between *coexistence* (= locally multi-type equilibria) and *clustering* (= locally mono-type equilibria), and identified the parameter regime for both. We found a change of the dichotomy due to the presence of the seed-bank. Without seed-bank, for migration in the domain of attraction of Brownian motion, clustering occurs in $d = 1, 2$ and coexistence in $d \geq 3$, i.e., the *critical dimension* for the dichotomy is $d = 2$. With seed-bank, however, clustering becomes more difficult and occurs in $d = 2$ only when the wake-up time of a typical individual in the seed-bank has finite mean, and in $d = 1$ only when the wake-up time has a sufficiently thin tail. In other words, the seed-bank has a tendency to *lower* the critical dimension.

In fact, in [43] we found that our technique of proof works for geographic spaces that are *arbitrary* countable Abelian groups endowed with the discrete topology. The reason is that the dichotomy can be formulated in terms of how the *degree* of the random walk that underlies the migration balances with the exponent of the *tail* of the typical wake-up time. This raises the question how we can better understand the behaviour of spatial models with seed-bank close to criticality.

In [44] we established the so-called *finite-systems scheme*, i.e., we identified how a finite truncation of the system (both in the geographic space and in the seed-bank) behaves as both the time and the truncation level tend to infinity, properly tuned together. We found that if the wake-up time has finite mean, then the scaling time is proportional to the volume of the system and there is a *single universality class* for the scaling limit, namely, the system moves through a succession of equilibria of the infinite system with a density that evolves according to a Fisher-Wright diffusion. On the other hand, we found that if the wake-up time has infinite mean, then the scaling time grows faster than the volume of the system, and there are *two universality classes* depending on how fast the truncation level of the seed-bank grows compared to the truncation level of the geographic space.

§4.1.2 Goals

In the present paper we take as geographic space the *hierarchical group* Ω_N of order N. The reason for this choice is that Ω_N allows for more detailed computations. At the same time, migration on Ω_N can be used to *approximate* migration on $\mathbb{Z}^d$ in the *hierarchical mean-field limit* $N \to \infty$. In particular, by playing with the migration kernel we can approximate two-dimensional migration in the sense of potential theory. We consider migration kernels that in the limit as $N \to \infty$ are *critically recurrent*, i.e., the degree of the class of hierarchical migrations that we consider in the present paper converges to 0, either from above or from below.

The present paper has three goals:

(1) We apply the results obtained in [43] to Ω_N with $N < \infty$ fixed. We again find that part of the coexistence regime without seed-bank shifts into the clustering regime with seed-bank when the average wake-up time of a typical individual is infinite.

(2) We analyse a *space-time renormalised* system in the limit as $N \to \infty$. Namely, we show that the block averages on successive space-time scales each perform a diffusion with a *renormalised diffusion function*. In other words, we establish a *multi-scale version of the finite-systems scheme*. Also, we compare the behaviour of the space-time renormalised system with seed-bank to the one analysed in [21] and [20] without seed-bank.

(3) We exhibit *universal behaviour* in the clustering regime close to criticality. To do so, we analyse the attracting orbits of the renormalisation transformation,

acting on the space of diffusion functions, that connects successive hierarchical levels. We show that, in the *clustering regime* and after appropriate scaling, the renormalised diffusion function converges to the Fisher-Wright diffusion function as we move up in the hierarchy, irrespective of the diffusion function controlling the resampling. This convergence shows that the hierarchical system exhibits *universality* on large space-time scales in terms of the scaling limit. For several subclasses of parameters we identify the scaling of the renormalised diffusion function, which reveals a delicate interplay between the parameters controlling the migration and the seed-bank. This rate in turn determines the speed at which *mono-type clusters grow in space and time.*

In the *coexistence regime*, universality does *not* hold and the equilibrium depends on the diffusion function. Since the seed-bank enhances genetic diversity, it may be expected that equilibrium correlations between far away colonies decay faster with seed-bank than without seed-bank, an issue that will not be addressed.

Remark 4.1.1 (More general types). Throughout the paper we consider the *two-type* Fisher-Wright model with seed-bank, in the *continuum limit* where the number of individuals per colony tends to infinity. The extension to a general type space, called the Fleming-Viot model (see [25]), requires only standard adaptations and will not be considered here. In what follows, we *only* work with continuum models. However, we motivate these models by viewing them as the *large-colony-size limit* of individual-based models. For earlier work on hierarchically interacting Fisher-Wright diffusions without seed-bank we refer the reader to [20, 21, 25, 22] and [5, 6, 26]. ■

§4.1.3 Outline

The present paper consist of two parts:

- **Part I: Model and main results.** Sections 4.2–4.5 collect the main propositions and theorems. In Section 4.2 we define the hierarchical model and state some basic properties: the *well-posedness* of the associated martingale problem (Proposition 4.2.6), the *duality relation* (Proposition 4.2.7), and the *clustering criterion* via duality (Proposition 4.2.12). These properties were all derived in [43]. In Section 4.3 we state our main results for $N < \infty$. In particular, we compute the scaling of the wake-up time and the migration kernel (Theorem 4.3.2) and identify the *clustering regime* in terms of the coefficients controlling the migration and the seed-bank under the assumption that these are asymptotically polynomial or pure exponential (Theorem 4.3.3)). In Section 4.4 we state our main results for $N \to \infty$, the hierarchical mean-field limit. In particular, we introduce *block averages on successive hierarchical space-time scales*, analyse their limiting dynamics (Theorems 4.4.2 and 4.4.4), offer a heuristic explanation how this limiting dynamics arises, introduce a path topology called the Meyer-Zheng topology that is needed for a proper formulation, and introduce an object called the interaction chain, which describes how the diffent hierarchical levels interact with each other. In Section 4.5 we identify the *orbit of the renormalisation transformation* in the clustering regime (Theorem 4.5.1), identify the rate of scaling for the renormalised diffusion function (Theorem 4.5.3), and link this scaling to the rate of growth of mono-type clusters.

- **Part II: Preparations and proofs.** Chapters 5–10 provide the proofs of the theorems stated in Part I. These proofs consist of a long series of propositions and lemmas needed to build up the argument. In Chapter 5 we prove our main results for $N < \infty$. In Chapter 6 we focus on the *mean-field model* (consisting of a single hierarchy) and, respectively, state and prove a number of results that serve as preparation. In Chapters 7–8 we consider extensions of the mean-field model (consisting of finitely hierarchies), which serve as further preparation. In Chapter 9 we use the results in Sections 6–8 to deal with the full hierarchical model (consisting of infinitely many hierarchies), and prove our main results for $N \to \infty$. In Chapter 10 we analyse the orbit of the renormalisation transformation controlling the multi-scaling. Appendix B.1 contains a technical computation needed for the identification of the clustering regime. Appendix B.2 contains a basic introduction to convergence of paths in the Meyer-Zheng topology, which is needed for the main theorems.

Part I contains all the main results and their interpretations, and can be read without reference to Part II.

§4.2 Introduction of model and basic properties

Section 4.2.1 introduces the model ingredients, Section 4.2.2 gives the evolution equations, Section 4.2.3 states the well-posedness, Section 4.2.4 introduces the dual and states the duality relation, while Section 4.2.5 formulates the dichotomy between clustering versus coexistence in terms of the dual.

§4.2.1 Model: geographic space Ω_N, hierachical group of order N

Single colony. Our building block is the single-colony Fisher-Wright model with seed-bank defined in [12]. In that model, each individual in the population carries one of two types, $\heartsuit$ or $\diamondsuit$, and each individual can be either active or dormant. Active individuals *resample* until they become dormant. Dormant individuals suspend resampling until they become actieve again. The repository for the dormant individuals is called the *seed-bank*. When an active individual resamples, it randomly chooses another active individual and *adopts its type*. When an active individual becomes dormant, it randomly chooses a dormant individual and forces it to becomes active, i.e., the active and the dormant population *exchange* individuals (see Fig. 4.1). This exchange guarantees that the sizes of the active and the dormant population stay fixed over time. During the swap both the active and the dormant individual *retain their type*.

The types of the active population evolve through resampling and through exchange with the dormant population. The types of the the dormant population evolve only through exchange with the active population. It was shown in [12] that in the large-colony-size limit, i.e., as the number of individuals per colony tends to infinity and time is speeded up by the size of the colony, the two quantities

- $x(t)$ = the fraction of active individuals of type $\heartsuit$ at time t,

- $y(t)$ = the fraction of dormant individuals of type $\heartsuit$ at time t,

satisfy the following system of coupled SDEs:

$$\begin{aligned} \mathrm{d}x(t) &= Ke\,[y(t)-x(t)]\,\mathrm{d}t + \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t), \\ \mathrm{d}y(t) &= e\,[x(t)-y(t)]\,\mathrm{d}t. \end{aligned} \tag{4.1}$$

Here, e denotes the rate at which an active individual *exchanges* with a dormant individual from the seed-bank, K denotes the relative size of the dormant population with respect to the active population, and $(w(t))_{t\geq 0}$ is a Brownian motion. The first term in the first equation describes the flow from the dormant population to the active population, the term in the second equation describes the flow from the active population to the dormant population, while the second term in the first equation describes the effect of resampling on the active population (see Fig. 4.1). Active individuals resample at rate 1. Since dormant individuals do not resample, we do not see such a term in the second equation. The formal derivation of the continuum equations can be found in [12] and in [43, Appendix A].

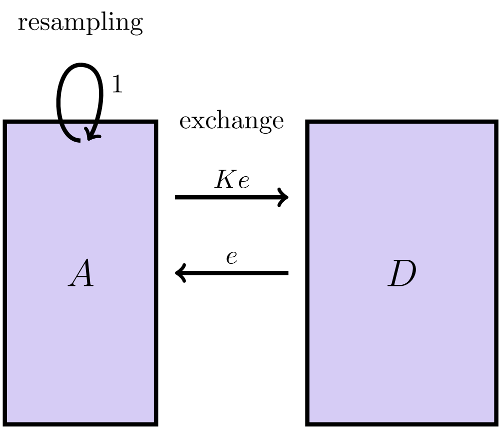

Figure 4.1: Active individuals resample at rate 1*. Active and dormant individuals exchange at rate e. The extra factor K arises from the fact that the dormant population is K times as large as the active population. Dormant individuals suspend resampling.*

Multi-colony. The present paper focuses on a *multi-colony* setting of the model described above, where the underlying geographic space is the hierarchical lattice of order N, given by ($\mathbb{N}_0 = \mathbb{N} \cup \{0\}$)

$$\Omega_N = \left\{ \xi = (\xi_k)_{k\in\mathbb{N}_0} \colon\ \xi_k \in \{0,1,\ldots,N-1\},\ \sum_{k\in\mathbb{N}_0} \xi_k < \infty \right\}, \tag{4.2}$$

which with addition modulo N becomes the hierarchical group of order N (see Fig. 4.2). The *hierarchical distance* on Ω_N is defined by

$$d_{\Omega_N}(\xi,\eta) = d_{\Omega_N}(0,\xi-\eta) = \min\left\{k\in\mathbb{N}_0 \colon\ \xi_l = \eta_l\ \forall\, l \geq k\right\}, \qquad \xi,\eta \in \Omega_N, \tag{4.3}$$

and is an ultra-metric, i.e.,

$$d_{\Omega_N}(\xi,\eta) \le \max\left\{d_{\Omega_N}(\xi,\zeta), d_{\Omega_N}(\eta,\zeta)\right\} \qquad \forall\, \xi,\eta,\zeta \in \Omega_N. \tag{4.4}$$

The choice of Ω_N as geographic space plays an important role for population models, and was first exploited in [65] in an attempt to formalise ideas coming from ecology. One interpretation is that the sequence $(\xi_k)_{k\in\mathbb{N}_0}$ encodes the 'address' of colony ξ: ξ_0 is the 'house', ξ_1 is the 'street', ξ_2 is the 'village', ξ_3 is the 'province', ξ_4 is the 'country', and so on. To describe the system on the hierarchical group we need three ingredients:

- Hierarchical migration.
- Layered seed-bank.
- Resampling rate.

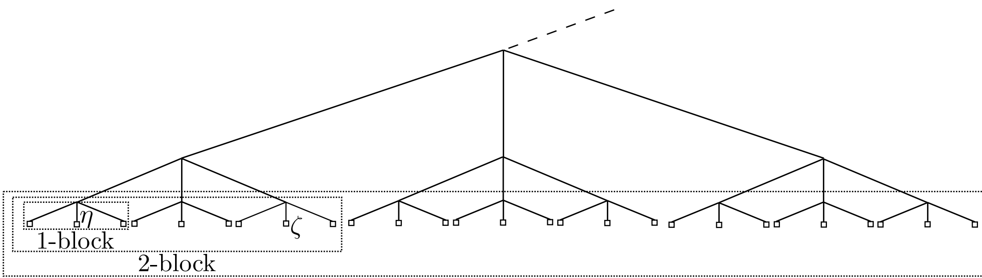

Figure 4.2: Close-ups of a 1-block, a 2-block and a 3-block in the hierarchical group of order $N = 3$. The elements of the group are the leaves of the tree (indicated by $\square$'s). The hierarchical distance between two elements in the group is the graph distance to the most recent common ancestor in the tree: $d_{\Omega_3}(\eta,\zeta) = 2$ for η and ζ in the picture.

- **Hierarchical migration**

We construct a migration kernel $a^{\Omega_N}(\cdot,\cdot)$ on the hierarchical group Ω_N built from a sequence of migration rates

$$\underline{c} = (c_k)_{k\in\mathbb{N}_0} \in (0,\infty)^{\mathbb{N}_0} \tag{4.5}$$

that do not depend on N. Individuals migrate as follows:

- For all $k \in \mathbb{N}$, each individual chooses at rate c_{k-1}/N^{k-1} the block of radius k around its present location and selects a colony uniformly at random from that block. Subsequently it selects an individual in this colony uniformly at random and adopts its type.

Note that the block of radius k contains N^k colonies, and that the migration kernel is therefore given by

$$a^{\Omega_N}(\eta,\xi) = \sum_{k\ge d_{\Omega_N}(\eta,\xi)} \frac{c_{k-1}}{N^{k-1}}\frac{1}{N^k}, \quad \eta,\xi\in\Omega_N, \eta\neq\xi, \qquad a^{\Omega_N}(\eta,\eta)=0, \quad \eta\in\Omega_N. \tag{4.6}$$

Throughout the paper, we assume that

$$\limsup_{k\to\infty} \frac{1}{k} \log c_k < \log N. \tag{4.7}$$

This guarantees that the total migration rate per individual is finite. Indeed, note that for every $\eta \in \Omega_N$,

$$\begin{aligned}
\sum_{\xi\in\Omega_N} a^{\Omega_N}(\eta,\xi) &= \sum_{\xi\in\Omega_N} \sum_{k\geq d_{\Omega_N}(\eta,\xi)} \frac{c_{k-1}}{N^{2k-1}} \\
&= \sum_{k\in\mathbb{N}} \left[\sum_{\xi\in\Omega_N} \mathbf{1}_{\{d_{\Omega_N}(\eta,\xi)\leq k\}} \right] \frac{c_{k-1}}{N^{2k-1}} \\
&= \sum_{k\in\mathbb{N}} \frac{c_{k-1}}{N^{k-1}},
\end{aligned} \tag{4.8}$$

which is finite because of (4.7).

Remark 4.2.1 (Degree of random walk). For a random walk on an Abelian group with time-t transition kernel $a_t^{\Omega_N}(\cdot,\cdot)$, the *degree* is defined as (see [19])

$$\delta = \sup\left\{ \zeta \in (-1,\infty) \colon \int_0^\infty \mathrm{d}t\, t^\zeta a_t^{\Omega_N}(0,0) < \infty \right\}. \tag{4.9}$$

The degree is said to be δ^+, respectively, δ^- when the integral is finite, respectively, infinite *at* the degree. If $\delta > 0$, then δ is called the *degree of transience*. If $\delta \in (-1,0)$, then $-\delta$ is called the *degree of recurrence*. If the degree is 0^-, then the random walk is called *critically recurrent*. (It would be interesting to have a version of (4.9) that includes a slowly varying function in front of the power t^ζ. However, such an extension appears not to have been explored in the literature.) ■

By playing with $\underline{c}$ and letting $N \to \infty$, we can approximate migration for which the corresponding random walk is critically recurrent, i.e., $\delta^- = 0$. In that case both the potential theory and the Green function for the hierarchical random walk have the same asymptotics as the potential theory and the Green function for a critically recurrent random walk on $\mathbb{Z}^2$ in the domain of attraction of Brownian motion. Therefore, by tuning $\underline{c}$ properly, we can mimic migration on the geographic space $\mathbb{Z}^2$ (for which $\delta^- = 0$), an idea that was exploited in [25], [22], [23], [41], [42].

- **Layered seed-bank**

To create a layered seed-bank, dormant individuals are labeled with a colour $m \in \mathbb{N}_0$. An active individual that becomes dormant is assigned a colour $m \in \mathbb{N}_0$. When an active individual becomes dormant with colour m, it exchanges with a dormant individual of colour m. This dormant individual becomes active, *loses* its colour, but *retains* its type. To describe the layered seed-bank we need two sequences

$$\begin{aligned}
\underline{K} &= (K_m)_{m\in\mathbb{N}_0} \in (0,\infty)^{\mathbb{N}_0}, \\
\underline{e} &= (e_m)_{m\in\mathbb{N}_0} \in (0,\infty)^{\mathbb{N}_0},
\end{aligned} \tag{4.10}$$

both not depending on N, which we interpret as follows:

- K_m is the relative size of the dormant population of colour m with respect to the active population, i.e.,

$$K_m = \frac{\text{size } m\text{-dormant population}}{\text{size active population}}. \tag{4.11}$$

- At rate $K_m \frac{e_m}{N^m}$ an active individual becomes dormant, is assigned colour m, and retains its type. At the same time a dormant individual with colour m becomes active, loses its colour, and retains its type. By defining the rates in this way, the layered structure of the seed-bank is tuned to the hierarchical structure of the geographic space.

By giving the seed-bank a layered structure, we are able to tune the distribution of the wake-up time, i.e., the time an individual spends in the seed-bank before waking up. In particular, we will see that a layered seed-bank enables us to model wake up times with a *fat tail*, while at the same time preserving the Markov property of the evolution.

Since active and dormant individuals *exchange*, K_m remains constant over time for all $m \in \mathbb{N}_0$. Throughout the paper we assume that

$$\limsup_{m\to\infty} \frac{1}{m} \log(K_m e_m) < \log N. \tag{4.12}$$

This guarantees that the total rate of exchange per individual, given by

$$\chi = \sum_{m\in\mathbb{N}_0} K_m \frac{e_m}{N^m}, \tag{4.13}$$

is finite. On the other hand, the relative size of the dormant population with respect to the active population

$$\rho = \sum_{m\in\mathbb{N}_0} K_m \tag{4.14}$$

can be either finite or infinite. We will see that $\rho < \infty$ and $\rho = \infty$ represent two *different regimes.*

• Resampling rate

To describe the resampling we use a diffusion function g that is taken from the set

$$\mathcal{G} = \Big\{ g(x)\colon\ [0,1] \to [0,\infty)\colon\ g(0) = g(1) = 0,\ g(x) > 0\ \forall\, x \in (0,1),\ g \text{ Lipschitz} \Big\}, \tag{4.15}$$

and think of $h(x) = g(x)/x(1-x)$ as the rate of resampling at type frequency x. The choice $g = dg_{\mathrm{FW}}$, $d \in (0,\infty)$, with $g_{\mathrm{FW}}(x) = x(1-x)$, $x \in [0,1]$, corresponds to Fisher-Wright resampling at rate d. We use a collection of independent Brownian motions

$$W = \big((w_\xi(t))_{t\geq 0}\big)_{\xi\in\Omega_N} \tag{4.16}$$

to describe the fluctuations of the type frequencies caused by the resampling in each colony.

§4.2.2 Evolution equations

- **Evolution of single colonies**

With the above three ingredients, we can now describe the evolution of the system. For $\xi \in \Omega_N$, define

$$\begin{aligned} x_\xi(t) &= \text{ the fraction of active individuals of type } \heartsuit \text{ at colony } \xi \text{ at time } t, \\ y_{\xi,m}(t) &= \text{ the fraction of } m\text{-dormant individuals of type } \heartsuit \text{ at colony } \xi \text{ and time } t. \end{aligned} \tag{4.17}$$

Note that $x_\xi(t) \in [0,1]$ and $y_{\xi,m}(t) \in [0,1]$ for all $\xi \in \Omega_N$, $m \in \mathbb{N}_0$, $t \geq 0$. Therefore the state space of a single colony is $\mathfrak{s} = [0,1] \times [0,1]^{\mathbb{N}_0}$, and the state space of the system is

$$S = \mathfrak{s}^{\Omega_N}. \tag{4.18}$$

Our object of interest is the random process taking values in S, written

$$(X^{\Omega_N}(t), Y^{\Omega_N}(t))_{t\geq 0}, \qquad (X^{\Omega_N}(t), Y^{\Omega_N}(t)) = \big(x_\xi(t), (y_{\xi,m}(t))_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N}, \tag{4.19}$$

whose components evolve according to the following SSDE (= system of stochastic differential equations):

$$\begin{aligned} \mathrm{d}x_\xi(t) &= \sum_{\eta\in\Omega_N} a^{\Omega_N}(\xi,\eta)[x_\eta(t) - x_\xi(t)]\,\mathrm{d}t + \sqrt{g(x_\xi(t))}\,\mathrm{d}w_\xi(t) \\ &\qquad + \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{N^m}\,[y_{\xi,m}(t) - x_\xi(t)]\,\mathrm{d}t, \\ \mathrm{d}y_{\xi,m}(t) &= \frac{e_m}{N^m}\,[x_\xi(t) - y_{\xi,m}(t)]\,\mathrm{d}t, \quad m \in \mathbb{N}_0, \qquad \xi \in \Omega_N. \end{aligned} \tag{4.20}$$

The first term in the first equation describes the evolution of the active population at colony ξ due to migration, the second term due to the resampling. The third term in the first equation and the term in the second equation describe the exchange between the active and the dormant population at colony ξ (see Fig. 4.3). Since dormant individuals are not subject to resampling or migration, the dynamics of the dormant population is completely determined by the exchange with the active population. For the *initial state* we assume that

$$\begin{aligned} &\mathcal{L}(X^{\Omega_N}(0), Y^{\Omega_N}(0)) = \mu^{\otimes\Omega_N} \\ &\text{with } \mathbb{E}^\mu[x_\xi(0)] = \theta_x,\ \mathbb{E}^\mu[y_{\xi,m}(0)] = \theta_{y_m} \text{ with } \lim_{m\to\infty} \theta_{y_m} = \theta \text{ for some } \theta \in [0,1]. \end{aligned} \tag{4.21}$$

The last assumption in (4.21), which in [43] was referred to as μ being *colour regular*, guarantees that for finite N the system in (4.20) converges to an ergodic equilibrium.

Remark 4.2.2. *[Notation]* Throughout the sequel we use lower case letters for *single components* and upper case letters for *systems of single components*. We exhibit the geographic space for the system, but suppress it from the components. ■

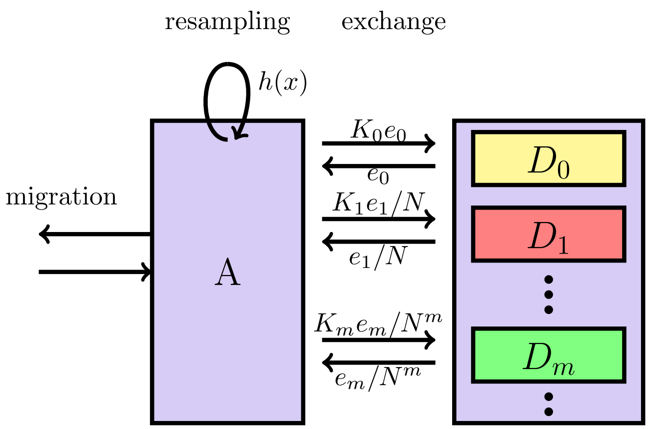

Figure 4.3: Active individuals (A) are subject to migration, resampling and exchange with dormant individuals (D). When active individuals become dormant they are assigned a colour (D_m, $m \in \mathbb{N}_0$), which they lose when they become active again. The resampling rate in the active state at type-♡ frequency x equals $h(x) = g(x)/x(1-x)$ with $g \in \mathcal{G}$ (e.g. for the standard Fisher-Wright diffusion the resampling rate is 1).

• **Evolution of block averages**

The choice of the migration kernel in (4.6) implies that, for every $k \in \mathbb{N}$, at rate $\asymp \frac{1}{N^k}$ individuals choose a space horizon of distance $k+1$ and subsequently choose a random colony from that space horizon. Therefore, in order to see interactions over a distance $k+1$, we need to speed up time by a factor N^k. A similar observation applies to the interaction with the seed-bank. Dormant individuals with colour k become active at rate $\asymp \frac{1}{N^k}$. Therefore, in order to see interactions with the k-dormant population, we need to speed up time by a factor N^k. To analyse the effective interaction on time scale N^k, we introduce *successive block averages* labelled by $k \in \mathbb{N}_0$.

Definition 4.2.3 (Block averages). For $k \in \mathbb{N}_0$, let
$B_k(0) = \{\eta \in \Omega_N \colon d_{\Omega_N}(0, \eta \leq k\}$ denote the k-block around 0. Define the k-block average around 0 at time $N^k t$ by

$$
\begin{aligned}
x_k^{\Omega_N}(t) &= \frac{1}{N^k} \sum_{\eta \in B_k(0)} x_\eta(N^k t), \\
y_{m,k}^{\Omega_N}(t) &= \frac{1}{N^k} \sum_{\eta \in B_k(0)} y_{\eta,m}(N^k t), \qquad m \in \mathbb{N}_0.
\end{aligned}
\tag{4.22}
$$

The k-block average represents the dynamics of the system on space-time scale k. ■

By translation invariance of the SSDE in (4.20), each $\xi \in \Omega_N$ can serve as the origin. In the remainder of the paper we consider without loss of generality the k-block average around $\xi = 0$, and suppress the center 0 from the notation.

Remark 4.2.4 (Notation). We use lower case letters for the block averages because they live in the space of components $\mathfrak{s} = [0,1] \times [0,1]^{\mathbb{N}_0}$. At the same time we exhibit the geographic space Ω_N for the block averages because they are functionals of the system of components (recall Remark 4.2.2). ■

Using Definition 4.2.3 and inserting the specific choice of the migration kernel defined in (4.6), we can rewrite (4.20) for $\xi = 0$ as follows (0-blocks are single components):

$$\begin{aligned} \mathrm{d}x_0^{\Omega_N}(t) &= \sum_{l\in\mathbb{N}} \frac{c_{l-1}}{N^{l-1}} \left[x_l^{\Omega_N}(N^{-l}t) - x_0^{\Omega_N}(t)\right] \mathrm{d}t + \sqrt{g\big(x_0^{\Omega_N}(t)\big)}\, \mathrm{d}w(t) \\ &\qquad + \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{N^m} \left[y_{m,0}^{\Omega_N}(t) - x_0^{\Omega_N}(t)\right] \mathrm{d}t, \\ \mathrm{d}y_{m,0}^{\Omega_N}(t) &= \frac{e_m}{N^m} \left[x_0^{\Omega_N}(t) - y_{m,0}^{\Omega_N}(t)\right] \mathrm{d}t, \qquad m \in \mathbb{N}_0. \end{aligned} \tag{4.23}$$

From (4.23) we see that migration between colonies can be expressed as a drift towards block averages at a higher hierarchical level.

The SSDE for the k-block average on time scale N^k reads as follows (recall (4.22)):

$$\begin{aligned} \mathrm{d}x_k^{\Omega_N}(t) &= \sum_{l\in\mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[x_{k+l}^{\Omega_N}(N^{-l}t) - x_k^{\Omega_N}(t)\right] \mathrm{d}t + \sqrt{\frac{1}{N^k} \sum_{i\in B_k(0)} g\big(x_i(N^k t)\big)}\; \mathrm{d}w_k(t) \\ &\qquad + \sum_{m\in\mathbb{N}_0} N^k \frac{K_m e_m}{N^m} \left[y_{m,k}^{\Omega_N}(t) - x_k^{\Omega_N}(t)\right] \mathrm{d}t, \\ \mathrm{d}y_{m,k}^{\Omega_N}(t) &= N^k \frac{e_m}{N^m} \left[x_k^{\Omega_N}(t) - y_{m,k}^{\Omega_N}(t)\right] \mathrm{d}t, \qquad m \in \mathbb{N}_0. \end{aligned} \tag{4.24}$$

To deduce these equations from (4.20), we sum over $\xi \in B_k$, speed up time by a factor N^k, insert the specific choice of the migration kernel in (4.6), and use the standard scaling properties of Brownian motion: $w(ct) =^d \sqrt{c}\, w(t)$ and $\sqrt{a}w(t) + \sqrt{b}w'(t) =^d \sqrt{a+b}\, w''(t)$, with $w(t)$ and $w'(t)$ independent Brownian motions, and with $=^d$ denoting equality in distribution. This computation is spelled out in Section 9.

Remark 4.2.5 (Separation space-time scales). The *block averages* and their evolution equations in (4.24) will be key objects in the analysis of the hierarchical mean-field limit $N \to \infty$. We will see that the limit $N \to \infty$ brings about considerable simplifications. In Section 4.4 we discuss these simplifications in detail. In particular, a *complete separation of space-time scales* takes places, in which each block average lives on its own time scale, effectively interacts with only one seed-bank, and effectively feels a drift towards the block average one hierarchical level up. ■

§4.2.3 Well-posedness

The generator of the system in (4.20) is given by

$$\begin{aligned} G \;=\; & \sum_{\xi\in\Omega_N}\Bigg(\sum_{\eta\in\Omega_N} a^{\Omega_N}(\xi,\eta)[x_\eta(t)-x_\xi(t)]\frac{\partial}{\partial x_\xi}+\frac{1}{2}g(x_\xi(t))\frac{\partial^2}{\partial x_\xi^2} \\ & +\sum_{m\in\mathbb{N}_0}\left[\frac{K_m e_m}{N^m}[y_{\xi,m}(t)-x_\xi(t)]\frac{\partial}{\partial x_\xi}+\frac{e_m}{N^m}[x_\xi(t)-y_{\xi,m}(t)]\frac{\partial}{\partial y_{\xi,m}}\right]\Bigg). \end{aligned} \tag{4.25}$$

Let

$$\begin{aligned} \mathcal{F}=\Big\{f\in C_b([0,\infty),S)\colon\ & f \text{ depends on finitely many components} \\ & \text{and is twice continuously differentiable in each component}\Big\}. \end{aligned} \tag{4.26}$$

Proposition 4.2.6 (Well-posedness). (a) *The SSDE in* (4.20) *has a unique strong solution in* $C([0,\infty),S)$*, whose law is the unique solution of the* $(G,\mathcal{F},\delta_u)$*-martingale problem for all* $u\in S$*.*

(b) *The process starting from* $u\in S$ *is Feller and strong Markov. Consequently, the SSDE in* (4.20) *defines a unique Borel Markov process starting from any initial law on* S*.*

Proof. Comparing with what is called model 2 in [43], we see that the Abelian group is chosen as in (4.2), the transition kernel is chosen as in (4.6), and the rates in and out of the seed-bank are $\frac{e_m}{N^m}$ and $\frac{K_m e_m}{N^m}$ for colour m. Hence the claim follows from [67], in the same way as shown in the proof of [43, Theorem 2.1]. □

Henceforth we write $\mathbb{P}$ and $\mathbb{E}$ to denote probability and expectation with respect to the random process in (4.19).

§4.2.4 Duality

If $g=dg_{\mathrm{FW}}$, then our model has a tractable dual, which turns out to play a crucial role in the analysis of the long-time behaviour. In this section we introduce the dual process following the same line of argument as in [43, Section 2.4]. There it was shown that the spatial Fisher-Wright diffusion with seed-bank is dual to a so-called *block-counting process* of a seed-bank coalescent. The latter describes the ancestral lines of $n\in\mathbb{N}$ individuals sampled from the current population backwards in time in terms of partition elements. At time zero the ancestral line of each individual is represented by a partition element. Traveling backwards in time, two partition elements merge as soon as their ancestral lines coalesce, i.e., two individuals have the same ancestor from that time onwards. Hence the seed-bank coalescent divides the ancestral lines of the $n\in\mathbb{N}$ individuals into subgroups of individuals with the same ancestor (i.e., individuals that are identical by descent). Therefore the seed-bank coalescent generates the ancestral lineages of the individuals evolving according to a Fisher-Wright diffusion with seed-bank, i.e., generates their full genealogy. The

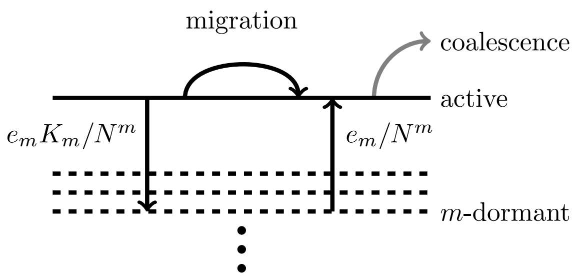

Figure 4.4: Transition scheme for an ancestral lineage in the dual, which moves according to the transition kernel $b(\cdot,\cdot)$ defined in (4.31)*. Two active ancestral lineages that are at the same colony coalesce at rate d.*

corresponding block-counting process counts the number of partition elements that are left when we travel backwards in time.

Formally, the spatial seed-bank coalescent is described as follows. Let

$$\mathbb{S} = \Omega_N \times \{A, (D_m)_{m\in\mathbb{N}_0}\} \tag{4.27}$$

be the *effective geographic space*. For $n \in \mathbb{N}$ the state space of the *n-spatial seed-bank coalescent* is the set of partitions of $\{1,\dots,n\}$, where the partition elements are marked with a position vector giving their locations. A state is written as π, where

$$\begin{aligned}
&\pi = ((\pi_1,\eta_1),\dots,(\pi_{\bar n},\eta_{\bar n})), \quad \bar n = |\pi|,\\
&\pi_\ell \subset \{1,\dots,n\}, \quad \{\pi_1,\cdots\pi_{\bar n}\} \text{ is a partition of } \{1,\dots,n\},\\
&\eta_\ell \in \mathbb{S}, \quad \ell \in \{1,\dots,\bar n\}, \quad 1 \le \bar n \le n.
\end{aligned} \tag{4.28}$$

A marked partition element (π_ℓ,η_ℓ) is called active if $\eta_\ell = (\xi, A)$ and m-dormant if $\eta_\ell = (\xi, D_m)$ for some $\xi \in \Omega_N$. The n-spatial seed-bank coalescent is denoted by

$$(\mathcal{C}^{(n)}(t))_{t\geqslant 0}, \tag{4.29}$$

and starts from

$$\mathcal{C}^{(n)}(0) = \pi(0), \qquad \pi(0) = \{(\{1\},\eta_{\ell_1}),\dots,(\{n\},\eta_{\ell_n})\}, \qquad \eta_{\ell_1},\dots,\eta_{\ell_n} \in \mathbb{S}. \tag{4.30}$$

The n-spatial seed-bank coalescent is the Markov process that evolves according to the following two rules (see Figs. 4.4–4.5).

(a) Each partition element moves independently of all other partition elements according to the transition kernel

$$b^{\Omega_N}((\xi,R_\xi),(\eta,R_\eta)) = \begin{cases} a^{\Omega_N}(\xi,\eta), & \text{if } R_\xi = R_\eta = A,\\ K_m \frac{e_m}{N^m}, & \text{if } \xi=\eta,\ R_\xi = A,\ R_\eta = D_m,\ \text{for } m\in\mathbb{N}_0,\\ \frac{e_m}{N^m}, & \text{if } \xi=\eta,\ R_\xi = D_m,\ R_\eta = A,\ \text{for } m\in\mathbb{N}_0,\\ 0, & \text{otherwise,}\end{cases} \tag{4.31}$$

where $a^{\Omega_N}(\cdot,\cdot)$ is the migration kernel defined in (4.6), K_m, $m \in \mathbb{N}_0$ are the relative sizes of the m-dormant population and the active population defined in

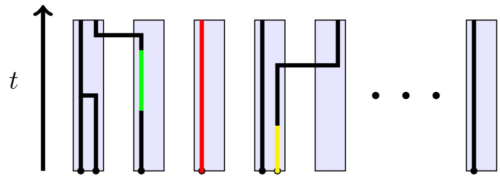

Figure 4.5: Picture of the evolution of lineages in the spatial coalescent. The purple blocks depict the colonies, the black lines the active lineages, and the coloured lines the dormant lineages. Blue lineages can migrate. Two black lineages can coalesce when they are at the same colony. Red dormant lineages first have to become black and active before they can migrate or coalesce with other black and active lineages. Note that the dual runs backwards *in time.*

(4.11), and e_m, $m \in \mathbb{N}_0$ are the coefficients controlling the exchange between the active and the dormant population defined in (4.10). Thus, an active partition element migrates according to the transition kernel $a^{\Omega_N}(\cdot,\cdot)$ and becomes m-dormant at rate $K_m \frac{e_m}{N^m}$, while an m-dormant partition element can only become active and does so at rate $\frac{e_m}{N^m}$.

(b) Independently of all other partition elements, two partition elements that are at the same colony and are both active coalesce with rate d, i.e., the two partition elements merge into one partition element.

Fig. 4.4 gives a schematic overview of the possible transitions of a single lineage, while Fig. 4.5 gives an example of the evolution in the dual. The *spatial seed-bank coalescent* $(\mathcal{C}(t))_{t\geq 0}$ is defined as the projective limit of the n-spatial seed-bank coalescents $(\mathcal{C}^{(n)}(t))_{t\geq 0}$ as $n \to \infty$. This object is well-defined by Kolmogorov's extension theorem (see [12, Section 3]).

For $n \in \mathbb{N}$ we define the block-counting process $(L(t))_{t\geq 0}$ corresponding to the n-spatial seed-bank coalescent as the process that counts at each site $(\xi, R_\xi) \in \Omega_N \times \{A, (D_m)_{m\in\mathbb{N}_0}\}$ the number of partition elements of $\mathcal{C}^{(n)}(t)$, i.e.,

$$
\begin{aligned}
&L(t) = \big(L_{(\xi,A)}(t), \big(L_{(\xi,D_m)}(t)\big)_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N},\\
&L_{(\xi,A)}(t) = L_{(\xi,A)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=(\xi,A)\}},\\
&L_{(\xi,D_m)}(t) = L_{(\xi,D_m)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=(\xi,D_m)\}}, \qquad m \in \mathbb{N}_0.
\end{aligned}
\tag{4.32}
$$

The state space of $(L(t))_{t\geq 0}$ is $S' = (\mathbb{N}_0 \times \mathbb{N}_0^{\mathbb{N}_0})^{\Omega_N}$. We denote the elements of S' by sequences $(m_\xi, (n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}$, and define $\delta_{(\eta,R_\eta)} \in S'$ to be the element of S' that is 0 at all sites $(\xi, R_\xi) \in \Omega_N \times \{A, (D_m)_{m\in\mathbb{N}_0}\}\backslash(\eta, R_\eta)$, and 1 at the site (η, R_η). From the evolution of $\mathcal{C}^{(n)}(t)$ described below (4.29) we see that the block-counting process has the following transition kernel:

$$(m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}\to$$
$$\begin{cases}(m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}-\delta_{(\eta,A)}+\delta_{(\zeta,A)}, & \text{at rate } m_\eta a(\eta,\zeta) \text{ for } \eta,\zeta\in\Omega_N,\\ (m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}-\delta_{(\eta,A)}, & \text{at rate } d\binom{m_\eta}{2} \text{ for } \eta\in\Omega_N,\\ (m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}-\delta_{(\eta,A)}+\delta_{(\eta,D_m)}, & \text{at rate } m_\eta K_m\frac{e_m}{N^m} \text{ for } \eta\in\Omega_N,\\ (m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}+\delta_{(\eta,A)}-\delta_{(\eta,D_m)}, & \text{at rate } n_{\eta,m}\frac{e_m}{N^m} \text{ for } \eta\in\Omega_N.\end{cases}\tag{4.33}$$

The process $(Z(t))_{t\geq0}$ defined in (4.20) is dual to the block-counting process $(L(t))_{t\geq0}$ with *duality function* $H\colon\ S\times S'\to\mathbb{R}$ defined by

$$H\Big(\big(x_\xi,(y_{\xi,m})_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N},\big(m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N}\Big)=\prod_{\xi\in\Omega_N}x_\xi^{m_\xi}\prod_{m\in\mathbb{N}_0}y_{\xi,m}^{n_{\xi,D_m}}.\tag{4.34}$$

Proposition 4.2.7 (Duality relation). *Let H be as in* (4.34). *Then, for all $(x_\xi,(y_{\xi,m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}\in S$ and $(m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0})_{\xi\in\Omega_N}\in S'$,*

$$\begin{aligned}&\mathbb{E}_{\left(x_\xi,(y_{\xi,m})_{m\in\mathbb{N}_0}\right)_{\xi\in\Omega_N}}\left[H\Big(\big(x_\xi(t),(y_{\xi,m}(t))_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N},(m_\xi,n_\xi)_{\xi\in\Omega_N}\Big)\right]\\&=\mathbb{E}_{\left(m_\xi,(n_{\xi,D_m})_{m\in\mathbb{N}_0}\right)_{\xi\in\Omega_N}}\left[H\Big(\big(x_\xi,(y_{\xi,m})_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N},\big(L_{(\xi,A)}(t),(L_{(\xi,D_m)}(t))_{m\in\mathbb{N}_0}\big)_{\xi\in\Omega_N}\Big)\right]\end{aligned}\tag{4.35}$$

with $\mathbb{E}$ the generic symbol for expectation (on the left over the original process, on the right over the dual process).

Proposition 4.2.7 was proved in [43, Section 2.4]. Since the duality function H captures all the mixed moments of $(Z(t))_{t\geq0}$, the duality relation is that of a *moment dual.*

Remark 4.2.8 (Duality relation in terms of the effective geographic space). Interpreting $(Z(t))_{t\geq0}$ as a process on the effective geographic space $\mathbb{S}=\Omega_N\times\{A,(D_m)_{m\in\mathbb{N}_0}\}$, we can rewrite (4.20) as

$$\begin{aligned}\mathrm{d}z_{(\xi,R_\xi)}(t)=&\sum_{(\xi,R_\xi)\in\mathbb{S}}b^{\Omega_N}((\xi,R_\xi),(\eta,R_\eta))[z_{(\eta,R_\eta)}(t)-z_{(\xi,R_\xi)}(t)]\,\mathrm{d}t\\&+1_{\{R_\xi=A\}}\sqrt{g(z_{(\xi,R_\xi)}(t))}\,\mathrm{d}w_\xi(t),\qquad(\xi,R_\xi)\in\mathbb{S},\end{aligned}\tag{4.36}$$

where $b^{\Omega_N}(\cdot,\cdot)$ is the transition kernel defined in (4.31). If $g=dg_{\mathrm{FW}}$, then we can write its dual process as follows. Let $(L(t))_{t\geq0}=(L(\mathcal{C}(t))_{t\geq0}$ be the block-counting process that at each site $(\xi,\mathbb{R}_\xi)\in\mathbb{S}$ counts the number of partition elements of $\mathcal{C}(t)$, i.e.,

$$\begin{aligned}L(t)&=(L_{(\xi,R_\xi)}(t))_{(\xi,R_\xi)\in\mathbb{S}},\\L_{(\xi,R_\xi)}(t)&=L_{(\xi,R_\xi)}(\mathcal{C}(t))=\sum_{\ell=1}^{\bar{n}}1_{\{\eta_\ell(t)=(\xi,R_\xi)\}}.\end{aligned}\tag{4.37}$$

Rewrite the duality function H in (4.34) as

$$H\Big((z_{(\xi,R_\xi)},l_{(\xi,R_\xi)})_{(\xi,R_\xi)\in\mathbb{S}}\Big)=\prod_{(\xi,R_\xi)\in\mathbb{S}}z_{(\xi,R_\xi)}^{l_{(\xi,R_\xi)}}.\tag{4.38}$$

Then, for $z \in S$ and $l \in S'$, the duality relation in (4.35) reads

$$\mathbb{E}_{z(\xi,R_\xi)}\big[H(z_{(\xi,R_\xi)}(t), l_{(\xi,R_\xi)})\big] = \mathbb{E}_{l(\xi,R_\xi)}\big[H(z_{(\xi,R_\xi)}, L_{(\xi,R_\xi)}(t))\big]. \tag{4.39}$$

Interpreting the duality relation in terms of the effective geographic space $\mathbb{S}$, we see that each ancestral lineage in the dual is a Markov process moving according to the transition kernel $b(\cdot,\cdot)$ defined in (4.31). Interpreting the duality relation in terms of the geographic space Ω_N, we see that an ancestral lineage is a random walk on Ω_N, with internal states A and $(D_m)_{m\in\mathbb{N}_0}$. Both interpretations turn out to be useful when we analyse the long-time behaviour of the system. ■

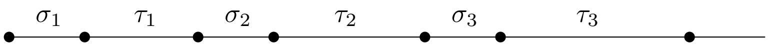

Figure 4.6: Renewal process induced by a lineage in the dual moving according to the transition kernel $b(\cdot,\cdot)$. For $k \in \mathbb{N}$, σ_k denotes the kth active period and τ_k the kth dormant period.

Remark 4.2.9 (The renewal process induced by the dual process). The partition elements describing the dual process give rise to a renewal process on the active state A and the dormant state $D = \bigcup_{m\in\mathbb{N}_0} D_m$. Since the only transition a dormant lineage can make is to become active, irrespectively of its colour, each dual lineage induces a sequence of active and dormant time lapses. Let $(\sigma_k)_{k\in\mathbb{N}}$ denote the successive active time periods and $(\tau_k)_{k\in\mathbb{N}}$ the successive dormant time periods (see Fig. 4.6). Then $(\sigma_k)_{k\in\mathbb{N}}$ and $(\tau_k)_{k\in\mathbb{N}}$ are sequences of i.i.d. random variables with marginal laws (recall (4.13))

$$\mathbb{P}(\sigma_1 > t) = \mathrm{e}^{-\chi t}, \quad \mathbb{P}(\tau_1 > t) = \sum_{m\in\mathbb{N}_0} \frac{K_m \frac{e_m}{N^m}}{\chi}\, \mathrm{e}^{-\frac{e_m}{N^m} t}, \qquad t \geq 0. \tag{4.40}$$

Remark 4.2.10 (Wake up times). The renewal process in Fig. 4.6 is key to understanding the long-time behaviour of the model (as we will see in Section 4.3). Note that

$$\tau = \tau_1 \tag{4.41}$$

represents *the typical wake-up time of a lineage in the dual.* By choosing specific sequences $(K_m)_{m\in\mathbb{N}_0}$ and $(e_m)_{m\in\mathbb{N}_0}$ we can mimic different wake-up time distributions. In particular, if we allow $\rho = \sum_{m\in\mathbb{N}_0} K_m = \infty$ (recall (4.14)), then τ may have a fat-tail (examples are given in Section 4.3). In other words, *the internal structure of the seed-bank allows us to model fat-tailed wake-up times without loosing the Markov property of the evolution.* ■

Note that even when there is no dual, i.e., $g \in \mathcal{G}$ with $g \neq dg_{\mathrm{FW}}$, we can still define τ by (4.41), since τ_1 in (4.40) is a random variable that depends only on the sequences $(K_m)_{m\in\mathbb{N}_0}$ and $(e_m)_{m\in\mathbb{N}_0}$, and we can still interpret τ as the typical wake-up time of an individual in the population. ■

§4.2.5 Clustering criterion

• Clustering criterion for Fisher-Wright diffusion function

In [43] we showed that the system exhibits a dichotomy between *coexistence* (= locally multi-type equilibria) and *clustering* (= locally mono-type equilibria). The clustering criterion is based on the dual and requires the notion of colour regularity. We call a law translation invariant when it is *invariant under the group action.*

Definition 4.2.11 (Colour regular initial measures). We say that a translation invariant initial measure $\mu(0)$ is *colour regular* when

$$\lim_{m\to\infty} \mathbb{E}_{\mu(0)}[y_{0,m}] \quad \text{exists.} \tag{4.42}$$

This condition is needed because, as time progresses, lineages starting from slower and slower seed-banks become active and bring new types into the active population. Without control on the initial states of the slow seed-banks, there may be no convergence to equilibrium. ■

The key clustering criterion is the following.

Proposition 4.2.12 (Clustering criterion). *Suppose that $\mu(0)$ is translation invariant. If $\rho = \infty$ (recall (4.14)), then additionally suppose that $\mu(0)$ is colour regular. Let $d \in (0,\infty)$. Then the system with $g = dg_{\mathrm{FW}}$ clusters if and only if in the dual two partition elements coalesce with probability 1.*

The idea behind Theorem 4.2.12 is as follows. If in the dual two partition elements coalesce with probability 1, then a random sample of n individuals drawn from the current population has a common ancestor some finite time backwards in time. Since individuals inherit their type from their parent individuals, this means that all n individuals have the same type. A formal proof was given in [43, Section 4.3]. The proof is valid for any geographic space given by a countable Abelian group endowed with the discrete topology, of which Ω_N is an example.

• Clustering criterion for general diffusion function

For $g \in \mathcal{G}$ with $g \neq dg_{\mathrm{FW}}$ no dual is available and hence we cannot use the clustering criterion in Proposition 4.2.12. However, as shown in [43], we can argue by duality comparison arguments (see [43, Lemma 5.5 and Lemma 6.3]) that the system evolving according to (4.20) with $g \in \mathcal{G}$ clusters if and only if the system with $g = dg_{\mathrm{FW}}$ for some $d \in (0,\infty)$ clusters. In particular, for $g = dg_{\mathrm{FW}}$, $d \in (0,\infty)$, whether or not the system clusters does not depend on the resampling rate d.

§4.3 Main results: $N < \infty$, identification of clustering regime

In this section we identify the *clustering regime*, i.e., the range of *parameters* for which the clustering criterion in Proposition 4.2.12 is met. In [43, Section 3.2, Theorem

3.3] we derived a necessary and sufficient condition for when clustering prevails, for any geometric space given by a countable Abelian group endowed with the discrete topology. Recall χ in (4.13), ρ in (4.14) and τ in (4.41),. From (4.40) it follows that

$$\mathbb{E}[\tau] = \sum_{m\in\mathbb{N}_0} \frac{K_m}{\chi} = \frac{\rho}{\chi}, \tag{4.43}$$

and hence the mean wake-up time is finite if $\rho < \infty$ and infinite if $\rho = \infty$. In Section 4.3.1 we look at $\rho < \infty$ and in Section 4.3.2 at $\rho = \infty$. In Section 4.3.3 we summarise our findings and identify the clustering regime.

§4.3.1 Finite mean wake-up time

Suppose that the system evolving according to (4.20) has a translation invariant initial measure $\mu(0)$ with density $\theta \in (0,1)$. Then [43, Theorem 3.3] says that for $\rho < \infty$ clustering occurs if and only if

$$\int_1^\infty \mathrm{d}t\, a_t^{\Omega_N}(0,0) = \infty. \tag{4.44}$$

It is known that (4.44) holds for the hierarchical migration defined in (4.6) if and only if [19, Section 3]

$$\sum_{k\in\mathbb{N}_0} \frac{1}{c_k} = \infty. \tag{4.45}$$

Hence, for $\rho < \infty$, the clustering criterion depends on the migration kernel *only* and the seed-bank has no effect.

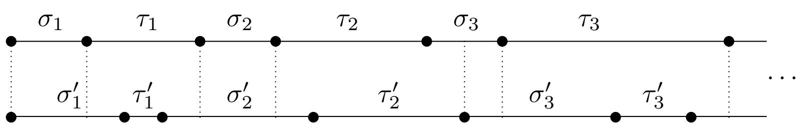

Figure 4.7: Successive periods during which the two random walks are active and dormant (recall (4.40) *and Fig. 4.6). The time lapses between successive pairs of dotted lines represent periods of* joint activity.

In view of Proposition 4.2.12, if $g = dg_{\mathrm{FW}}$, then clustering prevails if and only if two lineages in the dual coalesce with probability 1. Recall that two lineages in the dual can only coalesce when they are at the same site and are both active. Since the rate of coalescence is $d \in (0,\infty)$, each time this happens the two lineages have a positive probability to coalesce before moving or becoming dormant. Therefore, clustering prevails if and only if two lineages meet infinitely often while being active. This happens exactly when (4.44) holds. The fact that the seed-bank plays no role can be seen from the dual. Each lineage in the dual moves according to the transition kernel $b(\cdot,\cdot)$ (recall (4.31)). Looking at the renewal process induced by the dual process (recall Remark 4.2.9 and Fig. 4.7), we see that for $\rho < \infty$ the probability that a lineage in the dual is active at time t is approximately $\frac{1}{1+\rho}$ for large t. The total activity time of a lineage up to time t is therefore approximately $\frac{1}{1+\rho}t$ for large t. Hence the total time the two lineages in the dual are at the same site and are both active is approximately

$$\int_1^\infty \mathrm{d}t \left(\frac{1}{1+\rho}\right)^2 a_{2\frac{1}{1+\rho}t}(0,0), \tag{4.46}$$

By Polya's argument, if the integral in (4.46) is infinite, then two lineages in the dual meet infinitely often while being active. After a variable transformation, (4.46) becomes the integral in (4.44) up to a constant. (For a formal proof of the criterion in (4.44), we refer to [43].) By the above argument, we can think of the integral in (4.44) as the *total hazard of coalescence* of two dual lineages. To get the result for general $g \in \mathcal{G}$ we must invoke the duality comparison arguments mentioned in Section 4.2.5.

In terms of the degree of the random walk (recall Remark 4.2.1), (4.44) corresponds to hierarchical migration with degree 0^-. The same criterion as in (4.44) was found in [36] for interacting Fisher-Wright diffusions on the hierarchical lattice without seed-bank ($\rho = 0$). Hence we conclude that for $\rho < \infty$ the seed-bank does not affect the dichotomy.

§4.3.2 Infinite mean wake-up time

If $\rho = \infty$, then the seed-bank does affect the dichotomy. To apply the criterion in [43, Theorem 3.3], we assume that the system evolving according to (4.20) has a translation-invariant initial measure $\mu(0)$ with density $\theta \in (0,1)$ that is colour regular.

The criterion for clustering that was derived in [43] for $\rho = \infty$ applies to wake-up times τ (recall (4.41)) of the form

$$\frac{\mathbb{P}(\tau \in \mathrm{d}t)}{\mathrm{d}t} \sim \varphi(t)\, t^{-(1+\gamma)}, \quad t \to \infty, \quad \gamma \in (0,1], \tag{4.47}$$

with φ *slowly varying at infinity.* Define

$$\hat{\varphi}(t) = \begin{cases} \varphi(t), & \gamma \in (0,1), \\ \mathbb{E}\left[\tau \wedge t\right] & \gamma = 1. \end{cases} \tag{4.48}$$

As shown in [8, Section 1.3], every slowly varying function φ may be assumed to be infinitely differentiable and to be represented by the integral

$$\varphi(t) = \exp\left[\int_{(\cdot)}^t \frac{\mathrm{d}u}{u}\, \psi(u)\right] \tag{4.49}$$

for some $\psi\colon [0,\infty) \to \mathbb{R}$ such that $\lim_{u\to\infty} |\psi(u)| = 0$. From (4.47) we see that $\hat{\varphi}(t)$ is also slowly varying. If we assume that $|\psi(u)| \le C/\log u$ for some $C < \infty$, then the system clusters if and only if (see [43, Section 3.2])

$$\int_{(\cdot)}^\infty \mathrm{d}t\, \hat{\varphi}(t)^{-1/\gamma}\, t^{-(1-\gamma)/\gamma}\, \hat{a}_t(0,0) = \infty. \tag{4.50}$$

Note that γ in (4.47) is the *tail exponent* of the typical wake-up time τ (recall Remark 4.2.9) and depends on the sequences $\underline{e}, \underline{K}$ in (4.10) governing the exchange with the seed-bank. If $g = dg_{\mathrm{FW}}$, then in view of Theorem 4.2.12 the criterion in (4.50) determines whether two lineages in the dual coalesce with probability 1.

In Section 5 we will use the renewal process induced by the dual (recall Remark 4.2.9) to show that (4.50) indeed gives the *total hazard of coalescence* of two dual lineages. Therefore the integral in (4.50) is the counterpart of (4.44). The rate of coalescence again does not affect the dichotomy: in [43] a duality comparison argument was used to show that (4.50) gives the clustering criterion also for $g \in \mathcal{G}$ with $g \neq dg_{\mathrm{FW}}$. The effect of the seed-bank on the dichotomy is embodied by the term $\hat{\varphi}(t)^{-1/\gamma}\, t^{-(1-\gamma)/\gamma}$ in (4.50). The criterion in (4.50) shows that there is a competition between migration and exchange with the seed-bank.

For the special case where $\hat{\varphi}(t) \asymp 1$, the criterion in (4.50) says that (recall Remark 4.2.1)

$$\text{clustering} \quad \Longleftrightarrow \quad \text{either } \delta^- \leq -\frac{1-\gamma}{\gamma} \text{ or } \delta^+ < -\frac{1-\gamma}{\gamma}. \tag{4.51}$$

Condition (4.50) implies that for $\gamma \in (0, \frac{1}{2})$ no clustering is possible: the typical wake-up time has such a heavy tail that with a positive probability two dual lineages do not meet, irrespective of the migration.

Definition 4.3.1. In what follows we will focus on the following two specific parameter regimes:

- *Asymptotically polynomial*, i.e.,

$$\begin{aligned} &K_k \sim Ak^{-\alpha}, \quad e_k \sim Bk^{-\beta}, \quad c_k \sim Fk^{-\phi}, \quad k \to \infty, \\ &A, B, F \in (0,\infty),\ \alpha, \beta, \phi \in \mathbb{R}. \end{aligned} \tag{4.52}$$

- *Pure exponential*, i.e.,

$$K_k = K^k, \quad e_k = e^k, \quad c_k = c^k, \quad k \in \mathbb{N}_0,\ K, e, c \in (0,\infty). \tag{4.53}$$

Note that both (4.7) and (4.12) are satisfied for $N \to \infty$. Also note that an infinite seed-bank corresponds to $\alpha \in (-\infty, 1]$, respectively, $K \in [1, \infty)$. □

The scaling of the wake-up time and the migration kernel in these parameter regimes are as follows.

Theorem 4.3.2 (Scaling of wake-up time and migration kernel). *Suppose that $\rho = \infty$. Then*

(a) Subject to (4.52),

$$\begin{aligned} &\gamma = 1, \quad \varphi(t) \asymp (\log t)^{-\alpha}, \quad \hat{\varphi}(t) \asymp \begin{cases} (\log t)^{1-\alpha}, & \alpha \in (-\infty, 1), \\ \log\log t, & \alpha = 1, \end{cases} \\ &a_t^{\Omega_N}(0,0) \asymp t^{-1} \log^{\phi} t. \end{aligned} \tag{4.54}$$

(b) Subject to (4.53),

$$\begin{aligned} &\gamma = \gamma_{N,K,e} = \frac{\log(N/Ke)}{\log(N/e)}, \quad \varphi(t) \asymp 1, \quad \hat{\varphi}(t) \asymp \begin{cases} 1, & K \in (1,\infty), \\ \log t, & K = 1, \end{cases} \\ &a_t^{\Omega_N}(0,0) \asymp t^{-1-\delta_{N,c}}, \end{aligned} \tag{4.55}$$

where

$$\delta_{N,c} = \frac{\log c}{\log(N/c)}. \tag{4.56}$$

Theorem 4.3.2 will be proved in Section 5.

Note that, by (4.55), $\gamma_{N,K,e} = 1$ for all N when $K = 1$, while $\gamma_{N,K,e} < 1$ for all N when $K > 1$, but with $\gamma_{N,K,e} \uparrow 1$ as $N \to \infty$. Also note that, subject to (4.52), (4.54) says that the degree of the random walk is 0^- for $\phi \geq -1$ and 0^+ for $\phi < -1$, while subject to (4.53), by (4.55), the degree of the random walk is $\delta^-_{N,c}$, which is 0 for all N when $c = 1$, and tends to 0 as $N \to \infty$ from above when $c > 1$ and from below as $c < 1$. Thus, both (4.52) and (4.53) with $N \to \infty$ correspond to a *critically recurrent migration* and a *critically infinite seed-bank*.

§4.3.3 Clustering regime

Summarising the above discussion, we can now identify the clustering regime for both finite and infinite seed-banks.

Theorem 4.3.3 (Clustering regime). (1) *If $\rho < \infty$, then clustering prevails if and only if*

$$\sum_{k\in\mathbb{N}_0} \frac{1}{c_k} = \infty. \tag{4.57}$$

(2) *If $\rho = \infty$, then clustering prevails for N large enough*

(a) Subject to (4.52) *if and only if*

$$-\phi \leq \alpha \leq 1. \tag{4.58}$$

(b) Subject to (4.53) *if and only if*

$$Kc \leq 1 \leq K. \tag{4.59}$$

Also Theorem 4.3.3 will be proved in Section 5.

Note that for $\rho < \infty$ the clustering regime follows by combining (4.44) and (4.45), while for $\rho = \infty$ the clustering regime follows by substituting into (4.50) the scaling of the wake-up times and the migration kernel stated in Theorem 4.3.2.

Remark 4.3.4. Note that subject to (4.52), respectively, (4.53), $\rho < \infty$ implies that $\alpha > 1$, respectively, $K < 1$, and so the clustering regime is $-\phi \leq 1$, respectively, $c \leq 1$ (recall (4.57)), which are less stringent than (4.58), respectively, (4.59). ■

§4.4 Main results: $N \to \infty$, renormalisation and multi-scale limit

This section contains our multi-scale hierarchical limit theorems. The multi-scale hierarchical limit theorems analyse the evolution of the block averages defined in Definition 4.2.3. In Section 4.4.1 we recall a path topology referred to as the *Meyer-Zheng* *topology*, which we will need in part of our multi-scale hierarchical limit theorems. In Section 4.4.2 we present the conceptual ingredients needed for our theorems. In Section 4.4.3 we state two versions of the hierarchical multi-scale limit (Theorems 4.4.2 and 4.4.4), and comment on how they are related to each other. In Section 4.4.4 we explain how they arise from a heuristic analysis of the SSDE in (4.20).

§4.4.1 Intermezzo: Meyer-Zheng topology

Recall the block averages defined in (4.2.3) and their evolution equations in (4.24). In the limit as $N \to \infty$, some of the pre-factors in (4.24) diverge as a result of the speeding up of time. This makes the processes increasingly more volatile: paths becomes rougher and rougher during rarer and rarer times. Therefore we cannot work with weak convergence on path space $C([0,\infty),E)$ w.r.t. the topology generated by the *sup-norm* on compacts, or on path space $D([0,\infty),E)$ w.r.t. the Skorohod metric on compacts. Rather we must follow the methodology used in [23, pp. 792–794] and employ the so-called *Meyer-Zheng topology* on *pseudopaths*, (see [59]), which is based on the following idea. Consider functions $f\colon\ [0,\infty) \to E$, with (E,d) a Polish space, and sequences of functions $(f_n)_{n\in\mathbb{N}}$ that are càdlàg paths, i.e., functions in the Skorohod space $D([0,\infty),E)$. Then $(f_n)_{n\in\mathbb{N}}$ converges to f in the Meyer-Zheng topology if and only if

$$\lim_{n\to\infty} \int_a^b \mathrm{d}t \left[1 \wedge d(f(t),f_n(t))\right] = 0 \qquad \forall\ 0 \le a < b < \infty. \tag{4.60}$$

However, the topology induced by the metric in (4.60) does not turn $D([0,\infty),E)$ into a *closed* space (while in order to apply the classical theory of weak convergence of probability laws on path space we need the path space to be Polish).

To turn the idea from (4.60) into a manageable topology, we proceed by defining a *space of pseudopaths* equipped with the *Meyer-Zheng topology.* If (E,d) is a Polish space and $s \mapsto v(s)$ is a measurable map from $[0,\infty)$ to E, then the pseudopath ψ_v is the *probability measure* ρ on $[0,\infty)\times E$, defined by

$$\rho((a,b)\times B) = \int_a^b \mathrm{d}s\, \mathrm{e}^{-s}\, 1_B(v(s)), \qquad B \in \mathcal{B}(E), \tag{4.61}$$

Hence ψ_v is the image measure of $\mathrm{e}^{-t}\mathrm{d}t$ under the mapping $t \to (t,v(t))$. In other words, we consider the *weighted occupation measure* of the path in E in order to describe paths that are *regular representatives* in the space of functions once we take into account (4.60). Note that a piece-wise constant càdlàg path is uniquely determined by its occupation measure. So is a continuous path with continuous local times. The space of all pseudopaths is denoted by Ψ.

Since pseudopaths are measures on $[0,\infty]\times E$, convergence of pseudopaths is defined as *weak convergence* of *probability measures on* $[0,\infty]\times E$. A sequence $(v_n)_{n\in\mathbb{N}}$ of measurable maps from $[0,\infty)\times E$ is said to converge in the Meyer-Zheng topology to a measurable map v if $\lim_{n\to\infty}\psi_{v_n} = \psi_v$, i.e., $\lim_{n\to\infty}\psi_{v_n} f = \psi_v f$ for all continuous bounded functions f on $[0,\infty]\times E$.

Remark 4.4.1 (Pseudopaths). The space Ψ of pseudopaths endowed with the Meyer-Zheng topology is Polish, but the space $D([0,\infty), E)$ endowed with the Meyer-Zheng topology is not Polish (see [59, p. 372]). ■

In what follows, each time convergence holds in the Meyer-Zheng topology we will say so explicitly. If no topology is mentioned, then we mean convergence in $\mathcal{C}_b([0,\infty),[0,1])$. In Appendix B.2 we collect some basic facts about the Meyer-Zheng topology taken from [59] and [53].

§4.4.2 Main ingredients for the hierarchical multi-scale limit

Recall the definition of θ_x and θ_{y_m} in (4.21). Define

$$\vartheta_k = \frac{\theta_x + \sum_{m=0}^{k} K_m \theta_{y_m}}{1 + \sum_{m=0}^{k} K_m}, \qquad k \in \mathbb{N}_0. \tag{4.62}$$

For $\rho < \infty$, and for $\rho = \infty$ under the additional assumption of colour regularity (recall Proposition 4.2.12), we have

$$\lim_{k\to\infty} \vartheta_k = \theta \qquad \text{for some } \theta \in [0,1]. \tag{4.63}$$

Define the *slowing-down constants* ($E_0 = 1$)

$$E_k = \frac{1}{1 + \sum_{m=0}^{k-1} K_m}, \qquad k \in \mathbb{N}_0. \tag{4.64}$$

For $l \in \mathbb{N}_0$, let

$$\big(\theta, (y_{m,l})_{m\in\mathbb{N}_0}\big) \tag{4.65}$$

be a sequence of random variables taking values in $[0,1]$, and let

$$\Big(z_{l,(\theta,(y_{m,l})_{m\in\mathbb{N}_0})}(t)\Big)_{t\geq 0} = \Big(x_l(t), (y_{m,l}(t))_{m\in\mathbb{N}_0}\Big)_{t\geq 0} \tag{4.66}$$

be the *full process* evolving according to

$$\begin{aligned}
\mathrm{d}x_l(t) &= E_l\left[c_l[\theta - x_l(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(l)}g)(x_l(t))}\,\mathrm{d}w(t) + K_l e_l\,[y_{l,l}(t) - x_l(t)]\,\mathrm{d}t\right],\\
y_{m,l}(t) &= x_l(t), \qquad 0 \leq m < l,\\
\mathrm{d}y_{l,l}(t) &= e_l\,[x_l(t) - y_{l,l}(t)]\,\mathrm{d}t, \qquad m = l,\\
y_{m,l}(t) &= y_{m,l}, \qquad m > l.
\end{aligned} \tag{4.67}$$

where $\mathcal{F}^{(l)}g$ is an element of $\mathcal{G}$, (recall (4.15)), that will be defined in (4.76) below. By [72] the above SSDE has a unique solution for every initial measure. For $l \in \mathbb{N}_0$, let

$$(z^{\text{eff}}_{l,\theta}(t))_{t\geq 0} = \big(x^{\text{eff}}_l(t), y^{\text{eff}}_{l,l}(t)\big)_{t\geq 0} \tag{4.68}$$

be the *effective process* evolving according to

$$\begin{aligned}
\mathrm{d}x_l^{\mathrm{eff}}(t) &= E_l\left[c_l\,[\theta - x_l^{\mathrm{eff}}(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(l)}g)(x_l^{\mathrm{eff}}(t))}\,\mathrm{d}w(t) + K_l e_l\,[y_{l,l}^{\mathrm{eff}}(t) - x_l^{\mathrm{eff}}(t)]\,\mathrm{d}t\right],\\
\mathrm{d}y_{l,l}^{\mathrm{eff}}(t) &= e_l\,[x_l^{\mathrm{eff}}(t) - y_{l,l}^{\mathrm{eff}}(t)]\,\mathrm{d}t.
\end{aligned} \tag{4.69}$$

Comparing (4.67) with (4.69), we see that the effective process looks at the non-trivial components of the full process.

Apart from (4.66) and (4.68), we need the following list of *ingredients* to formally state the multi-scale limit:

(a) For $l \in \mathbb{N}_0$ and $t > 0$, define the *estimators* for the finite system by

$$\begin{aligned}
\bar{\Theta}^{(l),\Omega_N}(t) &= \frac{1}{N^l}\sum_{\xi\in B_l}\frac{x_\xi^{\Omega_N}(t) + \sum_{m=0}^{l-1}K_m y_{\xi,m}^{\Omega_N}(t)}{1+\sum_{m=0}^{l-1}K_m},\\
\Theta_x^{(l),\Omega_N}(t) &= \frac{1}{N^l}\sum_{\xi\in B_l} x_\xi^{\Omega_N}(t),\\
\Theta_{y_m}^{(l),\Omega_N}(t) &= \frac{1}{N^l}\sum_{\xi\in B_l} y_{\xi,m}^{\Omega_N}(t), \qquad m\in\mathbb{N}_0,
\end{aligned} \tag{4.70}$$

and put

$$\begin{aligned}
\mathbf{\Theta}^{(l),\Omega_N}(t) &= \big(\Theta_x^{(l),\Omega_N}(t), \big(\Theta_{y_m}^{(l),\Omega_N}(t)\big)_{m\in\mathbb{N}_0}\big),\\
\mathbf{\Theta}^{\mathrm{eff},(l),\Omega_N}(t) &= \big(\bar{\Theta}^{(l),\Omega_N}(t), \Theta_{y_l}^{(l),\Omega_N}(t)\big).
\end{aligned} \tag{4.71}$$

We call $(\mathbf{\Theta}^{(l),\Omega_N}(t))_{t\geq 0}$ the *full estimator process* and $(\mathbf{\Theta}^{\mathrm{eff},(l),\Omega_N}(t))_{t\geq 0}$ the *effective process*. Note that $\Theta_x^{(l),\Omega_N}(t)$ is the empirical average of the active components in the l-block, while $\Theta_{y_m}^{(l),\Omega_N}(t)$, is the empirical average of the m-dormant components in the l-block, both without scaling of time. Note that $\Theta_x^{(l),\Omega_N}(N^l t) = x_l^{\Omega_N}(t)$. The *level-$l$ estimator* $\bar{\Theta}^{(l),\Omega_N}(t)$ will play an important role in our analysis. Using (4.24), we can derive the evolution equations of $\bar{\Theta}^{(l),\Omega_N}(N^l t)$. We see that in the evolution of $\bar{\Theta}^{(l),\Omega_N}(N^l t)$ no rates appear that tend to infinity as $N\to\infty$. However, in the evolution of $\Theta_x^{(l),\Omega_N}(N^l t)$ and $\Theta_{y_m}^{(l),\Omega_N}(N^l t)$ for $m < l$ the rates describing the interaction between the active and the dormant population tend to infinity as $N\to\infty$.

(b) For $l\in\mathbb{N}_0$, consider *time scales* $N^l t_l$ such that

$$\mathcal{L}\big[\bar{\Theta}^{(l),\Omega_N}(N^l t_l - L(N)N^{l-1}) - \bar{\Theta}^{(l),\Omega_N}(N^l t_l)\big] = \delta_0 \tag{4.72}$$

for all $L(N)$ satisfying $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, but not for $L(N) = N$. In words, $N^l t_l$ is the time scale on which $(\bar{\Theta}^{(l),\Omega_N}(N^l t_l))_{t_l>0}$ is no longer a fixed process.

(c) For $l\in\mathbb{N}_0$ the *invariant measure* for the limiting evolution of the l-block averages in (4.67) is denoted by

$$\Gamma^{(l)}_{(\theta,y_l)}, \qquad y_l = (y_{m,l})_{m\in\mathbb{N}_0}. \tag{4.73}$$

(The existence of and convergence to this equilibrium will be proved in Section 9.2.) Note that $\Gamma^{(l)}_{(\theta,y_l)}$ depends on choice of the rates E_l, c_l, K_l, e_l in (4.67). The *invariant measure* of the limiting evolution for the effective l-block process in (4.69) is denoted by

$$\Gamma_\theta^{\mathrm{eff},(l)}. \tag{4.74}$$

Also $\Gamma_\theta^{\mathrm{eff},(l)}$ depends on the choice of the rates E_l, c_l, K_l, e_l.

(d) For $l \in \mathbb{N}_0$, let $\mathcal{F}^{E_l,c_l,K_l,e_l}$ denote the *renormalisation transformation* acting on $\mathcal{G}$ defined by

$$(\mathcal{F}^{E_l,c_l,K_l,e_l} g)(\theta) = \int_{[0,1]^2} g(x)\, \Gamma_\theta^{\mathrm{eff},(l)}(\mathrm{d}x), \qquad \theta \in [0,1]. \tag{4.75}$$

(In Section 6.3 we show that $\mathcal{F}g \in \mathcal{G}$.) For $k \in \mathbb{N}_0$, define the *iterate* of the renormalisation transformation as the composition

$$\mathcal{F}^{(k)} = \mathcal{F}^{E_{k-1},c_{k-1},K_{k-1},e_{k-1}} \circ \cdots \circ \mathcal{F}^{E_0,c_0,K_0,e_0}. \tag{4.76}$$

(e) For $k \in \mathbb{N}_0$, define the *interaction chain* [25]

$$(M^k_{-l})_{-l=-(k+1),-k,\dots,0} \tag{4.77}$$

as the *time-inhomogeneous* Markov chain on $[0,1] \times [0,1]^{\mathbb{N}_0}$ with initial state

$$M^k_{-(k+1)} = (\vartheta_k, \overbrace{\vartheta_k, \cdots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots) \tag{4.78}$$

that evolves from time $-(l+1)$ to time $-l$ according to the transition kernel $Q^{[l]}$ on $[0,1] \times [0,1]^{\mathbb{N}_0}$ given by

$$Q^{[l]}(u, \mathrm{d}v) = \Gamma_u^{(l)}(\mathrm{d}v). \tag{4.79}$$

(See Fig. 4.9.) For $k \in \mathbb{N}_0$, define the *effective interaction chain*

$$(M^{\mathrm{eff},k}_{-l})_{-l=-(k+1),-k,\dots,0} \tag{4.80}$$

as the *time-inhomogeneous* Markov chain on $[0,1] \times [0,1]$ with initial state

$$M^{\mathrm{eff},k}_{-(k+1)} = (\vartheta_k, \theta_{y_{k+1}}) \tag{4.81}$$

that evolves from time $-(l+1)$ to time $-l$ according to the transition kernel $Q^{[l]}$ on $[0,1] \times [0,1]$ given by

$$Q^{\mathrm{eff},[l]}(u, \mathrm{d}v) = \Gamma_{u_x}^{\mathrm{eff},(l)}(\mathrm{d}v), \tag{4.82}$$

where u_x denotes the first component of $u = (u_x, u_y)$.(See Fig. 4.8.) We denote the components of $\left(M^{\mathrm{eff},k}_{-l}\right)$ by

$$M^{\mathrm{eff},k}_{-l} = \left(M^{\mathrm{eff},k}_{-l,x}, M^{\mathrm{eff},k}_{-l,y}\right). \tag{4.83}$$

$$
\begin{array}{l}
M^{k,\mathrm{eff}}_{-(k+1)} = (\vartheta_k, \theta_{k+1}) \\
\quad \Big\downarrow \quad Q^{[k]}\left(M^{k,\mathrm{eff}}_{-(k+1)}, (du, dv)\right) = \Gamma^{\mathrm{eff}}_{\vartheta_k}(du, dv) \\
M^{k,\mathrm{eff}}_{-k} = (x_k(t_k), y_{k,k}(t_k)) \\
\quad \Big\downarrow \quad Q^{[k-1]}\left(M^{k,\mathrm{eff}}_{-k}, (du, dv)\right) = \Gamma^{\mathrm{eff}}_{x_k(t_k)}(du, dv) \\
M^{k,\mathrm{eff}}_{-(k-1)} = (x_{k-1}(t_{k-1}), y_{k-1,k-1}(t_{k-1})) \\
\quad \downarrow \quad Q^{[k-2]}\left(M^{k,\mathrm{eff}}_{-(k-1)}, (du, dv)\right) = \Gamma^{\mathrm{eff}}_{x_{k-1}(t_{k-1})}(du, dv) \\
\quad \vdots \\
\quad \downarrow \\
M^{k,\mathrm{eff}}_{-1} = (x_1(t_1), y_{1,1}(t_1)) \\
\quad \Big\downarrow \quad Q^{[0]}\left(M^{k,\mathrm{eff}}_{-1}, (du, dv)\right) = \Gamma^{\mathrm{eff}}_{x_1(t_1)}(du, dv) \\
M^{k,\mathrm{eff}}_{0} = (x_0(t_0), y_{0,0}(t_0))
\end{array}
$$

Figure 4.8: Effective interaction chain.

§4.4.3 Hierarchical multi-scale limit theorems

First we present and discuss the scaling of the effective process. Afterwards we do the same for the full process.

- **Effective process**

We present one of our main theorems, the hierarchical mean-field limit for the effective process. We will use the process and notation introduced in Section 4.4.2.

Theorem 4.4.2 (Hierarchical mean-field: the effective process). *Suppose that the initial state of the hierarchical system is given by* (4.21)*. Let* $L(N)$ *be such that* $\lim_{N\to\infty} L(N) = \infty$ *and* $\lim_{N\to\infty} L(N)/N = 0$*. For* $k \in \mathbb{N}$ *and* $t_k, \ldots, t_0 \in (0, \infty)$*, set* $\bar{t} = N^k L(N) + \sum_{n=0}^{k} N^n t_n$*.*

(a) For $k \in \mathbb{N}$*,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{eff},(l),\Omega_N}(\bar{t})\right)_{l=k+1,k,\ldots,0}\right] = \mathcal{L}\left[(M^{\mathrm{eff},k}_{-l})_{-l=-(k+1),-k,\ldots,0}\right]. \tag{4.84}$$

(b) *For $k \in \mathbb{N}$,*

$$
\begin{aligned}
l > k\colon\ & \lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\text{eff},(l),\Omega_N}(\bar{t}+N^k t)\right)_{t>0}\right] = \delta_{(\vartheta_l,\theta_{y_l})},\\
l = k\colon\ & \lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\text{eff},(l),\Omega_N}(\bar{t}+N^l t)\right)_{t>0}\right] = \mathcal{L}\left[\left(z^{\text{eff}}_{k,M^{\text{eff},k}_{-(k+1),x}}(t)\right)_{t>0}\right],\\
l < k\colon\ & \lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\text{eff},(l),\Omega_N}(\bar{t}+N^l t)\right)_{t>0}\right] = \mathcal{L}\left[\left(z^{\text{eff}}_{l,M^{\text{eff},k}_{-(l+1),x}}(t)\right)_{t>0}\right],
\end{aligned}
\tag{4.85}
$$

where the initial laws of the limiting processes are given by (see Fig. 4.8)

$$
\begin{aligned}
&\mathcal{L}\left[z^{\text{eff}}_{k,M^k_{-(k+1),x}}(0)\right] = \Gamma^{\text{eff},(k)}_{M^k_{-(k+1),x}},\\
&\mathcal{L}\left[z^{\text{eff}}_{k,M^k_{-(k+1),x}}(0)\right] = \Gamma^{\text{eff},(k)}_{M^k_{-(l+1),x}},\\
&\Gamma^{\text{eff},(l)}_{M^k_{-(l+1)}} = \int_{[0,1]^2}\cdots\int_{[0,1]^2}\Gamma^{\text{eff},(k)}_{M^k_{-(k+1)}}(\mathrm{d}u_k)\cdots\Gamma^{\text{eff},(l+1)}_{u_{l+2}}(\mathrm{d}u_{l+1})\Gamma^{\text{eff},(l)}_{u_{l+1}}
\end{aligned}
\tag{4.86}
$$

Theorem 4.4.2 can be interpreted as follows. The statement in (a) shows that if we look at the effective process on multiple space-time scales simultaneously, then the joint distribution of the different block averages is the law of the two-dimensional interaction chain defined in (4.80) and depicted in Fig. 4.8. Note that the process $(\boldsymbol{\Theta}^{\text{eff},(l),\Omega_N}(\bar{t}+N^l t))_{t>0}$ has at each level a different colour seed-bank average as second component, which is called *the effective seed-bank*. The statement in (b) describes the law of the path on different time scales.

- On time scale $\bar{t} + N^k t$ the l-block averages with $l > k$ are not moving, i.e., $(\boldsymbol{\Theta}^{\text{eff},(l),\Omega_N}(\bar{t} + N^k t))_{t>0}$ converges to the constant process taking the value $(\vartheta_l, \theta_{y_l}) = \boldsymbol{\Theta}^{\text{eff},(l),\Omega_N}(0)$.
- On time scale $\bar{t} + N^k t$ the k-block averages have reached equilibrium. The full k-block average feels a drift towards the full $(k+1)$-block average, which is still in its initial state ϑ_k. Therefore migration between the k-blocks in the hierarchical mean-field limit is replaced by a drift towards ϑ_k, and the k-blocks become independent. This phenomenon is called *decoupling* (or 'propagation of chaos'). The resampling function for the full estimator converges to $\mathcal{F}^{(k)}g$ (see (4.75)), the average diffusion function of the k-blocks. Finally, the full k-block exchanges individuals with the k-dormant population. Hence the k-dormant population is the effective seed-bank on space-timescale k Both the migration and the renormalisation are qualitatively similar to that of the hierarchical system without seed-bank [21]. However, the seed-bank still quantitatively affects the migration and the resampling through the slowing-down factor E_k. (In Section 4.4.4 we will see how the latter arises.)
- On time scale $\bar{t} + N^l t$ the l-block averages with $l < k$ are in a *quasi-equilibrium*. The full l-block averages feel a drift towards the instantaneous value of the $(l+1)$-block average at time $\bar{t}$. Therefore also the l-block averages decouple.

The $(l+1)$-block average is not moving on time scale $\bar{t}+N^l t$, and so for $t = L(N)$ we see that the l-block averages equilibrate faster than the $(l+1)$-block averages moves. The resampling function is given by $\mathcal{F}^{(l)}g$, which is to be interpreted as the average diffusion function of the l-blocks. The full average interacts with the l-blocks of the l-dormant population, which is the effective seed-bank on level l. Again the full l-block average feels a slowing-down factor E_l.

Note that Theorem 4.4.2 only describes the limiting process of the combined block average $\bar{\Theta}^{(l),\Omega_N}$ and the effective seed-bank $\Theta_{y_l}^{(l),\Omega_N}$. It does not provide a full description of the system, which is in Theorem 4.4.4 below. We will see later that Theorem 4.4.2 does describe all the *non-trivial* components of the system.

Remark 4.4.3 (Quasi equilibria). Note that Theorem 4.4.2 does not depend on the choice of $t_k, \ldots, t_0 \in (0,\infty)$. Since at each level $0 \leq l \leq k$ we start from time $\bar{t}$, the l-block averages have already reached a quasi-equilibrium. ■

$$
\begin{aligned}
M^k_{-(k+1)} &= (\vartheta_k, \overbrace{\vartheta_k, \cdots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots) \\
&\Big\downarrow \quad Q^{[k]}\left(M^k_{-(k+1)}, \mathrm{d}u\right) = \Gamma_{M^k_{-(k+1)}}(\mathrm{d}u) \\
M^k_{-k} &= (x_k(t_k), \overbrace{x_k(t_k), \cdots, x_k(t_k)}^{k \text{ times}}, y_{k,k}(t_k), \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots) \\
&\Big\downarrow \quad Q^{[k-1]}\left(M^k_{-k}, \mathrm{d}u\right) = \Gamma_{M^k_{-k}}(\mathrm{d}u) \\
M^k_{-(k-1)} &= \left(x_{k-1}(t_{k-1}), \overbrace{x_{k-1}(t_{k-1}), \cdots, x_{k-1}(t_{k-1})}^{k-1 \text{ times}}, y_{k-1,k-1}(t_{k-1}), y_{k,k}(t_k), \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots\right) \\
&\Big\downarrow \quad Q^{[k-2]}\left(M^k_{-(k-1)}, \mathrm{d}u\right) = \Gamma_{M^k_{-(k-1)}}(\mathrm{d}u) \\
&\ \vdots \\
&\Big\downarrow \\
M^k_{-1} &= \left(x_1(t_1), x_1(t_1), y_{1,1}(t_1), y_{2,2}(t_2), \cdots, y_{k-1,k-1}(t_{k-1}), y_{k,k}(t_k), \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots\right) \\
&\Big\downarrow \quad Q^{[0]}\left(M^k_{-1}, \mathrm{d}u\right) = \Gamma_{M^k_{-1}}(\mathrm{d}u) \\
M^k_0 &= \left(x_0(t_0), y_{0,0}(t_0), y_{1,1}(t_1), y_{2,2}(t_2), \cdots, y_{k-1,k-1}(t_{k-1}), y_{k,k}(t_k), \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots\right)
\end{aligned}
$$

Figure 4.9: Full interaction chain.

- **Full process**

To state our second main theorem, we will again use the process and the notation as defined in Section 4.4.2.

Theorem 4.4.4 (Hierarchical mean-field: full process). *Suppose that the initial state is given by* (4.21)*. Let* $L(N)$ *be such that* $\lim_{N\to\infty} L(N) = \infty$ *and* $\lim_{N\to\infty} L(N)/N = 0$. *For* $k \in \mathbb{N}$ *and* $t_k, \ldots, t_0 \in (0,\infty)$, *set* $\bar{t} = N^k L(N) + \sum_{n=0}^{k} N^n t_n$.

(a) For $k \in \mathbb{N}$,

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(l),\Omega_N}(\bar{t})\right)_{l=k+1,k,\dots,0}\right] = \mathcal{L}\left[(M^k_{-l})_{-l=-(k+1),-k,\dots,0}\right]. \tag{4.87}$$

(b) For $k \in \mathbb{N}$,

$$\begin{aligned} l > k\colon \ & \lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(l),\Omega_N}(\bar{t}+N^k t)\right)_{t>0}\right] = \delta_{(M^k_{-(k+1)})}, \\ l = k\colon \ & \lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(l),\Omega_N}(\bar{t}+N^l t)\right)_{t>0}\right] = \mathcal{L}\left[\left(z_{k,M^k_{-(k+1)}}(t)\right)_{t>0}\right], \\ l < k\colon \ & \lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(l),\Omega_N}(\bar{t}+N^l t)\right)_{t>0}\right] = \mathcal{L}\left[\left(z_{l,M^k_{-(l+1)}}(t)\right)_{t>0}\right], \end{aligned} \tag{4.88}$$

in the Meyer-Zheng topology,

where the initial laws of the limiting processes are given by (see Fig. 4.9)

$$\begin{aligned} &\mathcal{L}\left[z_{k,M^k_{-(k+1),x}}(0)\right] = \Gamma^{(k)}_{M^k_{-(k+1)}}, \\ &\mathcal{L}\left[z_{k,M^k_{-(k+1),x}}(0)\right] = \Gamma^{(l)}_{M^k_{-(l+1)}}, \\ &\Gamma^{(l)}_{M^k_{-(l+1)}} = \int_{\mathfrak{s}} \cdots \int_{\mathfrak{s}} \int_{\mathfrak{s}} \Gamma^{(k)}_{M^k_{-(k+1)}}(\mathrm{d}u_k)\Gamma^{(k-1)}_{u_k}(\mathrm{d}u_{k-1}) \cdots \Gamma^{(l+1)}_{l+2}(\mathrm{d}u_1)\Gamma^{(l)}_{u_{l+1}} \end{aligned} \tag{4.89}$$

Remark 4.4.5 (Convergence in the Meyer-Zheng topology). Note that Theorem 4.4.4(b) is stated in the Meyer-Zheng topology. This topology is needed because at time-scales $N^l t$ rates occur in(4.24) that tend to infinity as $N \to \infty$. In Section 4.4.4 we define the Meyer-Zheng topology and explain why it is neeeded. ■

The statement in (a) shows that if we look at multiple space-time scales simultaneously, then the joint distribution of the different block averages behaves like the infinite-dimensional interaction chain defined in (4.77). The statement in (b) describes the law of the path on different times scales.

- On time scale $N^k t$, the l-block averages with $l > k$ are not moving, i.e., $(\mathbf{\Theta}^{(l),\Omega_N}(\bar{t}+N^k t))_{t>0}$ is a constant process. However, there is a difference between seed-banks with colour $m > k$ and seed-banks with colour $0 \leq m \leq k$ in the way they interact with the active population. For $m > k$, even the m-dormant single colonies have not yet moved at time $\bar{t}+N^k t$, and hence are still in their initial states, with expectations $(\theta_{y_m})_{m=l+1}^{\infty}$. Therefore, also the l-block averages of m-dormant populations are still in their initial states, with expectations $(\theta_{y_m})_{m=l+1}^{\infty}$. For $l \leq k$ the m-dormant single colonies with $m \leq k$ at time $\bar{t}+N^k t$ have already interacted with the active population. Due to this interaction, for $l > k$ the l-block averages of m-dormant populations with $m \leq k$ are in state ϑ_k instead of their initial state θ_{y_m}. However, on time scale $\bar{t}+N^k t$

l-block averages of m-dormant populations are not moving. (In Section 4.4.4 we explain how the shift from θ_{y_m} to ϑ_k occurs.) Also the l-block averages of the active population are in state ϑ_k.

- On time scale $\bar{t} + N^k t$, the k-block averages have reached equilibrium. We see that the active k-block average and the k-dormant k-block average evolve together like the effective k-block process in Theorem 4.4.2. Therefore the evolution of the active k-block average is slowed down by a factor E_k, the active k-block feels a drift towards ϑ_k (the value of the active $(k+1)$-block average at time $\bar{t}$), resamples with diffusion function $\mathcal{F}^{(k)} g$, and exchanges individuals with the k-dormant k-block. The k-dormant k-block average evolves only via interaction with the active k-block. The m-dormant k-block averages with colour $0 \leq m < k$ equal the active k-block average and hence follow their evolution. The m-dormant k-blocks with colour $m > k$ are still in their initial states, since on time scale $\bar{t} + N^k t$ even single colony seed-banks with colour $m > k$ have not yet started to interact with the active population.

- On time scale $\bar{t} + N^l t$, for $0 \leq l < k$, the l-block averages are in a quasi-equilibrium. Again, the active l-block and the l-dormant l-block average, which is the effective seed-bank, behave as the effective process in 4.4.2. Hence, the active l-block average feels a drift towards the instantaneous value of the active $(l+1)$-block average, which is given by the first component of the interaction chain $M^k_{-(l+1)}$, resamples according to the renormalised diffusion function $\mathcal{F}^{(l)} g$, and exchanges with the l-block of the l-dormant population. The evolution of the active l-block average is slowed down by a factor E_l. The l-block of the l-dormant population exchanges individuals with the active population. The l-blocks of the m-dormant population with colours $0 \leq m < l$ follow the active population. The states of the m-dormant population with colour $m > l$ are given by the corresponding components in the interaction chain $M^k_{-(l+1)}$. Hence the l-block averages with colours $m > k$ are still in their initial states θ_{y_m}, because on time scale $\bar{t} + N^l t$ even the single dormant colonies with colour $m > k$ have not yet interacted with the active population. However, something interesting is happening with the colours $l < m \leq k$: they are in a state obtained on the time scale in which they where effective, i.e., for $l < m \leq k$ the m-dormant l-block average is in state $y_{m,m}(\bar{t})$. This happens because at time $\bar{t}$ the single colonies have already interacted with the active population, but on time scale $N^l t$ they do not interact anymore with the active population. (Also this effect will be further explained in Section 4.4.4.)

Remark 4.4.6 (Comparison to system without seed-bank). Comparing Theorem 4.4.4 with the multi-scale limit theorems derived for the hierarchical system without seed-bank [21], [20], [25], we see that the seed-bank affects the system both quantitatively and qualitatively. First, the active population is *slowed down* by the total size of the seed-banks it has interacted with, represented by the slowing-down factors $(E_l)_{l\in\mathbb{N}_0}$. Second, the interaction with the *effective seed-bank* on each space-time scale is special to the system with seed-bank. Still, the *decoupling* of the active component and the *renormalisation transformation* for the diffusion function are similar as in the system without seed-bank.

Remark 4.4.7 ($k \to \infty$ limit of the interaction chain). The result in (4.85) raises the question how the hierarchical multi-scale limit behaves for large k. We find the following dichotomy:

$$\lim_{k\to\infty} \mathcal{L}\left[(M^k_{-(k+1),-k,\cdots,0})\right] = \mathcal{L}\left[(M_k^\infty)_{k\in\mathbb{Z}^-}\right], \tag{4.90}$$

where in the *clustering regime*

$$\mathcal{L}\left[(M_k^\infty)_{k\in\mathbb{Z}^-}\right] = \theta\delta_{(1,1^{\mathbb{N}_0})^{\mathbb{Z}^-}} + (1-\theta)\delta_{(0,0^{\mathbb{N}_0})^{\mathbb{Z}^-}} \tag{4.91}$$

and in the *coexistence regime* $M^\infty = (M_k^\infty)_{k\in\mathbb{Z}^-}$ is a realisation of the unique entrance law of the interaction chain at time $-\infty$ with

$$\lim_{l\to\infty} M_{-l}^\infty = (\theta,\theta). \tag{4.92}$$

In the latter case, M^∞ corresponds to the equilibrium vector of block averages around site 0, whose law agrees with that of the equilibrium block averages for the mean-field model after we take the limit $N \to \infty$ (see [25, Proposition 6.2 and 6.3]). ■

Remark 4.4.8 (Interaction field). Theorem 4.4.4 looks at the tower of block averages over a fixed site, namely, 0. In order to study the cluster formation in the clustering regime or the equilibria in the coexistence regime, we must analyse the dependence structure between the towers of block averages over *different* sites. We can show that, in the limit as $N \to \infty$, an interacting random field emerges, indexed by a tree with countably many edges coming out of every site at every level. This random field has the property that the averages over any two points η,η' with $d(\eta,\eta') = l$, follow a single interaction chain in equilibrium from $k+1$ until l (or from $-\infty$ until l in the entrance law) and, conditional on the state in l, evolve independently as the interaction chain beyond l. This corresponds to what is called *propagation of chaos* of the $(l-1)$-block averages given the l-block average. For the model without seed-bank such results are described in [25, Section 0(e)]. Using our results for the model with seed-bank above, we can in principle follow an analogous line of argument. We refrain from spelling out the details. ■

§4.4.4 Heuristics behind the multi-scale limit

The proofs of Theorems 4.4.2 and 4.4.4 written out in Sections 6.1–9, are long and technical. In order to help the reader appreciate these proofs, we provide the heuristics in this section.

- **Evolution of the block-averages**

Recall the block averages introduced in Definition 4.2.3 and their evolution defined in (4.24). In the limit as $N \to \infty$, we heuristically obtain from the SSDE in (4.24) the following results for the k-block process

$$(x_k^{\Omega_N}(t), (y_{m,k}^{\Omega_N}(t))_{m\in\mathbb{N}_0})_{t\geq 0}. \tag{4.93}$$

- **Migration.** Recall that the migration is captured by the first term of the first equation in (4.24), i.e., the first term of the evolution of the active part of the population. Letting $N \to \infty$, we see that in the sum over l only the term $l = 1$ contributes. Hence we expect that the effective migration felt by the active k-block average is towards $x^{\Omega_N}_{k+1}(0)$, the initial state of the active $(k+1)$-block average. Note that the migration term in (4.24) can be written as

$$\sum_{l\in\mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[x^{\Omega_N}_{k+l}(N^{-l}t) - x^{\Omega_N}_k(t)\right] = \sum_{l\in\mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[\frac{1}{N^l}\sum_{k=0}^{N^l-1} x^{\Omega_N}_k(t) - x^{\Omega_N}_k(t)\right]. \tag{4.94}$$

  The drift towards the $(k+1)$-block average is therefore also a drift towards the current average of the k-blocks in the $(k+1)$-block. In the limit as $N \to \infty$, the latter can be approximated by $\mathbb{E}[x_k(t)]$. Effectively, as $N \to \infty$, the k-blocks become independent given the the value of $x^{\Omega_N}_{k+1}(0)$, i.e., there is *decoupling.*

- **Resampling.** Recall that the diffusion term in the evolution equation of the active population represents the resampling. Therefore we see that the active k-block resamples at a rate that is the average resampling rate over the k-block. For $k = 1$, the resampling rate of the 1-block is the average of the resampling ratse of the single colonies. Therefore, in the limit $N \to \infty$, due to the decoupling described above, we expect that the resampling rate for the 1 block is given by $\mathbb{E}[g]$, where the expectation is w.r.t. the quasi-equilibrium of the single colonies. This expectation is exaclty the renormalised diffusion function $\mathcal{F}g$ (see (4.75)). For the k-block, we may interpret the diffusion function to be the average of the diffusion function for the $(k-1)$-blocks. By "induction" we assume that the $(k-1)$ blocks resample at rate $\mathcal{F}^{(k-1)}g$. Hence, due to the decoupling of the $(k-1)$-blocks as $N \to \infty$, we expect the resampling rate for the k blocks to equal $\mathbb{E}[\mathcal{F}^{(k-1)}g]$, where the expectation is w.r.t. the quasi-equlibrium of the $(k-1)$-blocks. This yields another iteration of the renormalisation transformation (see (4.76)). Hence, we expect the diffusion function for the k-blocks to converge to $\mathcal{F}^{(k)}g$.

- **Exchange with the seed-bank.** Recall that the last term of the first equation in (4.24) and the second equation in (4.24) together describe the exchange of the active k-block with the m-dormant k-block. To describe the limiting behaviour as $N \to \infty$, we distinguish three cases: $0 \le m < k$, $m = k$, $m > k$.

  - If $0 \le m < k$, then we see that the rate of exchange between the active k-block and the m-dormant k-block tends to infinity as $N \to \infty$. We therefore expect them to equalise, i.e.,

$$\lim_{N\to\infty} \mathcal{L}\left[\left(x^{\Omega_N}_k(t) - y^{\Omega_N}_{m,k}(t)\right)_{t>0}\right] = \delta_0, \tag{4.95}$$

    where 0 should be interpreted as the process equal to 0, $(0)_{t>0}$. Hence we see that m-dormant k-block follows the active k-block immediately. (To formalise this fact, we need the *Meyer-Zheng topology* [59].)

  - If $m = k$, then there is a non-trivial exchange between the active k-block and the k-dormant k-block.

– If $m > k$, then the exchange rate between the active k-block and the m-dormant k-block tends to zero as $N \to \infty$.

Thus, only the k-dormant k-block has a non-trivial interaction with the active k-block. We express this by saying that on space-time scale k the k-dormant population plays the role of the *effective seed-bank*.

• **Limiting evolution of the block-averages**

To determine the limiting evolution of the full block-averages process, we first have a look at the limiting evolution of the effective process.

The effective process. To determine the limit as $N \to \infty$ of (4.24), we need to get rid of the diverging rates. Instead of only looking at the k-block process $(x_k^{\Omega_N}(t), (y_{m,k}^{\Omega_N}(t))_{m\in\mathbb{N}_0})_{t\geq 0}$, which evolves according to (4.24), we look at the *effective k-block process* defined as

$$\left(\bar{x}_k^{\Omega_N}(t), y_{k,k}^{\Omega_N}(t)\right)_{t\geq 0}, \tag{4.96}$$

where we abbreviate

$$\bar{x}_k^{\Omega_N}(t) = \frac{x_k^{\Omega_N}(t) + \sum_{m=0}^{k-1} K_m y_{m,k}^{\Omega_N}(t)}{1 + \sum_{m=0}^{k-1} K_m}. \tag{4.97}$$

By (4.95) and the heuristic discussion given above, the process in (4.96) equals $(x_k^{\Omega_N}(t), y_{k,k}^{\Omega_N}(t))_{t\geq 0}$ in the limit as $N \to \infty$, i.e., it describes the joint distribution of the active k-block and the effective dormant k-block, which is the k-dormant k-block. Using (4.24), we see that the process in (4.96) evolves according to the SSDE

$$\begin{aligned}
\mathrm{d}\bar{x}_k^{\Omega_N}(t) &= E_k \sum_{l\in\mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[x_{k+l}^{\Omega_N}(N^{-l}t) - x_k^{\Omega_N}(t)\right] \mathrm{d}t \\
&\quad + E_k \sqrt{\frac{1}{N^k} \sum_{\xi\in B_k(0)} g(x_\xi(N^k t))}\ \mathrm{d}w_k(t) \\
&\quad + E_k \sum_{m=k}^{\infty} N^k \frac{K_m e_m}{N^m} \left[y_{m,k}^{\Omega_N}(t) - x_k^{\Omega_N}(t)\right] \mathrm{d}t, \\
\mathrm{d}y_{k,k}^{\Omega_N}(t) &= e_k \left[x_k^{\Omega_N}(t) - y_{k,k}^{\Omega_N}(t)\right] \mathrm{d}t.
\end{aligned} \tag{4.98}$$

In (4.98) no infinite rates appear anymore. In the limit as $N \to \infty$, by (4.95) we can approximate

$$x_k^{\Omega_N}(t) \approx y_{m,k}^{\Omega_N}(t), \qquad 0 \leq m < k, \tag{4.99}$$

such that

$$x_k^{\Omega_N}(t) \approx \bar{x}_k^{\Omega_N}(t). \tag{4.100}$$

We can therefore approximate (4.98) by

$$
\begin{aligned}
\mathrm{d}\bar{x}_k^{\Omega_N}(t) &= E_k \sum_{l\in\mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[\bar{x}_{k+l}^{\Omega_N}(N^{-l}t) - \bar{x}_k^{\Omega_N}(t)\right] \mathrm{d}t \\
&\quad + E_k \sqrt{\frac{1}{N^k} \sum_{\xi\in B_k(0)} g\big(\bar{x}_\xi(N^k t)\big)}\, \mathrm{d}w_k(t) \\
&\quad + E_k \sum_{m=k}^{\infty} N^k \frac{K_m e_m}{N^m} \left[y_{m,k}^{\Omega_N}(t) - \bar{x}_k^{\Omega_N}(t)\right] \mathrm{d}t, \\
\mathrm{d}y_{k,k}^{\Omega_N}(t) &= e_k \left[\bar{x}_k^{\Omega_N}(t) - y_{k,k}^{\Omega_N}(t)\right] \mathrm{d}t.
\end{aligned}
\tag{4.101}
$$

Hence, in the limit as $N \to \infty$, the process in (4.96) becomes autonomous. Moreover, assuming that $\lim_{N\to\infty} \bar{x}_{l+1}^{\Omega_N}(0) = \vartheta_l$, we see that (4.101) approaches the effective process defined in (4.69), with

$$
\theta = \vartheta_l, \quad E = E_k, \quad c = c_k, \quad e = e_k, \quad K = K_k, \quad g = \mathcal{F}^{(k)} g. \tag{4.102}
$$

In particular, we see that the *slowing-down constant* E_k arises because the active population is the only part of the first component of (4.96). Note that therefore only a part, the active part, from the first component migrates, resamples and exchanges with the seed-bank. Due to the infinite rates, the active population "drags along all the fast seed-banks with total size $\sum_{m=0}^{k-1} K_m$". This causes the slowing down factors E_k.

Since there are no infinite rates in the evolution of the effective process, we can use the classical path space topology. This allows us in the proof in Sections 6.1–9 to build on techniques developed for the hierarchical mean-field model without seed-bank in [20], [25]. It turns out that the *effective process* is very useful in our analysis.

From the effective process to the full process. For large N, by (4.99) and (4.100), the evolution of our original process $(x_k^{\Omega_N}(t), (y_{m,k}^{\Omega_N}(t))_{m\in\mathbb{N}_0})_{t\geq 0}$ can be approximated by

$$
\begin{aligned}
\mathrm{d}\bar{x}_k^{\Omega_N}(t) &\approx E_k \sum_{l\in\mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[\bar{x}_{k+l}^{\Omega_N}(N^{-l}t) - \bar{x}_k^{\Omega_N}(t)\right] \mathrm{d}t \\
&\quad + E_k \sqrt{\frac{1}{N^k} \sum_{\xi\in B_k(0)} g(\bar{x}_\xi(N^k t))}\, \mathrm{d}w_k(t) \\
&\quad + E_k \sum_{m=k}^{\infty} N^k \frac{K_m e_m}{N^m} \left[y_{m,k}^{\Omega_N}(t) - \bar{x}_k^{\Omega_N}(t)\right] \mathrm{d}t, \\
y_{k,k}^{\Omega_N}(t) &= \bar{x}_k^{\Omega_N}(t), \\
\mathrm{d}y_{k,k}^{\Omega_N}(t) &= e_k \left[\bar{x}_k^{\Omega_N}(t) - y_{k,k}^{\Omega_N}(t)\right] \mathrm{d}t, \\
y_{k,k}^{\Omega_N}(t) &= y_{m,k}^{\Omega_N}(0).
\end{aligned}
\tag{4.103}
$$

By the ergodic theorem for exchangeable measures, we can assume that

$$\lim_{N\to\infty} \bar{x}_{k+1}^{\Omega_N}(0) = \vartheta_k \; a.s. \tag{4.104}$$

We expect that (4.103) approaches (4.67) with

$$\theta = \vartheta_k, \quad E = E_k, \quad c = c_k, \quad e = e_k, \quad K = K_k, \quad g = \mathcal{F}^{(k)} g. \tag{4.105}$$

To prove that

$$\lim_{N\to\infty} \mathcal{L}[y_{m,k}^{\Omega_N}(t)] = \lim_{N\to\infty} \mathcal{L}[\bar{x}_k^{\Omega_N}(t)], \qquad 0 \le m < k-1, \tag{4.106}$$

we need *the Meyer-Zheng topology* explained in Section 4.4.1. In the proof in Sections 6.2–9 we show how the above approximations can be made rigorous.

Conserved quantities. Note that, by (4.24) and (4.94), for each $k \in \mathbb{N}_0$

$$\mathbb{E}\left[\frac{x_k^{\Omega_N}(t) + \sum_{m\in\mathbb{N}_0} K_m y_k^{\Omega_N}(t)}{1 + \sum_{m\in\mathbb{N}_0} K_m}\right] = \theta, \qquad t \ge 0, \tag{4.107}$$

is a conserved quantity. For each $k \in \mathbb{N}$ we obtain that for $l \ge k$

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{x_k^{\Omega_N}(t) + \sum_{m=0}^{l} K_m y_k^{\Omega_N}(t)}{1 + \sum_{m=0}^{l} K_m}\right] = \vartheta_l, \tag{4.108}$$

is a conserved quantity. For the effective process, (4.101) implies that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\bar{x}_k^{\Omega_N}(t)] &= E_k K_k e_k \left(\mathbb{E}[y_{k,k}^{\Omega_N}(t)] - \mathbb{E}[\bar{x}_k^{\Omega_N}(t)]\right), \\ \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[y_{k,k}^{\Omega_N}(t)] &= e_k \left(\mathbb{E}[\bar{x}_k^{\Omega_N}(t)] - \mathbb{E}[y_{k,k}^{\Omega_N}(t)]\right). \end{aligned} \tag{4.109}$$

Recall that

$$\mathbb{E}[\bar{x}_k^{\Omega_N}(0)] = \mathbb{E}\left[\frac{x_k^{\Omega_N}(0) + \sum_{m=0}^{k-1} K_m y_{m,k}^{\Omega_N}(0)}{1 + \sum_{m=0}^{k-1} K_m}\right] = \vartheta_{k-1}, \quad \mathbb{E}[y_{k,k}^{\Omega_N}(0)] = \theta_{y_k}. \tag{4.110}$$

Therefore we can solve (4.109) explicitly as

$$\begin{aligned} \mathbb{E}[\bar{x}_k^{\Omega_N}(t)] &= \vartheta_k + \frac{E_k K_k}{1 + E_k K_k}(\vartheta_{k-1} - \theta_{y_k})\,\mathrm{e}^{-(E_k K_k + 1)e_k t}, \\ \mathbb{E}[y_{k,k}^{\Omega_N}(t)] &= \vartheta_k - \frac{1}{1 + E_k K_k}(\vartheta_{k-1} - \theta_{y_k})\,\mathrm{e}^{-(E_k K_k + 1)e_k t}. \end{aligned} \tag{4.111}$$

The above computation shows what happens to $\mathbb{E}[\bar{x}_k^{\Omega_N}(t)]$ if we move one space-time scale up in the hierarchy, namely, a new seed-bank starts interacting with the active population. This causes that ϑ_{k-1} is pulled a bit towards θ_{y_k}, so that also $\mathbb{E}[\bar{x}_k^{\Omega_N}(t)]$ changes a bit. Each new seed-bank that opens up changes the expectation of the active population, which results in the sequence $(\theta_x, \vartheta_0, \vartheta_1, \vartheta_2, \ldots)$ for the expectation of the active population on space-time scales $\{0,1,2,3,\cdots\}$. From (4.94) we see that the drift of $\bar{x}_k^{\Omega_N}(t)$ is towards

$$\bar{x}_{k+1}^{\Omega_N}(N^{-1}t) = \frac{1}{N^k}\sum_{k=0}^{N^k-1} x_k^{\Omega_N}(t) \approx \mathbb{E}[\bar{x}_k^{\Omega_N}(t)], \tag{4.112}$$

where the last approximation can be made because the k-blocks decouple. Hence, in the limit as $N\to\infty$, once the k-blocks are in a quasi-equilibrium we can replace the drift towards $\bar{x}_{k+1}^{\Omega_N}(N^{-1}t)$ by a drift towards $\mathbb{E}[\bar{x}_k^{\Omega_N}(t)] = \vartheta_k$.

Shifting averages. Recall the full estimator process $(\mathbf{\Theta}^{(l),\Omega_N}(t))_{t>0}$ defined in (4.71). Equation 4.23 implies that the evolution of the estimator process is given by

$$\begin{aligned} \mathrm{d}\Theta_x^{(l),\Omega_N}(t) &= \sum_{n=l+1}^{\infty}\frac{c_{n-1}}{N^{n-1}}[\Theta_x^{(n),\Omega_N}(t)-\Theta_x^{(l)}(t)]\,\mathrm{d}t \\ &\quad + \sqrt{\frac{1}{N^{2l}}\sum_{\xi\in B_l} g\big(x_\xi(t)\big)}\,\mathrm{d}w(t) \\ &\quad + \sum_{m\in\mathbb{N}_0}\frac{K_m e_m}{N^m}[\Theta_{y_m}^{(l),\Omega_N}(t)-\Theta_x^{(1),\Omega_N}(t)]\,\mathrm{d}t, \\ \mathrm{d}\Theta_{y_m}^{(l),\Omega_N}(t) &= \frac{e_m}{N^m}[\Theta_x^{(l),\Omega_N}(t)-\Theta_{y_m}^{(l),\Omega_N}(t)]\,\mathrm{d}t,\ m\in\mathbb{N}_0. \end{aligned} \tag{4.113}$$

Looking at the estimator process $(\mathbf{\Theta}^{(l),\Omega_N}(t))_{t>0}$ on time scale $N^k t$, we see that

$$\begin{aligned} \mathrm{d}\Theta_x^{(l),\Omega_N}(N^k t) &= \sum_{n=l+1}^{\infty}\frac{c_{n-1}}{N^{n-1-k}}\left[\Theta_x^{(n),\Omega_N}(N^k t)-\Theta_x^{(l),\Omega_N}(N^k t)\right]\,\mathrm{d}t \\ &\quad + \sqrt{\frac{N^k}{N^{2l}}\sum_{\xi\in B_l} g\big(x_\xi(N^k t)\big)}\,\mathrm{d}w(t) \\ &\quad + \sum_{m\in\mathbb{N}_0}\frac{K_m e_m}{N^{m-k}}\left[\Theta_{y_m}^{(l),\Omega_N}(N^k t)-\Theta_x^{(l),\Omega_N}(N^k t)\right]\,\mathrm{d}t, \\ \mathrm{d}\Theta_{y_m}^{(l),\Omega_N}(N^k t) &= \frac{e_m}{N^{m-k}}\left[\Theta_x^{(l),\Omega_N}(N^k t)-\Theta_{y_m}^{(l),\Omega_N}(N^k t)\right]\,\mathrm{d}t, \qquad m\in\mathbb{N}_0. \end{aligned} \tag{4.114}$$

From (4.114) we get that, on time scale $N^k t$, for all $l\geq k+1$,

$$\Theta^{(l),\Omega_N}(N^k t) = \frac{\Theta_x^{(l),\Omega_N}(N^k t)+\sum_{m=0}^{l-1}K_m\Theta_{y_m}^{(l),\Omega_N}(N^k t)}{1+\sum_{m=0}^{l-1}K_m}, \qquad t\geq 0, \tag{4.115}$$

is a conserved quantity in the limit as $N\to\infty$, and for $t\geq 0$,

$$\lim_{N\to\infty}\Theta^{(l),\Omega_N}(N^k t) = \frac{\theta_x+\sum_{m=0}^{l-1}K_m\theta_{y_m}}{1+\sum_{m=0}^{l-1}K_m} = \vartheta_l,\ \text{in probability}. \tag{4.116}$$

For $m > k$, also $\Theta^{(l),\Omega_N}_{y_m}(N^k t)$ is a conserved quantity, and

$$\lim_{N\to\infty} \Theta^{(l),\Omega_N}_{y_m}(N^k t) = \theta_{y_m}, \qquad t \geq 0. \tag{4.117}$$

However, for $l \geq k+1$, $\Theta^{(l),\Omega_N}_{x}(N^k t)$ and $(\Theta^{(l),\Omega_N}_{y_m}(N^k t))_{m=0}^{k}$ are not conserved quantities in the limit as $N \to \infty$. Note that from 4.24 we heuristically see that the full l-block estimator process $(\boldsymbol{\Theta}^{(l),\Omega_N}(N^k t))_{t\geq 0}$ with $l > k$ converges to the process

$$\left(\Theta^{(l)}_x(t), \left(\Theta^{(l)}_{y_m}(t)\right)_{m\in\mathbb{N}_0}\right)_{t>0}, \tag{4.118}$$

which evolves according to

$$\begin{aligned}
&\mathrm{d}\Theta^{(l)}_x(t) = E_{k-1}K_k e_k[\Theta^{(l)}_{y_k}(t) - \Theta^{(l),\Omega_N}_x(t)]\,\mathrm{d}t,\\
&\Theta^{(l)}_{y_m}(t) \;=\; \Theta^{(l)}_x(t), \qquad m < k,\\
&\mathrm{d}\Theta^{(l)}_{y_k}(t) \;=\; e_k[\Theta^{(l)}_x(t) - \Theta^{(l)}_{y_k}(t)]\,\mathrm{d}t, \quad m = k,\\
&\Theta^{(l)}_{y_m}(t) \;=\; \theta_{y_m}, \qquad m > k.
\end{aligned} \tag{4.119}$$

This system can be explicitly solved as

$$\begin{aligned}
\Theta^{(l)}_x(t) &= \frac{\Theta^{(l)}_x(0) + E_{k-1}K_k\Theta^{(l)}_{y_k}(0)}{1+E_{k-1}K_k}\\
&\quad + \frac{E_{k-1}K_k}{1+E_{k-1}K_k}[\Theta^{(l),\Omega_N}_x(0) - \Theta^{(l)}_{y_k}(0)]\,\mathrm{e}^{-(E_{k-1}K_k e_k + e_k)t},\\
\Theta^{(l)}_{y_m}(t) &= \Theta^{(l)}_x(t), \qquad m < k,\\
\Theta^{(l)}_{y_k}(t) &= \tfrac{\Theta^{(l)}_x(0)+E_{k-1}K_k\Theta^{(l)}_{y_k}(0)}{1+E_{k-1}K_k}\\
&\quad - \tfrac{1}{1+E_{k-1}K_k}[\Theta^{(l),\Omega_N}_x(0) - \Theta^{(l)}_{y_k}(0)]\,\mathrm{e}^{-(E_{k-1}K_k e_k + e_k)t}, \quad m = k,\\
\Theta^{(l)}_{y_m}(t) &= \theta_{y_m}, \qquad m > k.
\end{aligned} \tag{4.120}$$

The latter shows that, each time we enter a new space-time scale, all the large active blocks interact with the large effective dormant blocks until they equalise. Thus, on each space-time scale, all the active l-blocks and the dormant l-blocks of colour $m \leq l$ move for a short period of time. As a consequence, the value of $\Theta^{(l)}_x(0)$ depends on the scaling we choose. To illustrate this, we note that

$$\begin{aligned}
\lim_{N\to\infty} \Theta^{(l),\Omega_N}_x(0) &= \theta_x,\\
\lim_{N\to\infty} \Theta^{(l),\Omega_N}_x(L(N)+t) &= \vartheta_0,\\
\lim_{N\to\infty} \Theta^{(l),\Omega_N}_x((L(N)N^n + N^n t)) &= \vartheta_n, \quad 0 \leq n \leq k.
\end{aligned} \tag{4.121}$$

Hence under the scaling $N^k t$ the $\lim_{N\to\infty} \frac{L(N)N^n+N^n t}{N^k} = 0$ and therefore for consistentcy one would like to have

$$\begin{aligned}
\lim_{N\to\infty} \Theta^{(l),\Omega_N}_x(0) &= \lim_{N\to\infty} \Theta^{(l),\Omega_N}_x\left(N^k \frac{L(N)+t}{N^k}\right)\\
&= \lim_{N\to\infty} \Theta^{(l),\Omega_N}_x\left(N^k \frac{L(N)N^n + N^n t}{N^k}\right) \quad 0 \leq n \leq k,
\end{aligned} \tag{4.122}$$

but this contradicts with (4.121). Hence, if $(t(N))_{N\in\mathbb{N}}$ is a sequence such that $\lim_{N\to\infty} N^k t(N) = 0$, then the value of the limit

$$\lim_{N\to\infty} \Theta_x^{(l),\Omega_N}(N^k t(N)) \tag{4.123}$$

depends on $(t(N))_{N\in\mathbb{N}}$. Moreover, to obtain a limiting process for $(\boldsymbol{\Theta}^{(l),\Omega_N}(N^k t))_{t\geq 0}$ we need convergence also at time 0, while it is not clear what $N^k t \downarrow 0$ means. To circumvent these subtleties, we look at the process at times $t > 0$ and use as starting time $\bar{t}$ defined in Theorem 4.4.2.

From (4.120) it follows that, for $l \geq k$,

$$\begin{aligned}
&\lim_{N\to\infty} \Theta_x^{(l),\Omega_N}(\bar{t}\,) = \vartheta_k, \text{ in probability,}\\
&\lim_{N\to\infty} \Theta_{y_m}^{(l),\Omega_N}(\bar{t}\,) = \vartheta_k, \quad m \leq k, \text{ in probability,}\\
&\lim_{N\to\infty} \Theta_{y_m}^{(l),\Omega_N}(\bar{t}\,) = \theta_{y_m}, \quad m > k \text{ in probability.}
\end{aligned} \tag{4.124}$$

Note that the shifting of averages mentioned earlier is closely related to the conserved quantities discussed in Section 4.4.4 because, for large N,

$$\lim_{N\to\infty} \Theta_x^{(l),\Omega_N}(\bar{t}\,) \approx \mathbb{E}[x_k^{\Omega_N}], \tag{4.125}$$

where the expectation is taken in the quasi-equilibrium the k-blocks have attained after scaling with time $\bar{t}$.

• Formation of the interaction chain

In Section 4.4.4 we saw how subsequent space-time scales are connected via the migration term. In this section we show how the interaction chain arises from this connection. We first show how the effective interaction chain arises for the effective process. Then we show how the full interaction chain is formed, by studying the slow seed-banks.

Connections between different space-time scales Let $\bar{t}$ be as in Theorem 4.4.2. From (4.113) it follows that the process $(\boldsymbol{\Theta}^{(l),\Omega_N}(\bar{t}+N^l t))_{t>0}$ evolves according to

$$\begin{aligned}
\mathrm{d}\Theta_x^{(l),\Omega_N}(\bar{t}+N^l t) &= \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1-l}}[\Theta_x^{(n),\Omega_N}(\bar{t}+N^l t) - \Theta_x^{(l)}(\bar{t}+N^l t)]\,\mathrm{d}t\\
&\quad + \sqrt{\frac{1}{N^l}\sum_{\xi\in B_l} g\big(x_\xi(\bar{t}+N^l t)\big)}\,\mathrm{d}w(t)\\
&\quad + \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{N^{m-l}}[\Theta_{y_m}^{(l),\Omega_N}(\bar{t}+N^l t) - \Theta_x^{(1),\Omega_N}(\bar{t}+N^l t)]\,\mathrm{d}t,\\
\mathrm{d}\Theta_{y_m}^{(l),\Omega_N}(\bar{t}+N^l t) &= \frac{e_m}{N^{m-l}}[\Theta_x^{(l),\Omega_N}(\bar{t}+N^l t) - \Theta_{y_m}^{(l),\Omega_N}(\bar{t}+N^l t)]\,\mathrm{d}t, \qquad m\in\mathbb{N}_0.
\end{aligned} \tag{4.126}$$

Therefore, in the limit as $N \to \infty$, the active population $\Theta_x^{(l),\Omega_N}(\bar{t}+N^l t)$ feels a drift towards the $(l+1)$-block average $\Theta_x^{(l+1),\Omega_N}(\bar{t}+N^{l+1}t)$. If $l=k$, then

$$\lim_{N\to\infty} \mathcal{L}[\boldsymbol{\Theta}^{(k+1),\Omega_N}(\bar{t}+N^k t)] = \lim_{N\to\infty} \mathcal{L}[\boldsymbol{\Theta}^{(k+1),\Omega_N}(\bar{t})], \tag{4.127}$$

since the $(k+1)$-block has not yet started to move at time $\bar{t}+N^k t$. From (4.124) it follows that

$$\lim_{N\to\infty} \Theta_x^{(k+1),\Omega_N}(\bar{t}+N^k t) = \vartheta_k \, in \, probability. \tag{4.128}$$

Therefore the drift of the active population $\Theta_x^{(l),\Omega_N}(\bar{t}+N^l t)$ is towards ϑ_k. Since $\bar{t} > L(N)N^k$, the process $\boldsymbol{\Theta}^{(k),\Omega_N}(\bar{t}+N^k t)$ has, in the limit $N \to \infty$, already reached its equilibrium, which is denoted by $\Gamma_{\bar{\vartheta}_k}$, where

$$\bar{\vartheta}_k = (\vartheta_k, \overbrace{\vartheta_k, \cdots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y_{k+1}}, \theta_{y_{k+2}}, \cdots), \tag{4.129}$$

so that we recognise $(\bar{\vartheta}_k) = M^k_{-(k+1)}$. From (4.124) with $l=k+1$ we see that $(\bar{\vartheta}_k) = M^k_{-(k+1)}$ represents the state of $\boldsymbol{\Theta}^{(k+1),\Omega_N}(\bar{t})$.

If we look on time scale $\bar{t}+N^{k-1}t$, then we see that the active $(k-1)$-block averages feels a drift towards the active k-block average. The active k-block does not move on time scale N^{k-1}, but it has already moved at time $\bar{t}$. At time $\bar{t}$ the active k-block has even reached its quasi-equilibrium, given by $\Gamma^{(k)}_{\bar{\vartheta}_k}$. Thus, the drift of the active $(k-1)$-block average is towards the instantaneous state of the active k-block average, which has distribution $\Gamma^{(k)}_{\bar{\vartheta}_k}$. This explains the first step in the interaction chain.

For $0 \leq l < k$, the active l-block average feels a drift towards the $(l+1)$-block average. The latter does not evolve on time scale $N^l t$, but it has already moved at time $\bar{t}$. Therefore it is no longer in its initial state, but in a quasi-equilibrium $\Gamma_u^{(l+2)}$, where u is the value of the active $(l+1)$-block averages determined via the interaction chain, recall Figure 4.9. This explains how the different space-time scales are connected via the active block averages. The states of the different seed-bank averages is a little bit more complicated. Below we give a very short heuristic explanation of the different seed-banks in the interaction chain.

For the effective process $(\boldsymbol{\Theta}^{\mathrm{eff},(l),\Omega_N}(t))_{t>0}$ instead of the full process $(\boldsymbol{\Theta}^{(l),\Omega_N}(t))_{t>0}$, we can consider in (4.113) only the active block average and the effective seed-bank average with $m=l$ (recall that the full block average equals the active block average). According to the above explanation, we have to replace $\Gamma^{(k)}_{\bar{\vartheta}_k}$ by $\Gamma^{\mathrm{eff},(k)}_{\vartheta_k}$ and $\Gamma^{(l)}_u$ by $\Gamma^{\mathrm{eff},(l)}_u$. Hence we find the effective interaction chain defined in (4.80) and depicted in Figure 4.8.

Slow seed-banks. From (4.124) we see that if $l \geq k$ and we use the scaling $\bar{t}$, then all l-blocks of seed-banks with colour $0 \leq m \leq k$ equal ϑ_k, and all l-blocks of seed-banks with colour $m > k$ equal their initial values θ_{y_m}. Something interesting happens when we choose $0 \leq l < k$ and use the scaling $\bar{t}+N^l t$. The single colonies of seed-banks with colour $0 \leq m < l$ on time scale $\bar{t}+N^l t$ follow the active population, and hence their l-block averages equal the l-block average of the active population.

Chapter 4

The l-block average of the seed-bank with colour l has a non-trivial interaction with the active l-block. The l-blocks of seed-banks with colour $m > k$ have not yet moved and hence are still in their initial states $(\theta_{y_m})_{m=k+1}^{\infty}$. However, (4.126) implies that the single colonies of the seed-banks with colour $k \geq m > l$ are not moving on time scale $\bar{t} + N^l t$, even though they had already moved at time $\bar{t}$. Therefore the l-blocks averages of seed-banks with colour $l < m \leq k$ are no longer in their initial state at time $\bar{t}$. Note that they are also not in the state ϑ_k, since this is the state of their k-block averages and not of their l-block averages. The single colony seed-banks with colour $l < m \leq k$ are in the state given by

$$y_{m,0}(\bar{t}\,) = \int_0^{\bar{t}} \mathrm{d}s \, \frac{e_m}{N^m} [x_0(s) - y_{m,0}(s)]. \tag{4.130}$$

Hence, for large N,

$$y_{m,0}(\bar{t}) \approx \int_0^{L(N)N^k + N^k t_k + \cdots + N^m t_m} \mathrm{d}s \, \frac{e_m}{N^m} [x_0(s) - y_{m,0}(s)]. \tag{4.131}$$

Similarly, for the l-block average with colour m we have

$$\begin{aligned} \Theta_{y_m}^{(l),\Omega_N}(\bar{t}\,) &= \int_0^{\bar{t}} \mathrm{d}s \, \frac{e_m}{N^m} [\Theta_x^{(l),\Omega_N}(s) - \Theta_{y_m}^{(l),\Omega_N}(s)] \\ &\approx \int_0^{L(N)N^k + N^k t_k + \cdots + N^m t_m} \mathrm{d}s \, \frac{e_m}{N^m} [\Theta_x^{(l),\Omega_N}(s) - \Theta_{y_m}^{(l),\Omega_N}(s)]. \end{aligned} \tag{4.132}$$

Thus, we see that the state of $\Theta_{y_m}^{(l),\Omega_N}(\bar{t})$ is completely determined at time $L(N)N^k + N^k t_k + \cdots + N^m t_m$, i.e., the last time before $\bar{t} + N^l t$ that the single colony seed-banks of colour m had an opportunity to move. Up to time $L(N)N^k + N^k t_k + \cdots + N^m t_m$, the single colony seed-banks with colour m interact at a very slow rate with the active single colonies, and similarly for the l-blocks. Therefore effectively the colour-m seed-bank interacts with a "time-average on scale $N^m t_m$" of the active population. On time scale $N^m t_m$, a single active colony migrates very fast in its $(m-1)$-block. As a consequence at time $\bar{t} + N^m t_m$ individuals that start from a particular colony, e.g. site 0, are spread uniformly over the m-block containing this site. Hence the interaction of a single m-dormant colony with the active population can be intuitively interpreted as an interaction with the active m-block, and similarly for an m-dormant l-block. Once we move to lower time scales, the m-dormant single colonies do not interact with the active colony anymore. In the detailed proofs we show that one consequence of this is that, for $l < m$,

$$\begin{aligned} \Theta_{y_m}^{(l),\Omega_N}(\bar{t} + N^l t) &\approx \Theta_{y_m}^{(l),\Omega_N}(L(N)N^k + N^k t_k + \cdots + N^m t_m) \\ &\approx \Theta_{y_m}^{(m),\Omega_N}(L(N)N^k + N^k t_k + \cdots + N^m t_m). \end{aligned} \tag{4.133}$$

Thus, the l-block averages of colours $l \leq m \leq k$ equal the state of the corresponding m-block. This is the $(m+2)$-th component of the interaction chain at level l.

Conclusion. Combining the intuitive descriptions in Sections 4.4.4-4.4.4, we see how Theorems 4.4.2 and Theorems 4.4.4 come about. Their proofs will rely on coupling techniques and a detailed analysis of the SSDEs. This analysis will be done in several steps. In Sections 6.1–6.3 we first deal with a simplified system, the mean-field model, for which we derive the McKean-Vlasov limit and the mean-field finite-systems scheme. In Sections 7–9 we extend our analysis to finitely many hierarchical levels. In particular, we go through the following list of systems of increasing complexity, each being a simplified version of the system defined by (4.19) and each capturing a key feature:

(a) Two-colour mean-field finite-systems scheme (Section 7.1).

(b) Two-level hierarchical mean-field system (Section 8.1).

(c) Finite-level mean-field system (Section 9.1).

In Section 9 we put the pieces together to prove the multi-scaling for the infinite-level system given in Theorems 4.4.2 and 4.4.4.

§4.5 Main results $N \to \infty$: Orbit and cluster formation

In the hierarchical mean-field limit we say that the system clusters when the colonies gradually form larger and larger mono-type blocks. In Section 4.5.1 we determine whether, in the hierarchical mean-field limit, the system clusters along successive space-time scales. How this happens is captured by the interaction chain. We introduce a sequence of scaling factors $(A_k)_{k\in\mathbb{N}_0}$, where A_k is defined in terms of the rates $(c_k)_{k\in\mathbb{N}_0}$, $(e_k)_{k\in\mathbb{N}_0}$, $(K_k)_{k\in\mathbb{N}_0}$ and the factors $(E_k)_{k\in\mathbb{N}_0}$. Using these scaling factors, we analyse the orbit of the renormalisation transformation and establish *universality*: $A_k\mathcal{F}^{(k)}g$ converges as $k \to \infty$ to the Fisher-Wright diffusion function, irrespective of the choice of g. In Section 4.5.2 we show how the scaling factors A_k are connected to the *growth of mono-type clusters*. In Section 4.5.3 we identify the asymptotics of A_k as $k \to \infty$ in terms of the model parameters.

§4.5.1 Orbit of renormalisation transformations

To determine whether clustering occurs, we start from larger and larger time scales and use the interaction chain to see whether mono-type clusters are formed in the single colonies. Recall the kernels introduced in (4.79) that describe the connection between subsequent hierarchical levels in the interaction chain. Define the following composition of kernels (see Fig. 4.10):

$$Q^{(n)} = Q^{[n]} \circ \cdots \circ Q^{[0]}, \qquad n \in \mathbb{N}. \tag{4.134}$$

In words, $Q^{(n)}(z_n, \mathrm{d}z_0)$ is the probability density to see the population of a single colony in state z_0 given that the $(n+1)$-block average equals z_n.

In Section 4.3 we identified the clustering regime for fixed $N < \infty$. In this section we identify the clustering regime in the hierarchical mean-field limit. In the clustering regime, in the hierarchical mean-field limit, an interesting question is to determine how $\mathcal{F}^{(n)}g$ (recall (4.76)) scales with n. We identify the scaling and show that it does *not* depend on g (see Fig. 4.11).

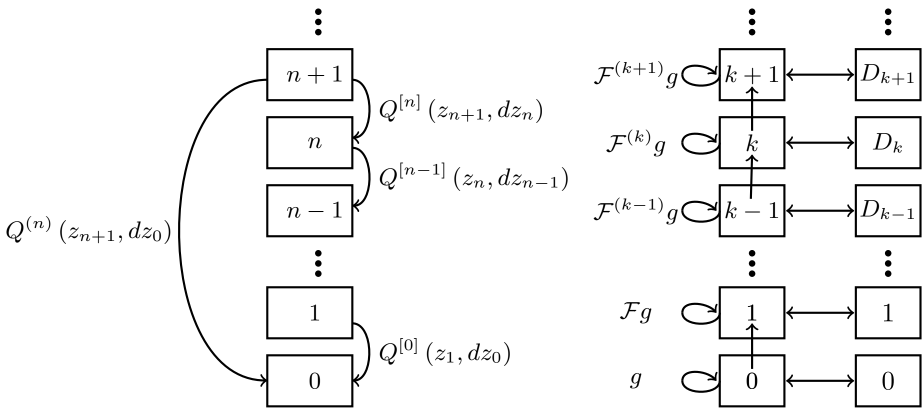

Figure 4.10: Left: *The interaction chain that connects successive hierarchical levels downwards. The arrows on the right correspond to* (4.79)*, the arrow on the left corresponds to* (4.134)*.* Right: *The renormalisation transformation that connects successive hierarchical levels upwards. The vertical arrows correspond to* (4.76)*. The horizontal arrows represent the interaction with the effective seed-bank. The arrows on the left represent the resampling driven by the renormalised diffusion function.*

To state the clustering result, abbreviate

$$\bar{\vartheta}^{(n)} = (\vartheta_n, \overbrace{\vartheta_n, \cdots, \vartheta_n}^{n+1 \text{ times}}, \theta_{y_{n+1}}, \theta_{y_{n+2}}, \cdots). \tag{4.135}$$

Theorem 4.5.1 (Renormalised scaling). *Let c_k be as in* (4.5)*, e_k and K_k as in* (4.10) *and E_k as in* (4.64)*. Define*

$$A_n = \frac{1}{2}\sum_{k=0}^{n-1} \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}, \qquad n \in \mathbb{N}. \tag{4.136}$$

Then

$$\lim_{n\to\infty} Q^{(n)}\big(\bar{\vartheta}^{(n)}, \cdot\big) = (1-\theta)\,\delta_{(0,0^{\mathbb{N}_0})} + \theta\,\delta_{(1,1^{\mathbb{N}_0})} \tag{4.137}$$

if and only if

$$\lim_{n\to\infty} A_n = \infty. \tag{4.138}$$

Moreover, if (4.138) *holds, then for all $g \in \mathcal{G}$,*

$$\lim_{n\to\infty} A_n \mathcal{F}^{(n)} g = g_{\mathrm{FW}} \quad \textit{pointwise}, \tag{4.139}$$

with $g_{\mathrm{FW}}(x) = x(1-x)$, $x \in [0,1]$.

The proof of Theorem 4.5.1 is given in Section 10. The scaling factors A_n can be interpreted as clustering coefficients: in Section 4.5.2 we will show that the faster the A_n grow to infinity, the faster we expect to see clusters grow. The property in (4.137) corresponds to the *clustering regime*. According to (4.139), even though A_n depends on the choice of the sequences $\underline{K}, \underline{e}, \underline{c}$ in (4.5) and (4.10), the limit $A_n \mathcal{F}^{(n)} g$ as $n \to \infty$ is *universal*: irrespective of the choice of $g \in \mathcal{G}$, the limit is the standard Fisher-Wright diffusion function g_{FW}. Thus, g_{FW} is the *global attractor of the renormalisation transformation* (see Fig. 4.11).

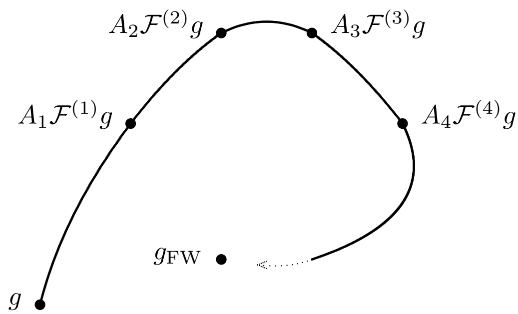

Figure 4.11: Flow of the iterates $\mathcal{F}^{(n)} g$, $n \in \mathbb{N}_0$, of the renormalisation transformation acting on the class $\mathcal{G}$. After multiplication by A_n, the flow is globally attracted by g_{FW}.

§4.5.2 Growth of mono-type clusters

In the *clustering regime* we are interested in how fast mono-type clusters grow in space over time. For the system on $\mathbb{Z}^d$ and Ω_N without seed-bank the growth rate has been studied in detail. Different growth rates were found for *strongly recurrent* and *critically recurrent* migration. Typical examples on Ω_N are migrations with coefficients $c_k = c^k$ with $c \in (0, 1)$, respectively, $c_k = C$ with $C \in (0, \infty)$. Typical examples on $\mathbb{Z}$, respectively, $\mathbb{Z}^2$ are migrations with zero mean and finite variance. For these models the following behaviour occurs.

- In the strongly recurrent case, mono-type clusters grow fast and cover a volume that increases at time t at a rate that is given by the Green function up to time t of the underlying random walk, times a certain random constant that can be determined explicitly and that is independent of the diffusion function $g \in \mathcal{G}$ [33], [51]. The cluster growth is monitored by considering families of balls growing in time at such a speed that, starting from a translation invariant and ergodic initial state, the mean of the configuration in the ball is still close to the starting mean but begins to move. *Fast clustering* means that the cluster covers multiples of a scale that eventually lies in every finite family of balls with the above property.

- In the critically recurrent case, the volume grows only moderately fast, like $N^{(1-U)t}$ as $t \to \infty$, with $U \in [0, 1]$ a random variable. In other words, the cluster sizes have random orders of magnitude, an effect known as *diffusive clustering*. For $c_k = C \in (0, \infty)$, $k \in \mathbb{N}_0$, the distribution of U can be identified by studying the fraction of active individuals of type 1 in a ball of size $N^{(1-u)t}$, which can be shown to converge to $V(\log \frac{1}{1-u})$ as $t \to \infty$ with $(V(s))_{s \geq 0}$ the standard Fisher-Wright diffusion, irrespective of the choice of $g \in \mathcal{G}$ [35], [36].

- For more general migration it is possible that mono-type clusters grow slower than any positive power of t as $t \to \infty$. This occurs for recurrent migration in which the Green function up to time t grows like $o(\log t)$. For this regime only few results are available [25].

From the perspective of explaining *universality* in $g \in \mathcal{G}$ in the hierarchical mean-field limit $N \to \infty$, the above type of behaviour has been studied in detail in [25] and [41] for the Fleming-Viot model, respectively, the Cannings model without seed-bank. The renormalisation analysis for the model with seed-bank allows us to study how the seed-bank affects the cluster growth. In what follows we give a sketch of *three regimes of cluster growth.*

Types of clustering. If a ball in Ω_N lies in a mono-type cluster, then the block average of the active and the dormant components in this ball are all close to either 0 or 1. We can therefore analyse the growth rate of mono-type clusters by analysing at which hierarchical level block averages hit 0 or 1 in the limit as $N \to \infty$. To that end, we look at the interaction chain $M^k_{-l(k)}$ for $k \to \infty$, where the *level scaling function* $l\colon \mathbb{N}_0 \to \mathbb{N}_0$ is non-decreasing with $\lim_{k\to\infty} l(k) = \infty$ and is suitably chosen such that we obtain a *non-trivial clustering limiting law*, i.e.,

$$\lim_{k\to\infty} \mathcal{L}\big[M^k_{-l(k)}\big] = \mathcal{L}\big[\hat{\theta}\,\big], \tag{4.140}$$

where the limiting sequence of random frequencies $\hat{\theta}$ satisfies

$$0 < \mathbb{P}\big(\hat{\theta} \in \{0^{\mathbb{N}_0}, 1^{\mathbb{N}_0}\}\big) < 1. \tag{4.141}$$

In line with [25] and [22], in order to analyse the growth of mono-type clusters on multiple space-time scales in the hierarchical mean-field limit, it is natural to consider a family of non-decreasing functions $l_\chi\colon \mathbb{N}_0 \to \mathbb{N}_0$, $\chi \in I \subseteq [0,\infty)$, called the *cluster scales*, satisfying (4.140)–(4.141):

(1) **F**ast clustering: $\lim_{k\to\infty} l_\chi(k)/k = 1$ for all $\chi \in I$.

(2) **D**iffusive clustering: $\lim_{k\to\infty} l_\chi(k)/k = \kappa(\chi)$ for all $\chi \in [0,1]$, where $\chi \mapsto \kappa(\chi)$ is continuous and non-increasing with $\kappa(0) = 1$ and $\kappa(1) = 0$.

(3) **S**low clustering: $\lim_{k\to\infty} l_\chi(k)/k = 0$ for all $\chi \in I$. (This regime borders with the regime of coexistence.)

We write $(M^\infty_\chi)_{\chi\in I}$ with $M^\infty_\chi = \lim_{k\to\infty} M^k_{-l_\chi(k)}$ to denote the *cluster process.*

Remark 4.5.2. Examples are:

(1) $I = \mathbb{N}_0$, $l_\chi(k) = k - \chi$.

(2) $I = [0,1]$, $l_\chi(k) = \lfloor (1-\chi)k \rfloor$.

(3) $I = [0,1]$, $l_\chi(k) = \lfloor L(k^{1-\chi}) \rfloor$ with $L(0) = 0$, L non-decreasing and sublinear.

In words, the clusters cover blocks of level: (1) $k-\chi$; (2) $\lfloor(1-\chi)k\rfloor$; (3) $L(k^{1-\chi})$. For the model without seed-bank and with migration coefficients $c_k = c^k$ with $c \in (0,1)$, case (1) is realised with a Markov chain $(M_l^\infty)_{l\in\mathbb{N}_0}$ as scaling limit, while for $c_k = C$, case (2) is realised with a time-transformed Fisher-Wright diffusion in χ as scaling limit. (For finite N, this corresponds to the first and the second example given in the first paragraph of this section.) For the model without seed-bank, these scales have been shown to satisfy the required conditions. Case (3) also appears for the model without seed-bank, but detailed information on scales and scaling limits is lacking. As we will see below, seed-banks can slow down cluster growth, so case (3) is worthwhile to be studied in more detail. ■

Recall (4.136). Fast clustering corresponds to $A_k \gg k$, diffusive clustering to $A_k \asymp k$, and slow clustering to $A_k \ll k$ for large k. Theorem 4.5.3 below shows that, subject to (4.52) and (4.53), all three regimes are possible for the model with seed-bank. The regimes are the same as for the model without seed-bank when $\rho < \infty$, but different when $\rho = \infty$.

For systems without seed-bank, examples of the three types of clustering can be found in the literature: diffusive clustering in [2], [16] (voter model on $\mathbb{Z}$, respectively, $\mathbb{Z}^2$) and in [20], [35], [25], [51] (interacting Fleming-Viot processes on Ω_N with $N < \infty$, respectively, $N \to \infty$), all types of clustering in [22], [52] , [71] (interacting Feller diffusions on Ω_N with $N \to \infty$) and in [41], [42] (interacting Cannings processes on Ω_N in non-random and random environment with $N \to \infty$).

For the model with seed-bank we have to use the asymptotics of $(A_k)_{k\in\mathbb{N}}$ to identify the set I and the family $(l_\chi(\cdot))_{\chi\in I}$, and show that $(M^k_{-l_\chi(k)})_{\chi\in I}$ converges as $k \to \infty$ to a Markov process, which we want to identify.

Computations. In the following we demonstrate how we can carry out the above task. The key idea is to study first and second moments of the interaction chain, as well as sums of variances in order to get a handle on the quadratic variation process. To that end we calculate

$$V^k_{-l} = \text{ the conditional variance of the active part of } M^k_{-l} \text{ given } M^k_{-(l+1)} \tag{4.142}$$

and consider the sum of random variables $A_{k,n} = \sum_{-(k+1)\leq -l\leq -n} V^k_{-l}$, $n \in \mathbb{N}_0$. In order for the system to cluster, we must have $\lim_{k\to\infty} A_{k,n} = 0$ for every $n \in \mathbb{N}_0$. The *volatility profile* is given by

$$(p_\chi(k))_{\chi\in[0,1]}, \qquad p_\chi(k) = A_{k,l_\chi(k)}/A_{k,0}. \tag{4.143}$$

This profile is a *random variable* that depends on the interaction chain up to $M^k_{-l_\chi(k)}$. Since $A_{k,l_\chi(k)} = A_{k,0} - A_{l_\chi(k)-1,0}$, we have $p_\chi(k) = (A_{k,0} - A_{l_\chi(k)-1,0})/A_{k,0}$. For diffusive clustering, for instance, we want to show that

$$\lim_{k\to\infty} p_\chi(k) = 1 - \kappa(\chi), \qquad \chi \in [0,1], \tag{4.144}$$

while for fast clustering the limit is 0 and for slow clustering the limit is 1. From (4.139) we know that the scaled renormalised diffusion function $A_n\mathcal{F}^{(n)}(g)$ tends to the standard Fisher-Wright diffusion function as $n \to \infty$. Since the latter hits the boundary $\{0,1\}$ after some finite time, the coefficients A_n describe the speed at which the interaction chain hits this boundary. We next make this idea precise and show how it can be used to obtain information about the growth of mono-type clusters.

The kernels defined in Section 4.5.1 allow us to compute the first and second moments of all the block averages, which will be done in Section 10.1 (Propositions 10.1.1–10.2.2). In particular, using the interaction chain starting between at $-n$ and running until $-m$ with $-n < -m \leq 0$, and considering the m-block averages on time scale $N^m t$ in the limit $N \to \infty$, we find that the variance of the active component x_m^n of M_{-m}^n equals

$$\mathbb{V}\mathrm{ar}(x_m^n) = \mathbb{E}\left[(x_m^n - \vartheta_n)^2\right] = A_m^n (\mathcal{F}^{(n+1)} g)(\vartheta_n), \tag{4.145}$$

where

$$A_m^n = \frac{1}{2} \sum_{k=m}^{n} \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}. \tag{4.146}$$

(Note that $A_n = A_0^{n-1}$.) On the other hand, since $x_m^n \in (0,1)$ we have $\mathbb{V}\mathrm{ar}(x_m^n) \in (0,1)$ and $A_m^n(\mathcal{F}^{(n+1)})(\vartheta_n) \in (0,1)$. Taking $m=0$, we get $(\mathcal{F}^{(n+1)} g)(\vartheta_n) \in (0, \frac{1}{A_0^n})$. This implies that

$$(\mathcal{F}^{(n+1)} g)(\vartheta_n) = \int_{[0,1]^2} (\mathcal{F}^m g)(x_m)\, Q_m^{(n)}((\vartheta_n, \theta_{y,n}), \mathrm{d}z_m) \in \left(0, \tfrac{1}{A_0^n}\right). \tag{4.147}$$

Since $\lim_{n\to\infty} A_n(\mathcal{F}^{(n+1)} g) = g_{\mathrm{FW}}$, for large enough m,n we can approximate $\mathcal{F}^{(m)} g \approx g_{\mathrm{FW}}/A_0^m$. Therefore

$$\int \frac{g_{\mathrm{FW}}}{A_0^m}(x_m)\, Q_m^{(n)}((\vartheta_n, \theta_{y,n}), \mathrm{d}z_m) \in \left(0, \tfrac{1}{A_0^n}\right), \tag{4.148}$$

or, equivalently,

$$\int_{[0,1]^2} x_m(1-x_m)\, Q_m^{(n)}((\vartheta_n, \theta_{y,n}), \mathrm{d}z_m) \in \left(0, \tfrac{A_0^m}{A_0^n}\right). \tag{4.149}$$

Hence, if $A_0^m/A_0^n < \epsilon$ with $\epsilon > 0$ small, then we know that with high probability the system on time scale n has clusters with a radius of size m. (Note that for the interaction chain this means that the variance is almost entirely centred between n and m.) Therefore the speed at which A_0^m/A_0^n converges to zero as $m,n \to \infty$ says something about the speed at which monotype clusters form.

To capture the cluster growth, we must decide how we let $m,n \to \infty$. For this we look for clusters of radius $l_\chi(n)$ with $\chi \in I$. Put

$$f^n(l_\chi(n)) = \frac{A_0^{l_\chi(n)}}{A_0^n}, \tag{4.150}$$

and define, for $\epsilon > 0$,

$$\mathcal{X}_\epsilon^n = \inf\{\chi \in I \colon\ f^n(l_\chi(n)) < \epsilon\}. \tag{4.151}$$

Then the three types of clustering correspond to:

(1) **F**ast clustering: $\lim_{n\to\infty} l_{\mathcal{X}_\epsilon^n}(n)/n = 1$.

(2) **D**iffusive clustering: $\lim_{n\to\infty} l_{\mathcal{X}_\epsilon^n}(n)/n = R$ for some random variable R taking values in $(0,1)$.

(3) **S**low clustering: $\lim_{n\to\infty} l_{\mathcal{X}_\epsilon^n}(n)/n = 0$.

In terms of the interaction chain starting from $-k$ with $k \to \infty$, in view of (4.145) this corresponds to the variance in the interaction chain being concentrated near the beginning, being spread out or being concentrated near the end.

§4.5.3 Rates of scaling for renormalised diffusion function

For the system without seed-bank, we have $K_k = e_k = 0$ and $E_k = 1$ for all $k \in \mathbb{N}_0$. Hence

$$A_n = \frac{1}{2}\sum_{k=0}^{n-1}\frac{1}{c_k} \tag{4.152}$$

and (4.137) holds if and only if

$$\sum_{k\in\mathbb{N}_0}\frac{1}{c_k} = \infty. \tag{4.153}$$

Various subcases were analysed in [5]. For the system with seed-bank because $E_0 = 1$ and $E_k < 1$ (see (4.64)), it follows from (4.136) that

$$A_n < \frac{1}{2}\sum_{k=0}^{n-1}\frac{1}{c_k}. \tag{4.154}$$

Thus we see that the seed-bank *weakens clustering*, i.e., enhances genetic diversity, even in the hierarchical mean-field limit.

We identify the clustering regime in the setting where the coefficients are *asymptotically polynomial*, as in (4.52), or are *pure exponential*, as in (4.53). It turns out that there is a delicate interplay between the migration and the seed-bank, resulting in 4 different scalings for asymptotically polynomial coefficients and 8 different scalings for pure exponential coefficients.

Theorem 4.5.3 (Rates of scaling for diffusion function). *Let ρ be as defined in* (4.14).

(I) *If $\rho < \infty$, then* (4.138) *holds if and only* (4.153) *hold, and*

$$A_n \sim \frac{1}{2(1+\rho)}\sum_{m=0}^{k-1}\frac{1}{c_k}. \tag{4.155}$$

(II) *If $\rho = \infty$, then* (4.138) *holds in the following cases:*

- *Subject to* (4.52) *if and only if* $-\phi \leq \alpha \leq 1$, *with*

$$\begin{aligned} -\phi < \alpha < 1&: \quad A_n \sim C_1\, n^{\alpha+\phi},\\ -\phi = \alpha < 1&: \quad A_n \sim C_2\, \log n,\\ -\phi < \alpha = 1&: \quad A_n \sim C_3\, \frac{n^{1+\phi}}{\log n},\\ -\phi = \alpha = 1&: \quad A_n \sim C_4\, \log\log n, \end{aligned} \tag{4.156}$$

where

$$C_1 = \tfrac{1}{2AF}\tfrac{1-\alpha}{\alpha+\phi}, \quad C_2 = \tfrac{1}{2AF}(1-\alpha), \quad C_3 = \tfrac{1}{2AF}\tfrac{1}{1+\phi}, \quad C_4 = \tfrac{1}{2AF}. \tag{4.157}$$

The values of B, β play no role for the clustering, nor for the asymptotics.

- *Subject to* (4.53) *if and only if* $Kc \leq 1 \leq K$, *with*

$$\begin{aligned} c < Ke,\ Kc < 1&: \quad A_n \sim \hat{C}_1\,(Kc)^{-(n-1)},\\ c < Ke,\ Kc = 1&: \quad A_n \sim \bar{C}_1\, n,\\ c = Ke,\ Kc < 1&: \quad A_n \sim \hat{C}_2\,(Kc)^{-(n-1)},\\ c = Ke,\ Kc = 1&: \quad A_n \sim \bar{C}_2\, n,\\ c > Ke,\ Kc < 1&: \quad A_n \sim \hat{C}_3\,(Kc)^{-(n-1)},\\ c > Ke,\ Kc = 1&: \quad A_n \sim \bar{C}_3\, n,\\ c < 1 = K&: \quad A_n \sim \tilde{C}_1\, n^{-1}\, c^{-(n-1)},\\ c = 1 = K&: \quad A_n \sim \tilde{C}_2 \log n, \end{aligned} \tag{4.158}$$

where

$$\begin{aligned} &\hat{C}_1 = \tfrac{K-1}{2K(1-Kc)}, \quad &&\hat{C}_2 = \tfrac{(K-1)^2}{2(2K-1)(1-Kc)}, \quad &&\hat{C}_3 = \tfrac{K-1}{2(1-Kc)},\\ &\bar{C}_1 = \tfrac{K-1}{2K}, &&\bar{C}_2 = \tfrac{(K-1)^2}{2(2K-1)}, &&\bar{C}_3 = \tfrac{K-1}{2},\\ &\tilde{C}_1 = \tfrac{1}{2(1-c)}, &&\tilde{C}_2 = \tfrac{1}{2}. \end{aligned} \tag{4.159}$$

The value of e plays no role for the clustering, but does for the asymptotics.

The proof of Theorem 4.5.3 is given in Section 10. Part (I) shows that for $\rho < \infty$ the clustering regime is the same as for the system without seed-bank. The scaling of A_n is controlled by the migration and is reduced by a factor $1/(1+\rho)$ with respect to the seed-bank. Part (II) shows that for $\rho = \infty$ the clustering regime is different from that for the system without seed-bank. Clustering is harder to achieve: since $\lim_{k\to\infty} E_k = 0$ the growth rate of A_n is *strictly smaller* than without seed-bank.

Furthermore, subject to (4.52), if $-\phi < \alpha < 1$, then the growth rate of A_n drops down from $\asymp n^{1+\phi}$ without seed-bank to $\asymp n^{\alpha+\phi}$ with seed-bank, while if $-\phi = \alpha = 1$, then it drops down from $\asymp \log n$ to $\asymp \log\log n$. Similarly, subject to (4.53), if $Kc < 1 < K$, then the growth rate of A_n drops down from $\asymp c^{-n}$ to $\asymp (Kc)^{-n}$, while if $c = K = 1$, then it drops down from $\asymp n$ to $\asymp \log n$.

Returning to the observations made in Section 4.5.2, we see that the three clustering regimes also appear in the model with seed-bank, both for $\rho < \infty$ and $\rho = \infty$, and in the latter case are accompanied by different migration coefficients. The scaling results mentioned in Section 4.5.2 can in principle be deduced from the asymptotics of A_n as $n \to \infty$ in Theorem 4.5.3. It would be interesting to work out the details and to identify the limiting processes that control the cluster growth.

CHAPTER 5

Proofs long-time behaviour $N < \infty$

In this chapter we prove Theorems 4.3.2–4.3.3. The integral criterion for $\rho = \infty$ in (4.50) is explained in Section 5.1. Theorem 4.3.2 is proved in Section 5.2 and Theorem 4.3.3 in Section 5.3.

§5.1 Explanation of clustering criterion for infinite seed-bank

Recall Fig. 4.6. Suppose that $g = dg_{\mathrm{FW}}$, so that we have a dual. We will show that the integral criterion in (4.50) determines whether or not two dual lineages coalesce with probability 1. Since two lineages in the dual can only coalesce when they are active at the same site, we need to keep track of the probabilities that the lineages are active at a given time. Because the lineages can only migrate when they are active, we also need to keep track of the total time they are active up to a given time.

Recall the renewal interpretation of the dual process (see Remark 4.2.9). We argue heuristically as follows. If $\rho = \infty$, then the activity times σ_k are much smaller than the sleeping times τ_k, and we may assume that $\tau_k + \sigma_k \asymp \tau_k$, $k \to \infty$. Discretising time, we can use the results from [1] for the intersection of two independent renewal processes. Then the integral criterion in (4.50) can be interpreted as follows:

- If $\gamma \in (0,1)$, then the probability for each of the lineages to be active at time s decays like $\asymp \varphi(s)^{-1} s^{-(1-\gamma)}$ [1]. Hence the total time they are active up to time s is $\asymp \varphi(s)^{-1} s^{\gamma}$. Because the lineages only move when they are active, the probability that the two lineages meet at time s is $\asymp a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. Hence the total hazard is $\asymp \int_1^\infty \mathrm{d}s\, [\varphi(s)^{-1} s^{-(1-\gamma)}]^2\, a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. After the transformation $t = t(s) = \varphi(s)^{-1} s^{\gamma}$, the latter turns into the integral in (4.50), modulo a constant. When carrying out this transformation, we need that $s\varphi'(s)/\varphi(s) \to 0$, which follows from (4.49), and $\varphi(t(s))/\varphi(s) \asymp 1$, which follows from the bound we imposed on ψ in (4.49) together with the fact that $\log \varphi(s)/\log s \to 0$. This computation is spelled out in Appendix B.1.

- If $\gamma = 1$, then the probability for each of the lineages to be active at time s decays like $\hat{\varphi}(s)^{-1}$ [1], and so the total time they are active up to time s is $\asymp s\hat{\varphi}(s)^{-1}$. Recall from (4.48) that $\hat{\varphi}(t) = \mathbb{E}[\tau \wedge t]$ is also slowly varying.) Hence the total hazard is $\asymp \int_1^\infty \mathrm{d}s\, [\hat{\varphi}(s)^{-1}]^2\, a^{(N)}_{\hat{\varphi}(s)^{-1}s}(0,0)$. After the transformation

$t = t(s) = \hat{\varphi}(s)^{-1}s$ (for which we can use the same type of computation as in Appendix B.1), the latter turns into the integral in (4.50), modulo a constant.

§5.2 Scaling of wake-up time and migration kernel for infinite seed-bank

We can prove Theorem 4.3.2 by direct computation via assumptions (4.52)–(4.53). We start by computing γ. Afterwards we compute $\hat{\varphi}(t)$ and $a_t^{\Omega_N}(0,0)$.

Computation of γ. Recall (4.40), which reads

$$\mathbb{P}(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m}\, \mathrm{e}^{-(e_m/N^m)t}. \tag{5.1}$$

Since we are interested in the asymptotic behaviour of $\mathbb{P}(\tau > t)$ as $t \to \infty$, we need to consider only large values of t. For large values of t, only large values of m (for which $\frac{e_m}{N^m}$ is small) contribute to the sum in (5.1). Hence we can estimate the latter by an integral and insert the assumptions made in (4.52)–(4.53). Subsequently, using the change of variable $s = \frac{e_m}{N^m}$ and taking the logarithm to express m in terms of s, we obtain the following values of γ after extracting the t-dependence:

$$\begin{aligned}
&(4.52) \quad \Longrightarrow \quad \gamma = 1, \quad \varphi(t) \asymp (\log t)^{-\alpha},\\
&(4.53) \quad \Longrightarrow \quad \gamma = \gamma_{N,K,e} = \frac{\log(N/Ke)}{\log(N/e)}, \quad \varphi(t) \asymp 1.
\end{aligned} \tag{5.2}$$

In order to guarantee that $\rho = \infty$, we must require that $\alpha \in (-\infty, 1]$, respectively, $K \in [1, \infty)$ (while β, respectively, e play no role). Subject to (4.52),

$$\hat{\varphi}(t) \asymp \begin{cases} (\log t)^{1-\alpha}, & \alpha \in (-\infty, 1),\\ \log\log t, & \alpha = 1, \end{cases} \tag{5.3}$$

while subject to (4.53),

$$\hat{\varphi}(t) \asymp \begin{cases} 1, & K \in (1, \infty),\\ \log t, & K = 1. \end{cases} \tag{5.4}$$

Computation of $a_t^{\Omega_N}(0,0)$. To compute $a_t^{\Omega_N}(0,0)$, we first rewrite the migration kernel $a^{\Omega_N}(\cdot,\cdot)$ in (4.6) as

$$a^{\Omega_N}(0,\eta) = \frac{r_{\|\eta\|}}{N^{\|\eta\|-1}(N-1)} \tag{5.5}$$

with

$$r_{\|\eta\|} = \frac{1}{D(N)} \frac{N-1}{N} \sum_{l \geq \|\eta\|} \frac{c_{l-1}}{N^{l-1}} \frac{1}{N^{l-\|\eta\|}}, \tag{5.6}$$

where $D(N)$ is a renormalisation constant such that $\sum_{j\in\mathbb{N}} r_j = 1$. For transition kernels of the form (5.5), the time-t transition kernel $a_t^{\Omega_N}(\cdot,\cdot)$ was computed in [35] with the help of Fourier analysis, see also [19]. Namely,

$$a_t^{\Omega_N}(0,\eta) = \sum_{j\geq k} K_{jk}(N)\,\frac{\exp[-h_j(N)t]}{N^j}, \qquad t\geq 0, \quad \eta\in\Omega_N\colon\ d_{\Omega_N}(0,\eta)=k\in\mathbb{N}_0, \tag{5.7}$$

where

$$K_{jk}(N) = \begin{cases} 0, & j=k=0,\\ -1, & j=k>0,\\ N-1, & \text{otherwise}, \end{cases} \qquad j,k\in\mathbb{N}_0, \tag{5.8}$$

and

$$h_j(N) = \frac{N}{N-1}\,r_j(N) + \sum_{i>j} r_i(N), \qquad j\in\mathbb{N}. \tag{5.9}$$

The expressions in (5.6)–(5.9) simplify considerably in the limit as $N\to\infty$, namely, the term with $i=j$ dominates and

$$h_j(N)\sim r_j(N)\sim \frac{c_{j-1}}{D(N)N^{j-1}}, \quad j\in\mathbb{N}, \qquad D(N)\sim c_0. \tag{5.10}$$

We show why this is true for $h_j(N)$ (the argument for $r_j(N)$ and $D(N)$ is similar). Write

$$\begin{aligned} h_j(N) &= \frac{N}{N-1}\,r_j(N) + \sum_{i>j} r_i(N)\\ &= \frac{1}{D(N)}\left(\sum_{l\geq j}\frac{c_{l-1}}{N^{l-1}}\frac{1}{N^{l-j}} + \frac{N-1}{N}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}}\sum_{i<j\leq l}\frac{1}{N^{l-i}}\right)\\ &= \frac{1}{D(N)}\frac{c_{j-1}}{N^{j-1}}\left(1+\left[1+O\left(\frac{1}{N}\right)\right]\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}}\right). \end{aligned} \tag{5.11}$$

Hence it suffices to show that

$$\limsup_{N\to\infty}\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}} = 0, \qquad j\in\mathbb{N}. \tag{5.12}$$

To do so, note that, since $\limsup_{k\to\infty}\frac{1}{k}\log c_k < \log N$ by (4.7), for N large enough we have

$$\sup_{k\in\mathbb{N}_0} c_k^{1/k} < N. \tag{5.13}$$

Let $\bar{N} = \inf\{N\in\mathbb{N}\colon\ \sup_{k\in\mathbb{N}_0} c_k^{1/k} < N\}$. Then

$$\begin{aligned} &\limsup_{N\to\infty}\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}} \leq \limsup_{N\to\infty}\frac{1}{c_{j-1}}\sum_{l>j}\frac{\bar{N}^{l-1}}{N^{l-1}}N^{j-1}\\ &= \frac{\bar{N}^{j-1}}{c_{j-1}}\limsup_{N\to\infty}\frac{\frac{\bar{N}}{N}}{1-\frac{\bar{N}}{N}} = 0, \qquad j\in\mathbb{N}, \end{aligned} \tag{5.14}$$

which settles (5.12).

To understand what (5.9) gives for finite N, note that for asymptotically polynomial coefficients (recall (4.52))

$$\begin{aligned}
\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} &= [1+o(1)]\, \frac{N^{j-1}}{F(j-1)^{-\phi}} \sum_{l>j} \frac{F(l-1)^{-\phi}}{N^{l-1}} \\
&= [1+o(1)] \sum_{l>j} \frac{(l-1)^{-\phi}}{(j-1)^{-\phi}} N^{-(l-j)} \qquad (5.15)\\
&= [1+o(1)] \sum_{k\geq 1} \left(1+\frac{k}{j-1}\right)^{-\phi} N^{-k}, \qquad j\in\mathbb{N}.
\end{aligned}$$

For $\phi \geq 0$ the right-hand side is bounded from above by $\sum_{k\geq 1} N^{-k} = \frac{1}{N-1}$ and for $\phi < 0$ by $N^{-1}\sum_{k\geq 1}(1+k)^{-\phi}N^{-(k-1)} \leq N^{-1}C_\phi$. On the other hand, for pure exponential coefficients (recall (4.53)),

$$\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} = \sum_{k\geq 1} \left(\frac{c}{N}\right)^{-k} = \frac{c}{N-c}. \qquad (5.16)$$

Hence, for both choices of coefficients we have the following:

> For $N \to \infty$ the quantities $h_j(N), r_j(N)$ are bounded from above and below by positive finite constants times the right-hand side of (5.10) uniformly in $j \in \mathbb{N}$. (5.17)

Picking $\eta = 0$ ($k = 0$) in (5.7), we obtain

$$a_t^{\Omega_N}(0,0) = \sum_{j\in\mathbb{N}} (N-1) \frac{\exp[-h_j(N)t]}{N^j}. \qquad (5.18)$$

Since we are interested in the asymptotic behaviour of $a_t^{\Omega_N}(0,0)$, only large values of j are relevant and we can estimate the sum in (5.18) by an integral. To do so, we change variables by putting $s = h_j(N)$ and exploit (5.17). Take the logarithm to express j in terms of s, compute $\mathrm{d}s/\mathrm{d}j$, and extract the t-dependence. This gives

$$\begin{aligned}
(4.52) \quad &\Longrightarrow \quad a_t^{\Omega_N}(0,0) \asymp t^{-1}\log^{\phi} t,\\
(4.53) \quad &\Longrightarrow \quad a_t^{\Omega_N}(0,0) \asymp t^{-1-\delta_{N,c}},
\end{aligned} \qquad (5.19)$$

where

$$\delta_{N,c} = \frac{\log c}{\log(N/c)}. \qquad (5.20)$$

§5.3 Hierarchical clustering

In this section we prove Theorem 4.3.3 by substituting the results of Theorem 4.3.2 into the clustering criterion in (4.50).

Combining (4.51), (5.2)–(5.4) and (5.19)–(5.20), we find the following clustering criterion for *fixed* N and infinite seed-bank:

- Subject to (4.52), clustering prevails if and only if

$$-\phi \leq \alpha \leq 1. \tag{5.21}$$

- Subject to (4.53), clustering prevails if and only

$$\delta_{N,c} \leq -\frac{1-\gamma_{N,K,e}}{\gamma_{N,K,e}}. \tag{5.22}$$

In view of (5.2) and (4.56), the condition in (5.22) amounts to

$$\log N \times \log(Kc) \leq \log c \times \log(K^2 e), \tag{5.23}$$

where we use that $c < N$ and $Ke < N$ (recall (4.7) and (4.12)). The condition in (5.23) holds for all N when

$$Kc = 1 \text{ with } \begin{cases} c = 1, & K^2 e \in (0, \infty), \\ c > 1, & K^2 e \geq 1, \\ c < 1, & K^2 e \leq 1. \end{cases} \tag{5.24}$$

It also holds for N large enough when $Kc < 1$ and fails for N large enough when $Kc > 1$. Thus, for infinite seed-bank, clustering prevails for N large enough if and only if

$$Kc \leq 1 \leq K, \tag{5.25}$$

which is the analogue of (5.21).

CHAPTER 6

Mean-field system

§6.1 Preparation: $N \to \infty$, McKean-Vlasov process and mean-field system

To analyse the scaling of our hierarchical system in the hierarchichal mean-field limit $N \to \infty$, we first need to understand simpler systems. In this section we consider the mean-field system consisting of a *single hierarchy*, and introduce the following:

(a) McKean-Vlasov process (Section 6.1.1).

(b) Mean-field system and McKean-Vlasov limit (Section 6.1.2).

For each we derive a key proposition that will play a crucial role in our analysis of the truncated system with *finitely many hierarchies* in Sections 7–9 and the full system with *infinitely many hierarchies* in Section 9. The proofs of the propositions stated in this section will be given in Sections 6.1.3 and 6.1.4.

§6.1.1 McKean-Vlasov process

In this section we introduce the McKean-Vlasov process, which will play an important role in our analysis of the mean-field system to be introduced in Sections 6.1.2–6.2.1. (In the full system the effective process introduced in (4.68) will be seen to be an example of a McKean-Vlasov process.)

For $g \in \mathcal{G}$ and $c, K, e \in (0, \infty)$, consider the single-colony process

$$z(t) = (x(t), y(t))_{t \geq 0}, \tag{6.1}$$

taking values in $[0,1]^2$, with initial law $\mathcal{L}[(x(0), y(0))] = \mu$ and with components evolving according to

$$\begin{aligned} \mathrm{d}x(t) &= c\,[\mathbb{E}[x(t)] - x(t)]\,\mathrm{d}t + \sqrt{g(x(t))}\,\mathrm{d}w(t) + Ke\,[y(t) - x(t)]\,\mathrm{d}t, \\ \mathrm{d}y(t) &= e\,[x(t) - y(t)]\,\mathrm{d}t, \end{aligned} \tag{6.2}$$

where $\mathbb{E}$ denotes expectation with respect to μ. With the help of Itô-calculus we can compute the expectation $\mathbb{E}[x(t)]$. Indeed, from (6.2) we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[x(t)] &= Ke\,\big[\mathbb{E}[y(t)] - \mathbb{E}[x(t)]\big], \\ \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[y(t)] &= e\,\big[\mathbb{E}[x(t)] - \mathbb{E}[y(t)]\big]. \end{aligned} \tag{6.3}$$

Define

$$\theta_x = \mathbb{E}^\mu[x(0)], \qquad \theta_y = \mathbb{E}^\mu[y(0)], \qquad \theta = \mathbb{E}^\mu\left[\frac{x(0)+Ky(0)}{1+K}\right]. \tag{6.4}$$

Note that (6.3) implies that θ is a preserved quantity, i.e.,

$$\mathbb{E}^\mu\left[\frac{x(0)+Ky(0)}{1+K}\right] = \mathbb{E}^\mu\left[\frac{x(t)+Ky(t)}{1+K}\right] = \theta, \qquad t \geq 0. \tag{6.5}$$

Solving (6.3), we find

$$\begin{aligned} \mathbb{E}[x(t)] &= \theta + \frac{K}{1+K}(\theta_x - \theta_y)\,\mathrm{e}^{-(K+1)et}, \\ \mathbb{E}[y(t)] &= \theta - \frac{1}{1+K}(\theta_x - \theta_y)\,\mathrm{e}^{-(K+1)et}. \end{aligned} \tag{6.6}$$

In particular, from (4.111) we see that

$$\lim_{t\to\infty}(\mathbb{E}[x(t)], \mathbb{E}[y(t)]) = (\theta, \theta). \tag{6.7}$$

Hence, in equilibrium we can replace $\mathbb{E}[x(t)]$ in (6.2) by θ. After inserting (6.6) into (6.2), we can use [72, Theorem 1, Remark on p.156] to show that for every deterministic initial state $(x(0), y(0)) \in [0,1]^2$ the SSDE in (6.2) has a unique strong solution. We will refer to this solution as the *McKean-Vlasov process.*

Remark 6.1.1 (Self-consistency). To prove uniqueness of the solution to (6.2) we can also use [38], where self-consistent mean-field dynamics are treated in detail. The solution has the Feller property. ■

Proposition 6.1.2 (McKean-Vlasov process: equilibrium). *For every initial law* $\mu \in \mathcal{P}([0,1]^2)$ *satisfying*

$$\mathbb{E}^\mu\left[\frac{x(0)+Ky(0)}{1+K}\right] = \theta, \qquad \theta \in [0,1], \tag{6.8}$$

the process in (6.1) *converges to a unique equilibrium,*

$$\lim_{t\to\infty} \mathcal{L}[(x(t), y(t))] = \Gamma_\theta, \tag{6.9}$$

and

$$\Gamma_\theta \in \mathcal{P}([0,1]^2), \tag{6.10}$$

satisfies

$$\theta = \int_{[0,1]^2} x\,\Gamma_\theta(\mathrm{d}x, \mathrm{d}y) = \int_{[0,1]^2} y\,\Gamma_\theta(\mathrm{d}x, \mathrm{d}y). \tag{6.11}$$

The proof of Proposition 6.1.2 is given in Section 6.1.3. Note that $\Gamma_\theta = \Gamma_\theta^{g,c,K,e}$ depends on all the parameters appearing in (6.2). In Section 6.2 we will see that Γ_θ is continuous as a function of θ.

Remark 6.1.3 (Non-linear Markov process). Note that (6.1) is a *non-linear* Markov process: the evolution not only depends on the current state $z(t)$, but also on the current law $\mathcal{L}[z(t)]$ via the expectation $\mathbb{E}[x(t)]$ appearing in the SSDE (6.2). This is different from the model without seed-bank, where the non-linearity is replaced by a drift towards θ that is constant in time. In equilibrium we can replace $\mathbb{E}[x(t)]$ by θ in (6.2), but before equilibrium is reached we cannot, because $t \mapsto \mathbb{E}[x(t)]$ is not constant, as is clear from (4.111). Note that $\mathbb{E}[x(t)]$ is a linear functional of $z(0)$. This fact will play an important role in the renormalisation analysis in Section 10. ■

§6.1.2 Mean-field system and McKean-Vlasov limit

In this section we consider a simplified version of the SSDE in (4.20), namely, we restrict to the finite geographic space

$$[N] = \{0, 1, \dots, N-1\}, \qquad N \in \mathbb{N}. \tag{6.12}$$

In this simplified version, the migration kernel $a^{\Omega_N}(\cdot,\cdot)$ is replaced by $a^{[N]}(\xi,\eta) = cN^{-1}$ for all $(\xi,\eta) \in [N]$, where $c \in (0,\infty)$ is a constant. The seed-bank consists of only *one colour* and the exchange rates between active and dormant are given by Ke, e. The state space is

$$S = \mathfrak{s}^{[N]}, \qquad \mathfrak{s} = [0,1]^2, \tag{6.13}$$

the system is denoted by

$$Z^{[N]}(t) = \big(X^{[N]}(t), Y^{[N]}(t)\big)_{t\geq 0}, \qquad \big(X^{[N]}(t), Y^{[N]}(t)\big) = \big(x_i^{[N]}(t), y_i^{[N]}(t)\big)_{i\in[N]}, \tag{6.14}$$

and its components evolve according to the SSDE

$$\begin{aligned} \mathrm{d}x_i^{[N]}(t) &= \frac{c}{N}\sum_{j\in[N]} \big[x_j^{[N]}(t) - x_i^{[N]}(t)\big]\,\mathrm{d}t + \sqrt{g\big(x_i^{[N]}(t)\big)}\,\mathrm{d}w_i(t) \\ &\quad + Ke\,\big[y_i^{[N]}(t) - x_i^{[N]}(t)\big]\,\mathrm{d}t, \\ \mathrm{d}y_i^{[N]}(t) &= e\,\big[x_i^{[N]}(t) - y_i^{[N]}(t)\big]\,\mathrm{d}t, \qquad i \in [N], \end{aligned} \tag{6.15}$$

which is the special case of (4.20) obtained by setting $a^{\Omega_N}(\eta,\xi) = 0$ if $d(\eta,\xi) > 1$ and $K_m = e_m = 0$ for $m \geq 1$. It is natural to take an *exchangeable random initial state*, because the evolution preserves exchangeability. According to De Finetti's theorem, there is no loss of generality in taking an i.i.d. initial state, i.e.,

$$\mathcal{L}\big[X^{[N]}(0), Y^{[N]}(0)\big] = \mu^{\otimes[N]}, \qquad \mu \in \mathcal{P}\left([0,1]^2\right). \tag{6.16}$$

By [67, Theorem 3.1], the SSDE in (6.15) is the unique weak solution of a well-posed martingale problem. By [67, Theorem 3.2], for every deterministic initial state $(X^{[N]}(0), Y^{[N]}(0))$, (6.15) has a unique strong solution. We are interested in the limit $N \to \infty$. For the limiting process we define

$$(Z(t))_{t\geq 0} = (X(t), Y(t))_{t\geq 0} = \big((x_i(t), y_i(t))_{i\in\mathbb{N}_0}\big)_{t\geq 0} \tag{6.17}$$

with components evolving according to (6.2), i.e.,

$$\begin{aligned}\mathrm{d}x_i(t) &= c\,[\mathbb{E}[x_i(t)] - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w(t)\\ &\qquad\qquad\qquad + Ke\,[y_i(t) - x_i(t)]\,\mathrm{d}t,\\ \mathrm{d}y_i(t) &= e\,[x_i(t) - y_i(t)]\,\mathrm{d}t, \qquad i \in \mathbb{N}_0,\end{aligned}\tag{6.18}$$

with $\mathcal{L}\left[(X(0), Y(0))\right] = \mu$ for some exchangeable $\mu \in \mathcal{P}(([0,1]^2)^{\otimes[\mathbb{N}_0]}$. Note that (6.18) consists of i.i.d. copies of the single-colony McKean-Vlasov process in (6.1), labelled by $i \in \mathbb{N}_0$.

Proposition 6.1.4 (Infinite-system McKean-Vlasov limit: convergence). *Suppose that* $\mathcal{L}[(X^{[N]}(0), Y^{[N]}(0))] = \mu^{[N]}$ *is exchangeable and*

$$\theta = \mathbb{E}^{\mu^{[N]}}\left[\frac{x(0) + Ky(0)}{1+K}\right].\tag{6.19}$$

Then

$$\lim_{N\to\infty} \mathcal{L}\left[\left(X^{[N]}(t), Y^{[N]}(t)\right)_{t\geq 0}\right] = \mathcal{L}\left[(X(t), Y(t))_{t\geq 0}\right]\tag{6.20}$$

with

$$\mathcal{L}\left[(X(0), Y(0))_{t\geq 0}\right] = \mu, \qquad \mu = \lim_{N\to\infty} \mu^{[N]},\tag{6.21}$$

where the limit is the McKean-Vlasov process in (6.1)–(6.2).

The proof of Proposition 6.1.4 is given in Section 6.1.4. For the system without seed-bank the McKean-Vlasov limit was proved in [38]. The fact that the components decouple is a property referred to as *propagation of chaos.*

CHAPTER 6

§6.1.3 Proof of equilibrium and ergodicity

In this section we prove Proposition 6.1.2.

Proof. Note that, by (4.111), we can rewrite (6.2) as

$$\begin{aligned}\mathrm{d}x(t) &= c\,\left[\theta + \frac{K}{1+K}(\theta_x - \theta_y)\,\mathrm{e}^{-(K+1)et} - x(t)\right]\,\mathrm{d}t + \sqrt{g(x(t))}\,\mathrm{d}w(t)\\ &\qquad + Ke\,[y(t) - x(t)]\,\mathrm{d}t,\\ \mathrm{d}y(t) &= e\,[x(t) - y(t)]\,\mathrm{d}t.\end{aligned}\tag{6.22}$$

Existence and uniqueness of a strong solution is again standard (see e.g. [72, Theorem 1] and recall Remark 6.1.1). We start by proving existence and uniqueness of the equilibrium. Afterwards we show that the solution converges to this equilibrium.

Consider two copies (x_1, y_1) and (x_2, y_2) of the system defined in (6.22), with $\mathcal{L}[(x_1(0), y_1(0))] = \mu_1$ and $\mathcal{L}[(x_2(0), y_2(0))] = \mu_2$, where μ_1 and μ_2 satisfy

$$\mathbb{E}^{\mu_1}\left[\frac{x_1(0) + Ky_1(0)}{1+K}\right] = \theta = \mathbb{E}^{\mu_2}\left[\frac{x_2(0) + Ky_2(0)}{1+K}\right]\tag{6.23}$$

for some $\theta \in [0,1]$. Write

$$\theta_{x_1} = \mathbb{E}^{\mu_1}[x_1(0)], \quad \theta_{y_1} = \mathbb{E}^{\mu_1}[y_1(0)], \quad \theta_{x_2} = \mathbb{E}^{\mu_2}[x_2(0)], \quad \theta_{y_2} = \mathbb{E}^{\mu_2}[y_2(0)].\tag{6.24}$$

Couple the two systems by coupling their Brownian motions. Denote the coupled process by

$$\begin{aligned}
&(\bar{z}(t))_{t\geq 0} = (z_1(t), z_2(t))_{t\geq 0}, \quad z_1(t) = (x_1(t), y_1(t)), \quad z_2(t) = (x_2(t), y_2(t)), \\
&\mathcal{L}(\bar{z}(0)) = \mu_1 \times \mu_2,
\end{aligned} \tag{6.25}$$

which has a unique strong solution. Put

$$\Delta(t) = x_1(t) - x_2(t), \qquad \delta(t) = y_1(t) - y_2(t). \tag{6.26}$$

To show that the equilibrium is unique, it is enough to show that

$$\lim_{t\to\infty} \mathbb{E}\left[|\Delta(t)| + EK|\delta(t)|\right] = 0. \tag{6.27}$$

Using a generalised form of Itô's formula, we find

$$\begin{aligned}
\mathrm{d}|\Delta(t)| &= (\mathrm{sgn}\ \Delta(t))\, \mathrm{d}\Delta(t) + \mathrm{d}L_t^0 \\
&= (\mathrm{sgn}\ \Delta(t))\, c\left[\frac{K}{1+K}\big((\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})\big)\, \mathrm{e}^{-(K+1)et} - \Delta(t)\right] \mathrm{d}t \\
&\quad\quad + (\mathrm{sgn}\ \Delta(t))\left(\sqrt{g(x_1(t))} - \sqrt{g(x_2(t))}\right) \mathrm{d}w(t) \\
&\quad\quad + (\mathrm{sgn}\ \Delta(t))\, Ke\left[\delta(t) - \Delta(t)\right] \mathrm{d}t,
\end{aligned} \tag{6.28}$$

where we use that the local time L_t^0 (see [63, Section IV.43]) of $\Delta(t)$ at 0 equals 0, since g is Lipschitz (see [63, Proposition V.39.3]). Again using Itô's formula, we also find

$$\mathrm{d}|\delta(t)| = (\mathrm{sgn}\ \delta(t))\, \mathrm{d}\delta(t) = (\mathrm{sgn}\ \delta(t))\, e\left[\Delta(t) - \delta(t)\right] \mathrm{d}t. \tag{6.29}$$

Taking expectations in (6.28)–(6.29), we get

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[|\Delta(t)| + K|\delta(t)|] \\
&= \mathbb{E}\left[c\left[(\mathrm{sgn}\ \Delta(t))\frac{K}{1+K}\big((\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})\big)\, \mathrm{e}^{-(K+1)et} - |\Delta(t)|\right]\right] \\
&\quad + Ke\,\mathbb{E}\Big[(\mathrm{sgn}\ \Delta(t) - \mathrm{sgn}\ \delta(t))\,(\delta(t) - \Delta(t))\Big] \\
&= \mathbb{E}\left[c\,(\mathrm{sgn}\ \Delta(t))\frac{K}{1+K}\big((\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})\big)\, \mathrm{e}^{-(K+1)et}\right] \\
&\quad - c\,\mathbb{E}[|\Delta(t)|] \\
&\quad - 2Ke\,\mathbb{E}\left[1_{\{\mathrm{sgn}\ \delta(t) \neq \mathrm{sgn}\,\Delta(t)\}}\,(|\delta(t)| + |\Delta(t)|)\right].
\end{aligned} \tag{6.30}$$

Define

$$h(t) = c\,\mathbb{E}[|\Delta(t)|] + 2Ke\,\mathbb{E}\left[1_{\{\mathrm{sgn}\ \delta(t)) \neq \mathrm{sgn}\ \Delta(t)\}}\,(|\delta(t)| + |\Delta(t)|)\right]. \tag{6.31}$$

Then $h(t)$ satisfies

(a) $h(t) > 0$.

(b) $0 \leq \int_0^\infty \mathrm{d}t\, h(t) \leq 1{+}K{+}c\,|(\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})|\, \frac{K}{K+1}\frac{1}{e(K+1)}\left[1 - \mathrm{e}^{-(K+1)et}\right]$.

(c) h is differentiable with h' bounded (see [43, Appendix D]).

Hence it follows that $\lim_{t\to\infty} h(t) = 0$, which implies that

$$\lim_{t\to\infty} \mathbb{E}\left[|\Delta(t)|\right] = 0. \tag{6.32}$$

We are left to prove that $\lim_{t\to\infty} \mathbb{E}[|\delta(t)|] = 0$. To do so, we define

$$f(t) = \mathbb{E}[|\delta(t)|], \qquad G(t) = e\,\mathbb{E}[(\mathrm{sgn}\ \delta(t))\Delta(t)]. \tag{6.33}$$

Note that G is bounded and continuous. Taking expectations in (6.29), we find

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = -ef(t) + G(t), \tag{6.34}$$

Solving (6.34) explicitly, we find that

$$f(t) = f(r)\,\mathrm{e}^{-e(t-r)} + \int_r^t \mathrm{d}s\,\mathrm{e}^{-e(t-s)} G(s), \qquad r,t \in \mathbb{R},\, t > r \geq 0. \tag{6.35}$$

By (6.32), for each $\epsilon > 0$ we can find an $r \in \mathbb{R}$ such that $\mathbb{E}[|\Delta(s)|] < \epsilon$ for all $s > r$, and hence $\sup_{t>r} |G(t)| < \epsilon$. Therefore

$$f(t) \leq f(r)\,\mathrm{e}^{-e(t-r)} + \epsilon \tag{6.36}$$

and, since $|f| < 1$, we find, for each $\epsilon > 0$,

$$\lim_{t\to\infty} f(t) < \epsilon. \tag{6.37}$$

Therefore $\lim_{t\to\infty} \mathbb{E}[|\delta(t)|] = 0$, which completes the proof of uniqueness of the equilibrium for given θ.

To prove existence of the equilibrium, let $(t_n)_{n\in\mathbb{N}}$ be any increasing sequenc eof times such that $\lim_{n\to\infty} t_n = \infty$. Let $\mu = \mathcal{L}[(x(0), y(0))]$ be any initial measure of the system in (6.22) with $\mathbb{E}^\mu[\frac{x(0)+Ky(0)}{1+K}] = \theta$, and let $\mu(t_n) = \mathcal{L}[(x(t_n), y(t_n))]$. Since the state space is compact, the sequence $(\mu(t_n))_{n\in\mathbb{N}}$ is tight, and by Prohorov's theorem we can find a converging subsequence $(\mu(t_{n_k}))_{k\in\mathbb{N}}$. Put $\nu = \lim_{k\to\infty} \mu(t_{n_k})$. We will show that ν is invariant. To that end, recall from Section 6.1.1 that

$$\mathbb{E}^\mu\left[\frac{x(t) + Ky(t)}{1+K}\right] = \theta, \qquad t \geq 0. \tag{6.38}$$

Hence we can use the coupling in (6.25) to show that the system starting in μ and the system starting $\mu(t)$ converge to the same law as $t \to \infty$, from which it follows that $\lim_{k\to\infty} \mu(t + t_{n_k}) = \nu$. Let $(S_t)_{t\geq 0}$ denote the semigroup of the system in (6.22). By the Feller property for semigroups,

$$S_t\nu = \lim_{k\to\infty} S_t\mu(t_{n_k}) = \lim_{k\to\infty} S_{t_{n_k}}(S_t\mu) = \nu, \tag{6.39}$$

where in the last equality we use the uniqueness of the equilibrium given θ. Thus, ν is an invariant measure. To exhibit its dependence on θ we write ν_θ. Using the same coupling as in (6.25), and starting from $\mu \times \nu_\theta$ with ν_θ the invariant measure just obtained, we see that for every θ the system in (6.2) converges to a unique equilibrium measure ν_θ, and so (6.11) is immediate from (6.38). □

§6.1.4 Proof of McKean-Vlasov limit

In this section we give a sketch of the proof of Proposition 6.1.4. In Chapters 6.2-9 we encounter more difficult versions of Proposition 6.1.4. There we will give the proofs in full detail.

Proof. Since we start from a distribution $\mu(0)$ that is exchangeable, Aldous's ergodic theorem gives

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j\in[N]} x_j(0) = \mathbb{E}^{\mu(0)}[x_0] \quad \mathbb{P}\text{-a.s.} \tag{6.40}$$

By Ioffe's theorem [25, Eqs. (1.1)–(1.2)], tightness of the associated sequence of processes (uniformly on the state space) follows from boundedness of the generator as an operator. To apply the generator criterion in [49] we must show propagation of chaos and prove the weak law of large numbers

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j\in[N]} x_j(t) = \mathbb{E}[x_0(t)]. \tag{6.41}$$

The propagation of chaos and the weak law of large numbers for $t > 0$ therefore follows from [38, Section 4]. Since the martingale problem is well-posed [38, Section 2], the limiting process exists and is unique. □

§6.2 Proofs: $N \to \infty$, mean-field finite-systems scheme

Chapter 6

In Sections 6.2.1 we introduce the so called mean-field finite-systems scheme for the mean-field system introduced in Section 6.1.2. In Section 6.2.2 we outline the *abstract scheme* behind the proof behind the mean-field finite-systems scheme. The computations in the proof of the abstract scheme are long and technical, and are deferred to Section 6.3.

§6.2.1 Mean-field finite-systems scheme

In this section we describe the limiting dynamics of the finite system in (6.14) from a *multiple space-time scale* viewpoint. To do so, we need the following limiting SSDE for the infinite system $Z(t) = (z_i(t))_{i\in\mathbb{N}_0} = (x_i(t), y_i(t))_{i\in\mathbb{N}_0}$, with initial law $\mathcal{L}[Z(0)] = \mu^{\otimes\mathbb{N}_0}$, evolving according to

$$\begin{aligned} \mathrm{d}x_i(t) &= c\,[\theta - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) + Ke\,[y_i(t) - x_i(t)]\,\mathrm{d}t, \\ \mathrm{d}y_i(t) &= e\,[x_i(t) - y_i(t)]\,\mathrm{d}t, \; i \in \mathbb{N}_0, \end{aligned} \tag{6.42}$$

where θ is defined in (6.11). Note that each component of (6.42) is an autonomous copy of the McKean-Vlasov process in (6.2) in equilibrium.

For the multiscale analysis we will need the following ingredients:

(a) The *estimator* for the finite system is defined by

$$\bar{\Theta}^{[N]}(t) = \bar{\Theta}^{[N]}\big(Z^{[N]}(t)\big) = \frac{1}{N} \sum_{i\in[N]} \frac{x_i^{[N]}(t) + Ky_i^{[N]}(t)}{1+K} \tag{6.43}$$

and its active and dormant counterparts

$$\begin{aligned}\bar{\Theta}_x^{[N]}(t) &= \frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(t),\\ \bar{\Theta}_y^{[N]}(t) &= \frac{1}{N}\sum_{i\in[N]} y_i^{[N]}(t).\end{aligned} \tag{6.44}$$

(b) The *time scale* N, on which $\lim_{N\to\infty}\mathcal{L}[\bar{\Theta}^{[N]}(L(N)) - \bar{\Theta}^{[N]}(0)] = \delta_0$ for all $L(N)$ such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, but not for $L(N) = N$. In words, N is the time scale on which $\bar{\Theta}^{[N]}(\cdot)$ starts evolving, i.e., $\big(\bar{\Theta}^{[N]}(Ns)\big)_{s>0}$ is not a fixed process. When we scale time by N, putting $t = Ns$, we view s as the "fast time scale" and t as the "slow time scale".

(c) The *invariant measure*, i.e., the equilibrium measure of a single component in (6.42) written

$$\Gamma_\theta, \tag{6.45}$$

and the *invariant measure* of the infinite system in (6.42), written $\nu_\theta = \Gamma_\theta^{\otimes\mathbb{N}_0}$, with $\theta\in[0,1]$ controlled by the initial measure (recall (6.4)–(4.111)).

(d) The *renormalisation transformation* $\mathcal{F}\colon \mathcal{G}\to\mathcal{G}$,

$$(\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x)\,\nu_\theta(\mathrm{d}x,\mathrm{d}y_0), \quad \theta\in[0,1], \tag{6.46}$$

where ν_θ is the equilibrium measure of 6.42. Note that $\mathcal{F}$ is the same transformation as defined in (4.75), but for the truncated system. Note that we can also write

$$(\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x)\,\Gamma_\theta(\mathrm{d}x,\mathrm{d}y_0), \quad \theta\in[0,1], \tag{6.47}$$

where Γ_θ is as defined in (6.45).

(e) The *macroscopic observable* $\big(\bar{\Theta}(s)\big)_{s>0}$ satisfying the SSDE

$$\mathrm{d}\bar{\Theta}(s) = \frac{1}{1+K}\sqrt{\mathbb{E}^{\Gamma_{\bar{\Theta}(s)}}[g(u)]}\,\mathrm{d}w(s) = \frac{1}{1+K}\sqrt{(\mathcal{F}g)(\bar{\Theta}(s))}\,\mathrm{d}w(s), \tag{6.48}$$

To obtain the multi-scale limit dynamics for the system in (6.14), we speed up time by a factor N and define the process

$$\left(x_1^{[N]}(s), y_1^{[N]}(s)\right)_{s>0} = \left(\Theta_x^{[N]}(Ns), \Theta_y^{[N]}(Ns)\right)_{s>0}, \tag{6.49}$$

which is the analogue of the 1-block average in (4.22). We use the lower index 1 to indicate that the average is taken over $[N]$ components. Using (6.15), we see that the dynamics of (6.49) is given by the SSDE

$$\begin{aligned}\mathrm{d}x_1^{[N]}(s) &= \sqrt{\frac{1}{N}\sum_{i\in[N]} g\big(x_i(Ns)\big)}\,\mathrm{d}w(s) + NKe\left[y_1^{[N]}(s) - x_1^{[N]}(s)\right]\mathrm{d}s,\\ \mathrm{d}y_1^{[N]}(s) &= Ne\left[x_1^{[N]}(s) - y_1^{[N]}(s)\right]\mathrm{d}s.\end{aligned} \tag{6.50}$$

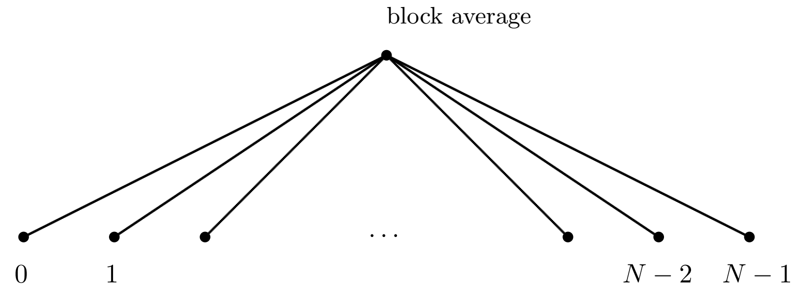

Figure 6.1: Given the value of the block average, the $N \gg 1$ constituent components equilibrate on a time scale that is fast with respect to the time scale on which the block average fluctuates. Consequently, the volatility of the block average is the expectation of the volatility of the constituent components under the conditional quasi-equilibrium.

In (6.50), in the limit as $N \to \infty$ infinite rates appear in the exchange between the active and the dormant population. However, looking at the process

$$\left(\frac{x_1^{[N]}(s) + K y_1^{[N]}(s)}{1+K}\right)_{s>0} = \left(\bar{\Theta}^{[N]}(Ns)\right)_{s>0} \tag{6.51}$$

we see that the terms carrying a factor N in front cancel out. Consequently, for the process in (6.51) we can use ideas from [20] to prove tightness as $N \to \infty$ in the *classical topology* of continuum path processes. We will show in Section 6.3.3 that

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\left[x_1^{[N]}(s) - y_1^{[N]}(s)\right]\right)_{s\geq 0}\right] = \mathcal{L}\left[(0)_{s\geq 0}\right] \tag{6.52}$$
in the *Meyer-Zheng topology*.

Combining (6.51) and (6.52), we obtain the multiple space-time scaling behaviour of the system in (6.14).

Proposition 6.2.1 (Mean-field: finite-systems scheme). *Suppose that the SSDE in* (6.15) *has initial measure $\mathcal{L}[Z^{[N]}(0)] = \mu^{\otimes[N]}$ for some $\mu \in \mathcal{P}\left([0,1]^2\right)$. Let*

$$\theta = \mathbb{E}^{\mu}\left[\frac{x + K y_0}{1+K}\right]. \tag{6.53}$$

(a) For the averages in (6.49)*,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(x_1^{[N]}(s), y_{0,1}^{[N]}(s)\right)_{s>0}\right] = \mathcal{L}\left[\left(x_1^{\mathbb{N}_0}(s), y_{0,1}^{\mathbb{N}_0}(s)\right)_{s>0}\right] \tag{6.54}$$
in the Meyer-Zheng topology,

where the limit process is the unique solution of the SSDE

$$\begin{aligned} \mathrm{d}x_1^{\mathbb{N}_0}(s) &= \frac{1}{1+K}\sqrt{(\mathcal{F}g)\big(x_1^{\mathbb{N}_0}(s)\big)}\,\mathrm{d}w(s), \\ y_{0,1}^{\mathbb{N}_0}(s) &= x_1^{\mathbb{N}_0}(s), \end{aligned} \tag{6.55}$$

with initial state

$$\left(x_1^{\mathbb{N}_0}(0), y_{0,1}^{\mathbb{N}_0}(0)\right) = (\theta, \theta)\,. \tag{6.56}$$

(b) For the weighted sum of the averages in (6.51),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\bar{\Theta}^{[N]}(Ns)\right)_{s>0}\right] = \mathcal{L}\left[\left(\bar{\Theta}(s)\right)_{s>0}\right], \tag{6.57}$$

where the limit is the macroscopic observable in (6.48) *with initial state*

$$\bar{\Theta}(0) = \theta. \tag{6.58}$$

(c) Define

$$\nu_\theta(s) = \int_{[0,1]} Q_s\big(\theta, \mathrm{d}\theta'\big)\, \nu_{\theta'} \in \mathcal{P}([0,1]^2), \tag{6.59}$$

where $Q_s(\theta, \cdot)$ is the time-s marginal law of the process $(\bar{\Theta}(s))_{s>0}$ starting from $\theta \in [0,1]$ (note that $\nu_\theta(0) = \nu_\theta$). Then, for every $s \in (0,\infty)$,

$$\lim_{N\to\infty} \mathcal{L}\left[\left(X^{[N]}(Ns+t), Y^{[N]}(Ns+t)\right)_{t>0}\right] = \mathcal{L}\left[(Z^{\nu_\theta(s)}(t))_{t>0}\right] \tag{6.60}$$

where, conditional on $\bar{\Theta}(s) = \theta$, $(z^{\nu_\theta(s)}(t))_{t\geq 0}$ is the random process in (6.42) *and $z^{\nu_\theta(s)}(0)$ is drawn according to $\nu_\theta(s)$ (which is a mixture of random processes in equilibrium).*

The proof of Proposition 6.2.1 is given in Section 6.2.2.

The result in Part (a) shows that the limit dynamics of the averages follows a similar type of diffusion as a single colony, but with four important changes:

- For the limit of the time-scaled average in (6.51) the diffusion function g is replaced by a *renormalised diffusion function* $\mathcal{F}g$, defined by (6.46) (recall Fig. 6.1). In section 6.2.2 we will show that $\mathcal{F}\mathcal{G} \subset \mathcal{G}$, i.e., $\mathcal{F}$ preserves the class of diffusion functions defined in (4.15).
- The average of the dormant population is the same as the average of the active population, and hence the term that accounts for the exchange between the active and the dormant population *vanishes*. This happens because when time is speeded up by a factor N also the rates of exchange between active and dormant are speeded up by a factor N (see (6.50)). Hence the exchange rates become infinitely large, which implies that the active and the dormant population equilibrate instantly in the Meyer-Zheng topology.
- Since we take the average over all the components, the migration terms in (6.15) cancel out against each other.
- Comparing the system in (6.14) with the system of interacting Fisher-Wright diffusions in the mean-field limit studied in [21], we see from (6.55) that the single-colour seed-bank *slows down the average* by a factor $1/(1+K)$, but does not change the system qualitatively. This is a direct consequence of the fact that the averages of the active and the dormant population equilibrate (due to the infinite rates), while only individuals in the active part of the population resample.

The result in Part (b) shows that the limit dynamics of the averages in (6.51) follows an autonomous SDE, with convergence in the classical topology, i.e., in $C_b([0,\infty),[0,1])$. The Brownian motion in (6.1) is taken to be independent of the initial state. The result in Part (c) says that, on time scale 1 and starting from time Ns with $N \to \infty$, the system has a *McKean-Vlasov limit*, i.e., exhibits propagation of chaos, with components that are versions of a McKean-Vlasov process with a random initial state whose law depends on s. So, in particular, the components become independent, and we see *decoupling* The proof of Part (c) will use Part (b). The proof of Part (a) will follow from Part (b) after we use the Meyer-Zheng topology.

Remark 6.2.2 (Basic multi-scale). Note that Proposition 6.2.1 already reveals several phenomena that we encountered in Theorems 4.4.2 and 4.4.4, capturing the hierarchical multiscale behaviour. Even for the one-layer mean-field system we find decoupling of components, the occurrence of a renormalisation transformation, equalisation of the seed-bank with the active population, and the need for the Meyer-Zheng topology. Later we will see that the role of the macroscopic observable $\bar{\Theta}$ is the same as that of the effective process. ■

Remark 6.2.3 (Interchange of limits). The notation $x_1^{\mathbb{N}_0}, y_{0,1}^{\mathbb{N}_0}$ indicates that the limit arises from taking averages over $[N]$ and letting $N \to \infty$. Note that, for i.i.d. initial states,

$$x_1^{\mathbb{N}_0}(0) = \lim_{N\to\infty} x_1^{[N]}(0) = \lim_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} x_i(0) = \theta_x \qquad \mathbb{P}\text{-a.s.} \tag{6.61}$$

On the other hand, picking any sequence of times $L(N)$ such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, we get

$$x_1^{\mathbb{N}_0}(0+) = \lim_{N\to\infty} x_1^{[N]}\big(\tfrac{L(N)}{N}t\big) = \lim_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} x_i(L(N)t) = \theta \qquad \mathbb{P}\text{-a.s.} \tag{6.62}$$

The mismatch between (6.61) and (6.62) indicates that we must be careful with interchanging the limits $N \to \infty$ and $s \downarrow 0$. This is why (6.54), which lives on the fast time scale, is restricted to $s > 0$. ■

§6.2.2 Abstract scheme behind finite-systems scheme

To prove Proposition 6.2.1, we follow the abstract scheme outlined in [25, p. 2314–2315] and based on [21], [20]. Below we state the abstract scheme for our model. The scheme consists of 4 steps, each of the steps consists of a series of propositions and lemmas. The proofs of these are given in Section 6.3.

Step 1. Equilibrium of the single components. This step fixes the one-dimensional distributions of the single components when $t, N \to \infty$ in a combined way, and is the equivalent of [21, Proposition 1]. Recall that $\bar{\Theta}^{[N]}$ is defined in (6.43).

Proposition 6.2.4 (Equilibrium for the infinite system). *Let $(N_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{N}$. Fix $s>0$. Let $L(N)$ be such that $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N=0$, and suppose that*

$$\begin{aligned}
&\lim_{k\to\infty} \mathcal{L}\left[\bar\Theta^{[N_k]}(N_k s)\right] = P_s,\\
&\lim_{k\to\infty} \mathcal{L}\left[\sup_{0\le t\le L(N_k)} \left|\bar\Theta^{[N_k]}(N_k s)-\bar\Theta^{[N_k]}(N_k s-t)\right|\right] = \delta_0,\\
&\lim_{k\to\infty} \mathcal{L}\big(X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s)\big) = \nu(s).
\end{aligned} \tag{6.63}$$

Then $\nu(s)$ is of the form

$$\nu(s) = \int_{[0,1]} P_s(\mathrm{d}\theta)\,\nu_\theta, \tag{6.64}$$

where ν_θ is the equilibrium measure of the process defined in (6.42).

Proposition 6.2.4 follows from the following seven lemmas, which are the analogues of the five lemmas used in [21, p. 477–478] for the system without seed-bank.

The first lemma establishes convergence of the infinite system in (6.42) to its equilibrium.

Lemma 6.2.5 (Convergence for the infinite system). *Let μ be an exchangeable probability measure on $([0,1]^2)^{\mathbb{N}_0}$. Then for the system $(Z(t))_{t\ge0}$ given by* (6.42) *with $\mathcal{L}(Z(0))=\mu$,*

$$\lim_{t\to\infty} \mathcal{L}[Z(t)] = \nu_\theta, \tag{6.65}$$

where ν_θ is of the form

$$\nu_\theta = \Gamma_\theta^{\otimes\mathbb{N}_0}, \tag{6.66}$$

with Γ_θ the equilibrium of the single-colony process defined in (6.45). *Moreover, ν_θ is ergodic.*

The second lemma establishes the continuity of the equilibrium with respect to its center of drift θ.

Lemma 6.2.6 (Continuity of the equilibrium). *Let $\mathcal{P}([0,1]^{\mathbb{N}_0})$ denote the space of probability measures on $[0,1]^{\mathbb{N}_0}$. The mapping $[0,1]\to\mathcal{P}([0,1]^{\mathbb{N}_0})$ given by*

$$\theta \mapsto \nu_\theta \tag{6.67}$$

is continuous. Furthermore, if h is a Lipschitz function on $[0,1]$, then also $\mathcal{F}h$ defined by

$$(\mathcal{F}h)(\theta) = \mathbb{E}^{\nu_\theta}[h(\cdot)] = \int_{([0,1]^2)^{\mathbb{N}_0}} \nu_\theta(\mathrm{d}z)\,h(x_0) \tag{6.68}$$

is a Lipschitz function on $[0,1]$.

The third lemma characterises the speed at which the estimators $\Theta_x^{[N]}$ and $\Theta_y^{[N]}$ converge to each other.

Chapter 6

Lemma 6.2.7 (Comparison of empirical averages). *Let* $(\Theta_x^{[N]}(t))_{t\geq 0}$ *and* $(\bar{\Theta}_y^{[N]}(t))_{t\geq 0}$ *be defined as in* (6.44)*, and define*

$$\Delta_{\bar{\Theta}}^{[N]}(t) = \Theta_x^{[N]}(t) - \Theta_y^{[N]}(t). \tag{6.69}$$

Then

$$\mathbb{E}\left[\left|\Delta_{\bar{\Theta}}^{[N]}(t)\right|\right] \leq \sqrt{\mathbb{E}\left[\left(\Delta_{\bar{\Theta}}^{[N]}(0)\right)^2\right]}\,\mathrm{e}^{-(Ke+e)t} + \sqrt{\frac{\|g\|}{N(Ke+e)}}. \tag{6.70}$$

Remark 6.2.8 (Key estimate for Meyer-Zheng convergence). The estimate in (6.70) in Lemma 6.2.7 will be the key estimate to show convergence of the active and dormant 1-block in Meyer-Zheng topology. Note that if we look at times Ns for $s > 0$, then (6.70) shows that $\mathbb{E}[|\Delta_{\bar{\Theta}}^{[N]}(Ns)|]$ is $\mathcal{O}(\sqrt{1/N})$. ■

The fourth lemma compares the finite system with an infinite system. To that end we construct both the finite and the infinite system on the same state-space by considering the finite system $(X^{[N]}(t), Y^{[N]}(t))$ as an element of $([0,1]^2)^{\mathbb{N}_0}$ via periodic continuation. Let $L(N)$ be such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, and define the distribution μ_N by continuing the configuration of $(X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N)))$ periodically to $([0,1]^2)^{\mathbb{N}_0}$. Define

$$\bar{\Theta}^{[N]} = \bar{\Theta}^{[N]}(Ns - L(N)). \tag{6.71}$$

Note that

$$\begin{aligned}\bar{\Theta}^{[N]} &= \frac{\frac{1}{N}\sum_{j\in[N]} x_j^{[N]}(Ns - L(N)) + \frac{K}{N}\sum_{j\in[N]} y_j^{[N]}(Ns - L(N))}{1+K}\\ &= \frac{1}{N}\sum_{j\in[N]} \frac{x_j^{\mu_N}(0) + K y_j^{\mu_N}(0)}{1+K}.\end{aligned} \tag{6.72}$$

Thus, $\bar{\Theta}^{[N]}$ is a random variable whose law depends on $\mathcal{L}\left[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))\right] = \mu_N$. The infinite system with initial law μ_N is denoted by

$$\left(X^{\mu_N}(t), Y^{\mu_N}(t)\right)_{i\in\mathbb{N}_0,\ t\geq 0} = \left(x_i^{\mu_N}(t), y_i^{\mu_N}(t)\right)_{t\geq 0} \tag{6.73}$$

and evolves according to

$$\begin{aligned}\mathrm{d}x_i^{\mu_N}(t) &= c\big[\bar{\Theta}^{[N]} - x_i^{\mu_N}(t)\big]\,\mathrm{d}t + \sqrt{g(x_i^{\mu_N}(t))}\,\mathrm{d}w_i(t) + Ke\,[y_i^{\mu_N}(t) - x_i^{\mu_N}(t)]\,\mathrm{d}t,\\ \mathrm{d}y_i^{\mu_N}(t) &= e\,[x_i^{\mu_N}(t) - y_i^{\mu_N}(t)]\,\mathrm{d}t, \qquad i\in\mathbb{N}_0,\end{aligned} \tag{6.74}$$

where $\{w_i\}_{i\in\mathbb{N}_0}$ is a collection of independent Brownian motions.

Lemma 6.2.9. *[Comparison of finite and infinite systems] Fix* $s > 0$ *and assume that, for any* $L(N)$ *satisfying* $\lim_{N\to\infty} L(N) = \infty$ *and* $\lim_{N\to\infty} L(N)/N = 0$,

$$\lim_{N\to\infty} \sup_{0\leq t\leq L(N)} \left|\bar{\Theta}^{[N]}(Ns) - \bar{\Theta}^{[N]}(Ns - t)\right| = 0 \quad \textit{in probability.} \tag{6.75}$$

Let

$$\left(X^{\mu_N}(t), Y^{\mu_N}(t)\right)_{t\geq 0} \tag{6.76}$$

be the infinite system defined in (6.74) *starting in the distribution* μ_N*, where* μ_N *is defined by continuing the configuration of* $(X^{[N]}(Ns-L(N)), Y^{[N]}(Ns-L(N)))$ *periodically to* $([0,1]^2)^{\mathbb{N}_0}$*. Similarly, view* $(X^{[N]}(t), Y^{[N]}(t))$ *as an element of* $([0,1]^2)^{\mathbb{N}_0}$ *by periodic continuation. Then, for all* $t \geq 0$*,*

$$\begin{aligned}&\lim_{N\to\infty}\left|\mathbb{E}\left[f\left(X^{\mu_N}(t), Y^{\mu_N}(t)\right) - f\left(X^{[N]}(Ns-L(N)+t), Y^{[N]}(Ns-L(N)+t)\right)\right]\right| = 0\\ &\qquad \forall\, f \in \mathcal{C}\left(([0,1]^2)^{\mathbb{N}_0}, \mathbb{R}\right).\end{aligned} \tag{6.77}$$

Before we can prove that the infinite system $(X^{\mu_N}(t), Y^{\mu_N}(t))_{t\geq 0}$ converges to some limiting system as $N \to \infty$, we need the following regularity property for the estimator $\bar{\Theta}^{[N]}$. This is stated in our fifth lemma.

Lemma 6.2.10 (Stability of the estimator for the conserved quantity). *Define* μ_N *as in Lemma 6.2.9. Let* $(x_i, y_i)_{i\in[N]}$ *be distributed according to the exchangeable probability measure* μ_N *on* $([0,1]^2)^{\mathbb{N}_0}$ *restricted to* $([0,1]^2)^{[N]}$*. Suppose that* $\lim_{N\to\infty}\mu_N = \mu$ *for some exchangeable probability measure* μ *on* $([0,1]^2)^{\mathbb{N}_0}$*. Define a random variable* ϕ *on* $(\mu, ([0,1]^2)^{\mathbb{N}_0})$ *by putting*

$$\phi = \lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]}\frac{x_i + Ky_i}{1+K}, \tag{6.78}$$

and a random variable ϕ_N *on* $(\mu_N, ([0,1]^2)^{\mathbb{N}_0})$ *by putting*

$$\phi_N = \frac{1}{N}\sum_{i\in[N]}\frac{x_i + Ky_i}{1+K}. \tag{6.79}$$

Then

$$\lim_{N\to\infty}\mathcal{L}[\phi_N] = \mathcal{L}[\phi]. \tag{6.80}$$

In the sixth lemma we state the convergence of the law $\mathcal{L}[(X^{\mu_N}(t), Y^{\mu_N}(t))]$ to the law of a limiting system as $N \to \infty$.

Lemma 6.2.11 (Uniformity of the ergodic theorem for the infinite system). *Let* μ_N *be defined as in Lemma 6.2.9. Since* $(\mu_N)_{N\in\mathbb{N}}$ *is tight, it has convergent subsequences. Let* $(N_k)_{k\in\mathbb{N}}$ *be a subsequence such that* $\mu = \lim_{k\to\infty}\mu_{N_k}$*. Define*

$$\bar{\Theta} = \lim_{N\to\infty}\frac{1}{N}\sum_{i\in[N]}\frac{x_i^\mu + Ky_i^\mu}{1+K} \qquad \text{in } L_2(\mu), \tag{6.81}$$

and let $(X^\mu(t), Y^\mu(t))_{t\geq 0}$ *be the infinite system evolving according to*

$$\begin{aligned}\mathrm{d}x_i^\mu(t) &= c\left[\bar{\Theta} - x_i^\mu(t)\right]\mathrm{d}t + \sqrt{g(x_i^\mu(t))}\,\mathrm{d}w_i(t) + Ke\left[y_i^\mu(t) - x_i^\mu(t)\right]\mathrm{d}t,\\ \mathrm{d}y_i^\mu(t) &= e\left[x_i^\mu(t) - y_i^\mu(t)\right]\mathrm{d}t, \qquad i \in \mathbb{N}_0.\end{aligned} \tag{6.82}$$

Then

(a) *For all* $t \geq 0$,

$$\begin{aligned} &\lim_{k\to\infty} \big|\mathbb{E}\big[f\big(X^{\mu_{N_k}}(t), Y^{\mu_{N_k}}(t)\big)\big] - \mathbb{E}\big[f\big(X^{\mu}(t), Y^{\mu}(t)\big)\big]\big| = 0, \\ &\qquad \forall\, f \in \mathcal{C}\big(([0,1]^2)^{\mathbb{N}_0}, \mathbb{R}\big). \end{aligned} \tag{6.83}$$

(b) *There exists a sequence* $(\bar{L}(N))_{N\in\mathbb{N}}$ *satisfying* $\lim_{N\to\infty} \bar{L}(N) = \infty$ *and* $\lim_{N\to\infty} \bar{L}(N)/N = 0$ *such that*

$$\begin{aligned} &\lim_{k\to\infty} \big|\mathbb{E}\big[f\big(X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big) \\ &\qquad - f\big(X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k))\big)\big|\big] \\ &+ \big|\mathbb{E}\big[f\big(X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k))\big)\big] - \mathbb{E}\big[f\big(X^{\mu}(\bar{L}(N_k)), Y^{\mu}(\bar{L}(N_k))\big)\big]\big| = 0 \\ &\qquad \forall\, f \in \mathcal{C}\big(([0,1]^2)^{\mathbb{N}_0}, \mathbb{R}\big). \end{aligned} \tag{6.84}$$

Remark 6.2.12 (Existence of $\bar{\Theta}$). Note that the limit in (6.81) is well-defined by the ergodic theorem in L_2, since μ is the limit of translation invariant measures and hence is itself translation invariant. ■

In the seventh lemma we provide a coupling of two copies of the finite system starting from different measures.

Lemma 6.2.13 (Coupling of finite systems). *Let* $(X^{[N],1}, Y^{[N],1})$ *be a finite system evolving according to* (6.15) *and starting from some exchangeable measure. Let* $\mu^{[N],1}$ *be the measure obtain by periodic continuation of the configuration of* $(X^{[N],1}(0), Y^{[N],1}(0))$. *Similarly, let* $(X^{[N],2}, Y^{[N],2})$ *be a finite system evolving according to* (6.15) *and starting from some exchangeable measure. Let* $\mu^{[N],2}$ *be the measure obtain by periodic continuation of the configuration of* $(X^{[N],2}(0), Y^{[N],2}(0))$. *Let* $\tilde{\mu}$ *be any weak limit point of the sequence of measures* $\{\mu^{[N],1} \times \mu^{[N],2}\}_{N\in\mathbb{N}}$. *Define random variables* $\bar{\Theta}^{[N],1}$ *on* $(\mu^{[N],1}, ([0,1]^2)^{\mathbb{N}_0})$, $\bar{\Theta}^{[N],2}$ *on* $(\mu^{[N],1}, ([0,1]^2)^{\mathbb{N}_0})$ *and* $\bar{\Theta}_1$ *and* $\bar{\Theta}_2$ *on* $(\mu, ([0,1]^2)^{\mathbb{N}_0})$ *by*

$$\begin{aligned} &\bar{\Theta}^{[N],1} = \frac{1}{N} \sum_{i\in[N]} \frac{x_i^{[N],1} + K y_i^{[N],1}}{1+K}, \qquad \bar{\Theta}^{[N],2} = \frac{1}{N} \sum_{i\in[N]} \frac{x_i^{[N],2} + K y_i^{[N],2}}{1+K}, \\ &\bar{\Theta}_1 = \lim_{n\to\infty} \frac{1}{n} \sum_{i\in[n]} \frac{x_i^1 + K y_i^1}{1+K}, \qquad \bar{\Theta}_2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i\in[n]} \frac{x_i^2 + K y_i^2}{1+K}, \end{aligned} \tag{6.85}$$

and let $(\bar{\Theta}^{[N],1}(t))_{t\geq 0}$ *and* $(\bar{\Theta}^{[N],2}(t))_{t\geq 0}$ *be defined according to* (6.43) *for* $(X_1^{[N]}, Y_1^{[N]})$, *respectively,* $(X_2^{[N]}, Y_2^{[N]})$. *Assume that*

$$\lim_{N\to\infty} \sup_{0\leq t\leq L(N)} \big|\bar{\Theta}^{[N],k}(0) - \bar{\Theta}^{[N],k}(t)\big| = 0 \quad \textit{in probability}, \quad k \in \{1,2\}, \tag{6.86}$$

and suppose that $\tilde{\mu}(\{\bar{\Theta}_1 = \bar{\Theta}_2\}) = 1$. *Then, for any sequence* $t(N) \to \infty$,

$$\lim_{N\to\infty} \mathbb{E}\left[\big|x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))\big| + K\big|y_i^{[N],1}(t(N)) - y_i^{[N],2}(t(N))\big|\right] = 0. \tag{6.87}$$

Chapter 6

Step 2. Convergence of the estimator. This step is the equivalent of [21, Proposition 2]. We first prove the tightness of the estimator $\bar{\Theta}^{[N]}$ in path space. After that we settle convergence of the finite-dimensional distributions and identify the limit.

Proposition 6.2.14 (Convergence of average sum process).

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\bar{\Theta}^{[N]}(Ns)\right)_{s>0}\right] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}], \tag{6.88}$$

where $(\bar{\Theta}(s))_{s>0}$ *evolves according to*

$$\mathrm{d}\bar{\Theta}(s) = \frac{1}{(1+K)}\sqrt{(\mathcal{F}g)(\bar{\Theta}(s))}\,\mathrm{d}w(s). \tag{6.89}$$

Proposition 6.2.14 follows from the following three lemmas, which are the equivalent of the three lemmas used in [21, p. 488–493] for the system without seed-bank.

Lemma 6.2.15 (Martingale property of average sum process).

(1) *The process* $(\bar{\Theta}^{[N]}(Ns))_{s>0}$ *is a square-integrable martingale with continuous paths and increasing process*

$$\left\langle \bar{\Theta}^{[N]}(Ns)\right\rangle_{s>0} = \frac{1}{(1+K)^2}\int_0^s \mathrm{d}r\,\frac{1}{N}\sum_{i\in[N]} g\big(x_i^{[N]}(Nr)\big). \tag{6.90}$$

(2) *Let* $L(N)$ *be such that* $\lim_{N\to\infty} L(N) = \infty$ *and* $\lim_{N\to\infty} L(N)/N = 0$*. Then*

$$\lim_{N\to\infty} \sup_{0\le t\le L(N)} \left|\bar{\Theta}^{[N]}(Ns) - \bar{\Theta}^{[N]}(Ns-t)\right| = 0 \text{ in probability.} \tag{6.91}$$

(3) $\big(\mathcal{L}[(\bar{\Theta}^{[N]}(Ns))_{s>0}]\big)_{N\in\mathbb{N}}$ *is tight as a sequence of probability measures on* $\mathcal{C}([0,\infty),[0,1])$.

Lemma 6.2.16 (Martingale property of limit process). *Let* $(N_k)_{k\in\mathbb{N}}$ *be any subsequence such that*

$$\lim_{k\to\infty} \mathcal{L}\left[\left(\bar{\Theta}^{[N_k]}(N_k s)\right)_{s>0}\right] = \mathcal{L}\big[(\bar{\Theta}(s))_{s>0}\big]. \tag{6.92}$$

Then $(\bar{\Theta}(s))_{s>0}$ *is a square-integrable martingale with continuous paths, and*

$$\left(\bar{\Theta}^2(s) - \int_0^s \mathrm{d}r\,\frac{1}{(1+K)^2}\,\mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right)_{s>0} \tag{6.93}$$

is a martingale.

Lemma 6.2.17 (Uniqueness). *The following martingale problem has a unique solution:*

$$\begin{gathered}(\bar{\Theta}_s)_{s>0} \text{ is a continuous martingale with values in } [0,1],\\ \left(\bar{\Theta}^2(s) - \frac{1}{(1+K)^2}\int_0^s \mathrm{d}r\,\mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right)_{s>0} \text{ is a martingale.}\end{gathered} \tag{6.94}$$

The solution of (6.94) *is given by the diffusion generated by* $\mathbb{E}^{\nu_u}[g(\cdot)]\frac{\partial^2}{\partial u^2}$.

Step 3. Convergence of the averages in the Meyer-Zheng topology. Recall the definition of the Meyer-Zheng topology in Section 4.4.1. We have to prove the following proposition.

Proposition 6.2.18 (Convergence in Meyer-Zheng topology). *If*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\bar{\Theta}(Ns)\right)_{s>0}\right] = \mathcal{L}\left[\left(\bar{\Theta}(s)\right)_{s>0}\right], \tag{6.95}$$

then

$$\begin{gathered}\lim_{N\to\infty} \mathcal{L}\left[\left(x_1^{[N]}(t), y_1^{[N]}(t)\right)_{t\geq 0}\right] = \mathcal{L}\left[\left(x_1^{\mathbb{N}_0}(t), y_1^{\mathbb{N}_0}(t)\right)_{t\geq 0}\right] \\ \textit{in the Meyer-Zheng topology,}\end{gathered} \tag{6.96}$$

where $(x_1^{\mathbb{N}_0}(t), y_1^{\mathbb{N}_0}(t))_{t\geq 0})$ *evolves according to* (6.55).

To prove Proposition 6.2.18 we will use Lemma 6.2.7 in combination with the following three general lemmas about the Meyer-Zheng topology, which are proven in Appendix B.2.3.

Lemma 6.2.19 (Convergence in probability in the Meyer-Zheng topology). *Let* $((Z_n(t))_{t\geq 0})_{n\in\mathbb{N}}$ *and* $(Z(t))_{t\geq 0}$ *be stochastic processes on the Polish space* (E, d)*. If, for all* $t \geq 0$,

$$\lim_{n\to\infty} \mathbb{E}\left[d(Z_n(t), Z(t))\right] = 0, \tag{6.97}$$

then,

$$\lim_{n\to\infty} (Z_n(t))_{t\geq 0} = (Z(t))_{t\geq 0} \text{ in probability in the Meyer-Zheng topology.} \tag{6.98}$$

Lemma 6.2.20 (Convergence of the joint law). *Let* $((X_n(t))t \geq 0)_{n\in\mathbb{N}}$, $((Y_n(t))_{t\geq 0})_{n\in\mathbb{N}}, (X(t))_[t \geq 0]$ *be stochastic processes on a metric space* (E, d) *and let* $c \in E$ *be a constant. If* $\lim_{n\to\infty} \mathcal{L}[X_n] = \mathcal{L}[X]$ *in the Meyer-Zheng topology and for all* $t \geq 0$, $\lim_{n\to\infty} \mathbb{E}[d(Y_n(t), c)] = 0$*, then* $\lim_{n\to\infty} \mathcal{L}[(X_n, Y_n)] = \mathcal{L}[(X, c)]$ *in the Meyer-Zheng topology.*

Lemma 6.2.21 (Continuous mapping theorem). *Let* $f\colon E \to E$ *be a continuous function and* $x \in M_E[0, \infty)$.

(a) The function

$$h\colon \Psi \to \Psi, \qquad \psi_x \to \psi_{f(x)}, \tag{6.99}$$

is continuous.

(b) If the stochastic processes $(X_n)_{n\in\mathbb{N}}, X$ *on state space* (E, d) *satisfy*

$$\lim_{n\to\infty} \mathcal{L}[X_n] = \mathcal{L}[X] \text{ in the Meyer-Zheng topology,} \tag{6.100}$$

then

$$\lim_{n\to\infty} \mathcal{L}[f(X_n)] = \mathcal{L}[f(X)] \text{ in the Meyer-Zheng topology.} \tag{6.101}$$

Note that Lemma 6.2.21 allows us to use the continuous mapping theorem in the Meyer-Zheng topology.

Step 4. Mean-field finite-systems scheme. Use Steps 1– 4 to prove Proposition 6.2.1.

Having completed the abstract scheme of steps 1–4, we set out to prove the constituent propositions and lemmas.

§6.3 Proofs: $N \to \infty$, mean-field, proof of abstract scheme

In Sections 6.3.1–6.3.4 we prove the propositions and the lemmas stated in Steps 1–4 in Section 6.2.2.

§6.3.1 Proof of step 1. Equilibrium of the single components

We start by proving Proposition 6.2.4 with the help of the seven lemmas stated in Step 1 of Section 6.2.2. Afterwards we prove each of the lemmas.

- **Proof of Proposition 6.2.4**

Proof. We use an argument similar to the one used in [21, Section 2 (i)]. Let $(L(N))_{N\in\mathbb{N}}$ be any sequence satisfying $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$. Let μ_N be the measure on $([0,1]^2)^{\mathbb{N}_0}$ obtained by periodic continuation of $\mathcal{L}[X^{[N]}(Ns-L(N)), Y^{[N]}(Ns-L(N))]$. Note that $([0,1]^2)^{\mathbb{N}_0}$ is compact. Hence, letting $(N_k)_{k\in\mathbb{N}}$ be the subsequence in Proposition 6.2.4, we can pass to a further subsequence and obtain

$$\lim_{k\to\infty} \mu_{N_k} = \mu. \tag{6.102}$$

Since we assumed that $\mathcal{L}[X^{[N]}(0), Y^{[N]}(0)]$ is exchangeable and the dynamics preserves exchangeability, the measures μ_{N_k} are exchangeable and also the limiting law μ is exchangeable. Define ϕ as in (6.78) in Lemma 6.2.10. Then we can condition on ϕ and write

$$\mu = \int_{[0,1]} \mu_\rho \, \mathrm{d}\Lambda(\rho), \tag{6.103}$$

where $\Lambda(\cdot) = \mathcal{L}[\phi]$. By assumption we know that

$$\lim_{k\to\infty} \mathcal{L}\big[\bar\Theta^{[N_k]}(N_k s)\big] = P_s \tag{6.104}$$

and

$$\lim_{k\to\infty} \mathcal{L}\left[\sup_{0\le t\le L(N_k)} \left|\bar\Theta^{[N_k]}(N_k s) - \bar\Theta^{[N_k]}(N_k s - t)\right|\right] = \delta_0. \tag{6.105}$$

Hence

$$\lim_{k\to\infty} \mathcal{L}\left[\bar\Theta^{[N_k]}(N_k s - L(N_k))\right] = P_s. \tag{6.106}$$

Recall that

$$\Lambda = \mathcal{L}[\phi] = \mathcal{L}\left[\lim_{n\to\infty} \frac{1}{n} \sum_{i\in[n]} \frac{x_i + Ky_i}{1+K}\right] \qquad \text{on } (\mu, ([0,1]^2)^{\mathbb{N}_0}). \tag{6.107}$$

By Lemma 6.2.10, if $\phi_{N_k} = \frac{1}{N_k}\sum_{i\in[N_k]} \frac{x_i+Ky_i}{1+K}$ on $(\mu_{N_k}, ([0,1]^2)^{\mathbb{N}_0})$, then $\lim_{k\to\infty} \mathcal{L}[\phi_{N_k}] = \mathcal{L}[\phi]$. Taking the subsequence $(\mu_{N_k})_{k\in\mathbb{N}}$, we get $\Lambda(\cdot) = P_s(\cdot)$, and hence

$$\mu = \int_{[0,1]} \mu_\rho \, \mathrm{d}P_s(\rho). \tag{6.108}$$

Let $\bar{L}(N)$ be the sequence constructed in Lemma 6.2.11[b]. We can require that $\bar{L}(N) \leq L(N)$ for all $N \in \mathbb{N}$. Write

$$\begin{aligned}
&\mathcal{L}\big[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big]\\
&= \mathcal{L}\big[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{N_k}(N_k s - L(N_k) + \bar{L}(N_k))\big]\\
&\quad - \mathcal{L}\big[X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k))\big],\\
&\quad + \mathcal{L}\big[X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k))\big] - \mathcal{L}\big[X^{\mu}(\bar{L}(N_k)), Y^{\mu}(\bar{L}(N_k))\big]\\
&\quad + \mathcal{L}\big[X^{\mu}(\bar{L}(N_k)), Y^{\mu}(\bar{L}(N_k))\big].
\end{aligned} \tag{6.109}$$

By Lemma 6.2.11, the first and the second term tend to zero as $k \to \infty$. Hence

$$\mathcal{L}\big[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big] \tag{6.110}$$

tends to $\mathcal{L}\big[X^\mu(L(N_k)), Y^\mu(L(N_k))\big]$ as $k \to \infty$. By (6.108),

$$\mathcal{L}\big[X^{\mu}(\bar{L}(N_k)), Y^{\mu}(\bar{L}(N_k))\big] = \int_{[0,1]} \mathcal{L}\big[X^{\mu_\rho}(\bar{L}(N_k)), Y^{\mu_\rho}(\bar{L}(N_k))\big] \, \mathrm{d}P_s(\rho). \tag{6.111}$$

Since $\lim_{k\to\infty} \bar{L}(N_k) = \infty$, by Lemma 6.2.5 we have

$$\lim_{k\to\infty} \mathcal{L}\big[X^{\mu_\rho}(\bar{L}(N_k)), Y^{\mu_\rho}(\bar{L}(N_k))\big] = \nu_\rho. \tag{6.112}$$

Therefore, by (6.109) and Lemma 6.2.6,

$$\begin{aligned}
&\lim_{k\to\infty} \mathcal{L}\big[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big]\\
&\quad = \int_{[0,1]} \nu_\rho \, \mathrm{d}P_s(\rho).
\end{aligned} \tag{6.113}$$

To show that

$$\lim_{k\to\infty} \mathcal{L}\big[X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s)\big] = \int_{[0,1]} \nu_\rho \, \mathrm{d}P_s(\rho). \tag{6.114}$$

we invoke Lemma 6.2.13. Let $(X^{[N],1}, Y^{[N],1})$ be the finite system starting from

$$\mathcal{L}\big[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))\big], \tag{6.115}$$

and let $(\bar{L}(N))_{N\in\mathbb{N}}$ be the sequence such that (6.113) holds. Let $(X^{[N],2}.Y^{[N],2})$ be the finite system starting from

$$\mathcal{L}\big[X^{[N]}(Ns-\bar{L}(N)),Y^{[N]}(Ns-\bar{L}(N))\big]. \tag{6.116}$$

Choose for the sequence $t(N)$ in Lemma 6.2.13 the sequence $\bar{L}(N)$. Let $\mu^{[N],1}$ be defined by periodic continuation of $(X^{[N]}(Ns-L(N)),Y^{[N]}(Ns-L(N)))$, and $\mu^{[N],2}$ by periodic continuation of $(X^{[N]}(Ns-\bar{L}(N)),Y^{[N]}(Ns-\bar{L}(N)))$. Defining $\bar{\Theta}_1$ and $\bar{\Theta}_2$ according to (6.85), where for $\mu^{[N],2}$ we replace $L(N)$ by $\bar{L}(N)$, we get

$$\lim_{k\to\infty}|\bar{\Theta}_1^{N_k}-\bar{\Theta}_2^{N_k}|=\lim_{k\to\infty}\big|\bar{\Theta}^{N_k}(N_ks-L(N_k))-\bar{\Theta}^{N_k}(N_ks-\bar{L}(N_k))\big|=0 \text{ in probability} \tag{6.117}$$

by the assumptions in (6.63). Hence, if μ is any weak limit point of the sequence $(\mu^{[N_k],1}\times\mu^{[N_k],2})_{k\in\mathbb{N}}$, then

$$\mu(\bar{\Theta}_1=\bar{\Theta}_2)=1. \tag{6.118}$$

By passing to a further subsequence, we can now apply Lemma 6.2.13, to obtain

$$\lim_{k\to\infty}\mathbb{E}\left[|x_{i,1}^{N_k}(\bar{L}(N_k))-x_{i,2}^{N_k}(\bar{L}(N_k))|+K|y_{i,1}^{N_k}(\bar{L}(N_k))-y_{i,2}^{N_k}(\bar{L}(N_k))|\right]=0. \tag{6.119}$$

Note that

$$\begin{aligned}&\mathcal{L}\big[X_1(\bar{L}(N_k),Y_1(\bar{L}(N_k)))\big]\\&=\mathcal{L}\big[X^{[N_k]}(N_ks-L(N_k)+\bar{L}(N_k)),Y^{[N_k]}(N_ks-L(N_k)+\bar{L}(N_k))\big],\\&\mathcal{L}\big[X_2(\bar{L}(N_k),Y_2(\bar{L}(N_k)))\big]=\mathcal{L}\big[X^{[N_k]}(N_ks),Y^{[N_k]}(N_ks)\big].\end{aligned} \tag{6.120}$$

Moreover, we know from (6.113) that

$$\begin{aligned}&\lim_{k\to\infty}\mathcal{L}\left[X^{[N_k]}(N_ks-L(N_k)+\bar{L}(N_k)),Y^{[N_k]}(N_ks-L(N_k)+\bar{L}(N_k))\right]\\&=\int_{[0,1]}\nu_\rho P_s(\mathrm{d}\rho).\end{aligned} \tag{6.121}$$

Combining (6.119)–(6.121), we find that

$$\mathcal{L}\big[X^{[N_k]}(N_ks),Y^{[N_k]}(N_ks)\big]=\int_{[0,1]}\nu_\rho P_s(\mathrm{d}\rho). \tag{6.122}$$

□

In the remainder of this section we prove Lemmas 6.2.5–6.2.11 and 6.2.13.

• Proof of Lemma 6.2.5

Proof. Since the components of the infinite system in (6.42) evolve independently, it is enough to show that each component converges to Γ_θ. This convergence follows from the proof of Proposition 6.1.2 (see Section 6.1.3). Hence the infinite system defined by (6.18) converges to $\nu_\theta=\Gamma_\theta^{\otimes\mathbb{N}_0}$. Ergodicity of ν_θ with respect to translations follows from Kolmogorov's zero-one law. □

• **Proof of Lemma 6.2.7**

Proof. Using the definition of $\Theta_x^{[N]}(t)$, $\Theta_y^{[N]}(t)$ in (6.44) and the SSDE in (6.15), we find the following evolution for the averages:

$$\begin{aligned} \mathrm{d}\Theta_x^{[N]}(t) &= \frac{1}{N}\sum_{i\in[N]} \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) + Ke\,[\Theta_y^{[N]}(t) - \Theta_x^{[N]}(t)]\,\mathrm{d}t, \\ \mathrm{d}\Theta_y^{[N]}(t) &= e\,[\Theta_x^{[N]}(t) - \Theta_y^{[N]}(t)]\,\mathrm{d}t. \end{aligned} \tag{6.123}$$

Consequently,

$$\begin{aligned} \mathrm{d}\big(\Delta_{\bar\Theta}^{[N]}(t)\big)^2 &= 2\Delta_{\bar\Theta}^{[N]}(t)\,\mathrm{d}\Delta_{\bar\Theta}^{[N]}(t) + 2\mathrm{d}\langle\Delta_{\bar\Theta}^{[N]}\rangle(t) \\ &= \Delta_{\bar\Theta}^{[N]}(t)\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) - (Ke+e)\left(\Delta_{\bar\Theta}^{[N]}(t)\right)^2\mathrm{d}t \\ &\qquad + 2\frac{1}{N^2}\sum_{i\in[N]} g(x_i(t))\;\mathrm{d}t, \end{aligned} \tag{6.124}$$

and hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbb{E}\left[\big(\Delta_{\bar\Theta}^{[N]}(t)\big)^2\right] = -2(Ke+e)\,\mathbb{E}\left[\big(\Delta_{\bar\Theta}^{[N]}(t)\big)^2\right] + \frac{2}{N^2}\sum_{i\in[N]} g(x_i(t)) \tag{6.125}$$

and

$$\mathbb{E}\left[\big(\Delta_{\bar\Theta}^{[N]}(t)\big)^2\right] = \mathbb{E}\left[\big(\Delta_{\bar\Theta}^{[N]}(0)\big)^2\right]\mathrm{e}^{-2(Ke+e)t} + \int_0^t \mathrm{d}r\,\mathrm{e}^{-2(Ke+e)(t-r)}\,\frac{2}{N^2}\sum_{i\in[N]} g(x_i(r)). \tag{6.126}$$

Therefore we get the bound

$$\mathbb{E}\left[\big|\Delta_{\bar\Theta}^{[N]}(t)\big|\right] \leq \sqrt{\mathbb{E}\left[\big(\Delta_{\bar\Theta}^{[N]}(0)\big)^2\right]\mathrm{e}^{-(Ke+e)t}} + \sqrt{\frac{2\|g\|}{N(Ke+e)}}. \tag{6.127}$$

□

• **Proof of Lemma 6.2.9**

Proof. To compare the systems in (6.15) and (6.74), we couple them via their Brownian motions. Therefore for all $i \in [N]$ we assume that the evolution in (6.15) and (6.74) is driven by the same Brownian motion, $\tilde w_i = w_i$. If $i \notin [N]$, then we set $w_i = w_j$ for $j = i \mod N$. We denote the coupled process by $\tilde z(t) = (\tilde z_i(t))_{i\in\mathbb{N}_0} = (\tilde z_i^{[N]}(t), \tilde z_i^{\mu_N}(t))_{i\in\mathbb{N}_0}$, where $\tilde z_i^{[N]}(t) = (\tilde x_i^{[N]}(t), \tilde y_i^{[N]}(t))$ and $\tilde z_i^{\mu_N}(t) = (\tilde x_i^{\mu_N}(t), \tilde y_i^{\mu_N}(t))$. The tilde indicates that we are considering the coupled process, and

$$\begin{aligned} \mathcal{L}[\tilde z(0)] &= \mathcal{L}\big[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))\big] \times \mu_N \\ &= \mathcal{L}\big[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))\big]^2. \end{aligned} \tag{6.128}$$

Define

$$\Delta_i^N(t) = \tilde x_i^{[N]}(t) - \tilde x_i^{\mu_N}(t), \qquad \delta_i^N(t) = \tilde y_i^{[N]}(t) - \tilde y_i^{\mu_N}(t). \tag{6.129}$$

To prove that the coupling is successful, we show that, for all $t \geq 0$,

$$\lim_{N\to\infty} \mathbb{E}\left[|\Delta_i^N(t)| + K|\delta_i^N(t)|\right] = 0 \quad \forall\, i \in \mathbb{N}_0. \tag{6.130}$$

From now on we will only consider sites $i \in [0, N]$ for which both infinite systems have the same Brownian motion.

From (6.15) and (6.74) it follows that

$$\begin{aligned}
&\mathrm{d}\left[|\Delta_i^N(t)| + K|\delta_i^N(t)|\right] \\
&\quad = (\mathrm{sgn}\ \Delta_i^N(t))\ \mathrm{d}\Delta_i^N(t) + \mathrm{d}L_t^0 + K\,\mathrm{sgn}\ \delta_i^N(t)\ \mathrm{d}\delta_i^N(t) \\
&\quad = -c\ (\mathrm{sgn}\ \Delta_i^N(t))\ \Delta_i^N(t)\,\mathrm{d}t + c\,(\mathrm{sgn}\ \Delta_i^N(t))\ \left[\bar{\Theta}^{[N]}(t) - \bar{\Theta}^{[N]}\right]\ \mathrm{d}t \\
&\qquad + c\,(\mathrm{sgn}\ \Delta_i^N(t))\ \left[\Theta_x^{[N]}(t) - \bar{\Theta}^{[N]}(t)\right]\ \mathrm{d}t \\
&\qquad + (\mathrm{sgn}\ \Delta_i^N(t))\ \left(\sqrt{g(x_i^{[N]}(t))} - \sqrt{g(x_i^{\mu_N}(t))}\right)\mathrm{d}w_i(t) \\
&\qquad + (\mathrm{sgn}\ \Delta_i^N(t))\,Ke\,[\delta_i^N(t) - \Delta_i^N(t)]\,\mathrm{d}t \\
&\qquad + (\mathrm{sgn}\ \delta_i^N(t))\,Ke\,[\Delta_i^N(t) - \delta_i^N(t)]\,\mathrm{d}t,
\end{aligned} \tag{6.131}$$

where we use that the local time L_t^0 is zero, since g is Lipschitz (see [63, Proposition V.39.3]).

Taking expectations in (6.131), we find

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] = &-c\,\mathbb{E}\left[|\Delta_i^N(t)|\right] \\
&+ c\ \mathbb{E}\left[(\mathrm{sgn}\ \Delta_i^N(t))\ \left[\bar{\Theta}^{[N]}(t) - \bar{\Theta}^{[N]}\right]\right] \\
&+ c\,\mathbb{E}\left[(\mathrm{sgn}\ \Delta_i^N(t))\ \left[\Theta_x^{[N]}(t) - \bar{\Theta}^{[N]}(t)\right]\right] \\
&+ Ke\,\mathbb{E}\left[\left(\mathrm{sgn}\ \Delta_i^N(t) - \mathrm{sgn}\ \delta_i^N(t)\right)\left[\delta_i^N(t) - \Delta_i^N(t)\right]\right].
\end{aligned} \tag{6.132}$$

Note that we can rewrite (6.132) as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] = &-c\,\mathbb{E}\left[|\Delta_i^N(t)|\right] \\
&- 2Ke\,\mathbb{E}\left[1_{\mathrm{sgn}\ \Delta_i^N(t) \neq \mathrm{sgn}\ \delta_i^N(t)}\ \left[|\delta_i^N(t)| + |\Delta_i^N(t)|\right]\right] \\
&+ c\ \mathbb{E}\left[(\mathrm{sgn}\ \Delta_i^N(t))\ \left[\bar{\Theta}^{[N]}(t) - \bar{\Theta}^{[N]}\right]\right] \\
&+ c\,\mathbb{E}\left[(\mathrm{sgn}\ \Delta_i^N(t))\ \left[\Theta_x^{[N]}(t) - \bar{\Theta}^{[N]}(t)\right]\right].
\end{aligned} \tag{6.133}$$

It therefore follows that

$$\begin{aligned}\mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] = \mathbb{E}[|\Delta_i^N(0)| + K|\delta_i^N(0)|] \\ - c\int_0^t \mathrm{d}r\, \mathbb{E}\left[|\Delta_i^N(r)|\right] \\ - 2Ke\int_0^t \mathrm{d}r\, \mathbb{E}\left[1_{\text{sgn } \Delta_i^N(r) \neq \text{sgn } \delta_i^N(r)}\ [|\delta_i^N(r)| + |\Delta_i^N(r)|]\right] \\ + \int_0^t \mathrm{d}r\, c\,\mathbb{E}\left[(\text{sgn } \Delta_i^N(r))\ \left[\bar{\Theta}^{[N]}(r) - \bar{\Theta}^{[N]}\right]\right] \\ + \int_0^t \mathrm{d}r\, c\,\mathbb{E}\left[(\text{sgn } \Delta_i^N(r))\ \left[\Theta_x^{[N]}(r) - \bar{\Theta}^{[N]}(r)\right]\right].\end{aligned} \tag{6.134}$$

Note that, by the choice of initial distribution for the coupling, we have

$$\mathbb{E}[|\Delta_i^N(0)| + K|\delta_i^N(0)|] = 0. \tag{6.135}$$

Therefore we get

$$\begin{aligned} 0 &\leq \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] \\ &\leq -c\int_0^t \mathrm{d}r\, \mathbb{E}\left[|\Delta_i^N(r)|\right] \\ &\quad - 2Ke\int_0^t \mathrm{d}r\, \mathbb{E}\left[1_{\text{sgn } \Delta_i^N(r) \neq \text{sgn } \delta_i^N(r)}\ [|\delta_i^N(r)| + |\Delta_i^N(r)|]\right] \\ &\quad + \int_0^t \mathrm{d}r\, c\,\mathbb{E}\left[\left|\bar{\Theta}^{[N]}(r) - \bar{\Theta}^{[N]}\right|\right] \\ &\quad + \int_0^t \mathrm{d}r\, c\,\mathbb{E}\left[\left|\Theta_x^{[N]}(r) - \bar{\Theta}^{[N]}(r)\right|\right] \\ &\leq t\left(\sup_{0\leq r\leq t} c\ \mathbb{E}\left[\left|\bar{\Theta}^{[N]}(r) - \bar{\Theta}^{[N]}\right|\right] + c\,\mathbb{E}\left[\left|\Theta_x^{[N]}(r) - \bar{\Theta}^{[N]}(r)\right|\right]\right). \end{aligned} \tag{6.136}$$

Hence, by the assumption in (6.75) and Lemma 6.2.7 (recall (6.128)) , we see that, for all $t > 0$,

$$\lim_{N\to\infty} \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] = 0. \tag{6.137}$$

Therefore, for every Lipschitz function $f \in \mathcal{C}\big(([0,1]),\mathbb{R}\big)$ of $x_i(t)$,

$$\lim_{n\to\infty} \left|\mathbb{E}[f(x_i^{[N]}(t)) - f(x_i^{\mu_N}(t))]\right| \leq \lim_{n\to\infty} \text{Lip} f\, \mathbb{E}[||\Delta_i^N(L(N))||] = 0, \tag{6.138}$$

and the same holds for Lipschitz functions of y_i. Using that the Lipschitz functions are dense in $\mathcal{C}\big(([0,1]),\mathbb{R}\big)$, we obtain that the result actually holds for all $f \in \mathcal{C}\big(([0,1]^2)^{\mathbb{N}_0},\mathbb{R}\big)$ depending on finitely many components. This in turn implies that the result holds for all $f \in \mathcal{C}\big(([0,1]^2)^{\mathbb{N}_0},\mathbb{R}\big)$. □

- **Proof of Lemma 6.2.10**

Proof. Define

$$D^N(Z) = \frac{1}{N}\sum_{j\in[N]} \frac{x_j + Ky_j}{1+K}, \qquad D(Z) = \lim_{N\to\infty} D^N(Z) \text{ in } L_2(\mu). \tag{6.139}$$

Since μ is translation invariant with $\int_{[0,1]^2} \frac{x_0+Ky_0}{1+K}\,\mathrm{d}\mu < 1$, the $L_2(\mu)$-limit $D(Z)$ exists by the ergodic theorem. Since, by assumption, $\mu_N \to \mu$ as $N \to \infty$ for all fixed $M \in \mathbb{N}_0$, we have

$$\lim_{N\to\infty} \mathcal{L}_{\mu_N}[D^M(Z)] = \mathcal{L}_\mu[D^M(Z)]. \tag{6.140}$$

Therefore, in order to prove Lemma 6.2.10, we are left to show

$$\lim_{M\to\infty} \sup_{N\geq M} \|D^M(Z) - D^N(Z)\|_{L_2(\mu_N)} = 0. \tag{6.141}$$

This can be done by using Fourier transforms and spectral densities, and to do so we follow the same strategy as in [21, Lemma 2.5].

Define

$$\bar{\theta}^N = \mathbb{E}^{\mu_N}\left[\frac{x_0+Ky_0}{1+K}\right]. \tag{6.142}$$

Since μ_N is translation invariant on $\mathbb{N}_0$, by Herglotz's theorem there exists a unique measure λ_N such that, for all $j,k \in \mathbb{N}_0$,

$$\mathbb{E}^{\mu_N}\left[\left(\frac{x_j+Ky_j}{1+K} - \bar{\theta}^N\right)\left(\frac{x_k+Ky_k}{1+K} - \bar{\theta}^N\right)\right] = \int_{(-\pi,\pi]} \lambda_N(\mathrm{d}u)\, \mathrm{e}^{\mathrm{i}(j-k)u}. \tag{6.143}$$

For $N \in \mathbb{N}_0$, define

$$D^N(u) = \frac{1}{N}\sum_{j\in[N]} \mathrm{e}^{\mathrm{i}ju}. \tag{6.144}$$

By (6.143), it follows that

$$\left\|D^M(Z) - D^N(Z)\right\|_{L_2(\mu_N)} = \left\|D^M(u) - D^N(u)\right\|_{L_2(\lambda_N)}. \tag{6.145}$$

Polynomials of the type $D^N(u)$ are called trigonometric polynomials and satisfy:

(a) $\lim_{N\to\infty} D^N(u) = 1_{\{0\}}(u)$.

(b) For $\delta > 0$ and $M < \infty$ there exists an $\epsilon(M,\delta)$ such that, for all $N \geq M$,

$$|D^N(u) - 1_{\{0\}}(u)| \leq 1_{(-\delta,\delta)\backslash\{0\}} + \epsilon(M,\delta) \text{ with } \epsilon(M,\delta) \to 0 \text{ as } M \to \infty. \tag{6.146}$$

Hence it follows that

$$\left\|D^M(u) - D^N(u)\right\|^2_{L_2(\lambda_N)} \leq 2\lambda_N((-\delta,\delta)\backslash\{0\}) + 2\epsilon(M,\delta). \tag{6.147}$$

Now let $M \to \infty$, to obtain

$$\sup_{N\geq M} \|D^M(u) - D^N(u)\|_{L_2(\lambda_N)} \leq 2\lambda_N((-\delta,\delta)\backslash\{0\}). \tag{6.148}$$

Subsequently let $\delta \to 0$, so that $(-\delta,\delta)\backslash\{0\} \to \emptyset$ and

$$\lim_{M\to\infty} \sup_{N\geq M} \|D^M(u) - D^N(u)\|^2_{L_2(\lambda_N)} = 0. \tag{6.149}$$

□

- **Proof of Lemma 6.2.11**

Proof. We first prove Lemma 6.2.11[1]. Afterwards we construct $(\bar{L}(N))_{N\in\mathbb{N}}$ to prove Lemma 6.2.11[2].

Since $\lim_{k\to\infty} \mu_{N_k} = \mu$, Lemma 6.2.10 implies that $\lim_{k\to\infty} \mathcal{L}[\bar{\Theta}^{[N_k]}] = \mathcal{L}[\bar{\Theta}]$. For ease of notation we drop the subsequence notation in the remainder of this proof. By Skohorod's theorem we can construct the random variables $(z_i^{\mu_N})_{N\in\mathbb{N}}$ and z_i^{μ} on one probability space such that $\lim_{N\to\infty} z_i^{\mu_N} = z_i$ a.s. Then, as in the proof of Lemma 6.2.10, we obtain

$$\lim_{N\to\infty} \mathbb{E}[|\bar{\Theta}^{[N]} - \bar{\Theta}|] = 0. \tag{6.150}$$

To prove the claim we couple the two infinite systems, namely,

$$\begin{aligned} \mathrm{d}x_i^{\mu_N}(t) &= c\left[\bar{\Theta}^{[N]} - x_i^{\mu_N}(t)\right]\mathrm{d}t + \sqrt{g(x_i^{\mu_N}(t))}\,\mathrm{d}w_i(t) + Ke\,[y_i^{\mu_N}(t) - x_i^{\mu_N}(t)]\,\mathrm{d}t,\\ \mathrm{d}y_i^{\mu_N}(t) &= e\,[x_i^{\mu_N}(t) - y_i^{\mu_N}(t)]\,\mathrm{d}t, \qquad i\in\mathbb{N}_0, \end{aligned} \tag{6.151}$$

and

$$\begin{aligned} \mathrm{d}x_i^{\mu}(t) &= c\left[\bar{\Theta} - x_i^{\mu}(t)\right]\mathrm{d}t + \sqrt{g(x_i^{\mu}(t))}\,\mathrm{d}w_i(t) + Ke\,[y_i^{\mu}(t) - x_i^{\mu}(t)]\,\mathrm{d}t,\\ \mathrm{d}y_i^{\mu}(t) &= e\,[x_i^{\mu}(t) - y_i^{\mu}(t)]\,\mathrm{d}t, \qquad i\in\mathbb{N}_0, \end{aligned} \tag{6.152}$$

are coupled by using the same Brownian motions in (6.151) and (6.152). Like before, define $\Delta_i^{\mu_N} = x_i^{\mu_N} - x_i^{\mu}$ and $\delta_i^{\mu_N} = y_i^{\mu_N} - y_i^{\mu}$. By the construction with Skohorod's theorem, we have that

$$\lim_{N\to\infty} \mathbb{E}\big[|\Delta_i^{\mu_N}(0)| + K|\delta_i^{\mu_N}(0)|\big] = 0. \tag{6.153}$$

To prove that, for all $t \geq 0$,

$$\lim_{N\to\infty} \mathbb{E}\big[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|\big] = 0, \tag{6.154}$$

we proceed as in the proof of Lemma 6.2.9. By Itô-calculus, we find that

$$\begin{aligned} &\mathbb{E}\big[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|\big]\\ &= \mathbb{E}\big[|\Delta_i^{\mu_N}(0)| + K|\delta_i^{\mu_N}(0)|\big]\\ &\quad - c\int_0^t \mathrm{d}r\,\mathbb{E}\left[|\Delta_i^{\mu_N}(r)|\right]\\ &\quad - 2Ke\int_0^t \mathrm{d}r\,\mathbb{E}\left[1_{\mathrm{sgn}\ \Delta_i^{\mu_N}(r)\neq \mathrm{sgn}\ \delta_i^{\mu_N}(r)}\ [|\delta_i^{\mu_N}(r)| + |\Delta_i^{\mu_N}(r)|]\right]\\ &\quad + \int_0^t \mathrm{d}r\, c\,\mathbb{E}\left[\left|\bar{\Theta}^{[N]} - \bar{\Theta}\right|\right]\\ &\leq \mathbb{E}\big[|\Delta_i^{\mu_N}(0)| + K|\delta_i^{\mu_N}(0)|\big] + tc\,\mathbb{E}\big[\big|\bar{\Theta}^{[N]} - \bar{\Theta}\big|\big]. \end{aligned} \tag{6.155}$$

From (6.155) it the follows that, for every $t \geq 0$,

$$\lim_{N\to\infty} \mathbb{E}\big[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|\big] = 0. \tag{6.156}$$

Chapter 6

We next construct the sequence $(\bar{L}(N))_{N\in\mathbb{N}}$. From (6.137) and (6.156),we have

$$\lim_{N\to\infty} \mathbb{E}\big[|\Delta_i^N(t)| + K|\delta_i^N(t)|\big] + \mathbb{E}\big[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|\big] = 0. \tag{6.157}$$

Let $(t_k)_{k\in\mathbb{N}}$ be an increasing sequence such that $\lim_{k\to\infty} t_k = \infty$ and $\lim_{k\to\infty} t_k/k = 0$. By (6.157), we can for each k find an $N_k \in \mathbb{N}$ such that, for all $N \geq N_k$,

$$\mathbb{E}\big[|\Delta_i^N(t_k)| + K|\delta_i^N(t_k)|\big] + \mathbb{E}\big[|\Delta_i^{\mu_N}(t_k)| + K|\delta_i^{\mu_N}(t_k)|\big] < \frac{1}{k}. \tag{6.158}$$

Requiring that $N_{k+1} > N_k$, we obtain a strictly increasing sequence $(N_k)_{k\in\mathbb{N}}$ that partitions $\mathbb{N}$. Set

$$\bar{L}(N) = \sum_{k\in\mathbb{N}} t_k 1_{\{N_k,\cdots,N_{k+1}-1\}}(N). \tag{6.159}$$

We show that $\bar{L}(N)$ satisfies the required properties:

- $\lim_{N\to\infty} \mathbb{E}[|\Delta_i^N(\bar{L}(N))|+K|\delta_i^N(\bar{L}(N))|]+\mathbb{E}[|\Delta_i^{\mu_N}(\bar{L}(N))|+K|\delta_i^{\mu_N}(\bar{L}(N))|] = 0$. To proof this, we fix $\epsilon > 0$ and let K be such that $\frac{1}{K} < \epsilon$. Then, for all $N \geq N_K$,

$$\begin{aligned}
&\mathbb{E}\big[|\Delta_i^N(\bar{L}(N))| + K|\delta_i^N(\bar{L}(N))|\big] + \mathbb{E}\big[|\Delta_i^{\mu_N}(\bar{L}(N))| + K|\delta_i^{\mu_N}(\bar{L}(N))|\big] \\
&\quad = \sum_{k\in\mathbb{N}} \mathbb{E}\big[|\Delta_i^N(t_k)| + K|\delta_i^N(t_k)|\big] + \mathbb{E}\big[|\Delta_i^{\mu_N}(t_k)| + K|\delta_i^{\mu_N}(t_k)|\big]\, 1_{\{N_k,\cdots,N_{k-1}\}}(N) \\
&\quad < \frac{1}{K} < \epsilon.
\end{aligned} \tag{6.160}$$

  We conclude that

$$\lim_{N\to\infty} \mathbb{E}\big[|\Delta_i^N(\bar{L}(N))| + K|\delta_i^N(\bar{L}(N))|\big] + \mathbb{E}\big[|\Delta_i^{\mu_N}(\bar{L}(N))|v + K|\delta_i^{\mu_N}(\bar{L}(N))|\big] = 0. \tag{6.161}$$

- $\lim_{N\to\infty} \bar{L}(N) = \infty$. By (6.159), for each $k \in \mathbb{N}$ there exists an $N_k \in \mathbb{N}$ such that, for all $N \geq N_k$, $\bar{L}(N) \geq t_k$ and $t_k \to \infty$. We conclude that $\lim_{N\to\infty} \bar{L}(N) = \infty$.

- $\lim_{N\to\infty} \bar{L}(N)/N = 0$. Recall that $\lim_{k\to\infty} t_k/k = 0$ and $N_k \geq k$ by construction. Hence $\lim_{N\to\infty} \bar{L}(N)/N \leq \lim_{N\to\infty} \sum_{k\in\mathbb{N}}(t_k/k)]\, 1_{\{N_k,\cdots,N_{k-1}\}}(N) = 0$.

Choosing $(t_k)_{k\in\mathbb{N}} = (L(N))_{N\in\mathbb{N}}$, we complete the proof of Lemma 6.2.11. □

- **Proof of Lemma 6.2.6**

Proof. The goal is to prove that ν_θ is continuous in θ. To do so, let $(\theta_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1]$ (note that θ_n is not a random variable) such that $\lim_{n\to\infty} \theta_n = \theta$. Couple the two infinite systems

$$\begin{aligned}
&\mathrm{d}x_i^n(t) = c\,[\theta_n - x_i^n(t)]\,\mathrm{d}t + \sqrt{g(x_i^n(t))}\,\mathrm{d}w_i(t) + Ke\,[y_i^n(t) - x_i^n(t)]\,\mathrm{d}t, \\
&\mathrm{d}y_i^n(t) = e\,[x_i^n(t) - y_i^n(t)]\,\mathrm{d}t, \qquad i \in \mathbb{N}_0,
\end{aligned} \tag{6.162}$$

and

$$\begin{aligned} \mathrm{d}x_i(t) &= c\,[\theta - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) + Ke\,[y_i(t) - x_i(t)]\,\mathrm{d}t, \\ \mathrm{d}y_i(t) &= e\,[x_i(t) - y_i(t)]\,\mathrm{d}t, \qquad i \in \mathbb{N}_0, \end{aligned} \tag{6.163}$$

via their Brownian motions, like in the proof of Lemma 6.2.11. Let $\mathcal{L}[(x_i^n(0), y_i^n(0))] = \delta_{(\theta_n,\theta_n)}$ and $\mathcal{L}[(x_i(0), y_i(0))] = \delta_{(\theta,\theta)}$. As before, define $\Delta_i^n = x_i^n - x_i$ and $\delta_i^n = y_i^n - y_i$. Note that

$$\lim_{n\to\infty} \mathbb{E}\big[|\Delta_i^n(0)| + K|\delta_i^n(0)|\big] = 0. \tag{6.164}$$

By a similar argument as in the proof of Lemma 6.2.11, we obtain that, for all $t \geq 0$,

$$\lim_{n\to\infty} \mathbb{E}\big[|\Delta_i^n(t)| + K|\delta_i^n(t)|\big] = 0. \tag{6.165}$$

Hence we can construct a sequence $(L(n))_{n\in\mathbb{N}}$ satisfying $\lim_{n\to\infty} L(n) = \infty$ and $\lim_{n\to\infty} L(n)/n = 0$ such that

$$\lim_{n\to\infty} \mathbb{E}\big[|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|\big] = 0. \tag{6.166}$$

To prove the continuity of the equilibrium ν_θ in θ, we reason as follows. First note that we can couple the system in (6.162) starting from $\delta_{(\theta_n,\theta_n)}$ with the system in (6.162) starting from ν_{θ_n}. By the uniqueness and convergence to equilibrium (see Lemma 6.2.5), we see that this coupling is successful. Similarly, we can couple the system in (6.163) starting from $\delta_{(\theta,\theta)}$ with the system in (6.163) starting from ν_{θ_n}, and see the coupling succesful. Finally, we use (6.166) to obtain

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{E}^{\nu_{\theta_n}\times\nu_\theta}\big[|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|\big] \\ &\leq \lim_{n\to\infty} \mathbb{E}^{\nu_{\theta_n}\times\delta_{(\theta_n,\theta_n)}}\big[|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|\big] \\ &\quad + \lim_{n\to\infty} \mathbb{E}^{\delta_{(\theta,\theta)}\times\delta_{(\theta_n,\theta_n)}}\big[|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|\big] \\ &\quad + \lim_{n\to\infty} \mathbb{E}^{\nu_\theta\times\delta_{(\theta,\theta)}}\big[|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|\big] = 0. \end{aligned} \tag{6.167}$$

Let f be a Lipschitz function. Then, by the equilibrium property of ν_{θ_n} and ν_θ,

$$\begin{aligned} &\lim_{n\to\infty} \big|\mathbb{E}^{\nu_{\theta_n}}\big[f(x^n(0))\big] - \mathbb{E}^{\nu_\theta}\big[f(x(0))\big]\big| \\ &= \lim_{n\to\infty} \big|\mathbb{E}^{\nu_{\theta_n}}\big[f(x^n(L(n)))\big] - \mathbb{E}^{\nu_\theta}\big[f(x(L(n)))\big]\big| \\ &= \lim_{n\to\infty} \big|\mathbb{E}^{\nu_{\theta_n}\times\nu_\theta}\big[f(x^n(L(n))) - f(x(L(n)))\big]\big| \\ &\leq \lim_{n\to\infty} (\mathrm{Lip}f)\, \mathbb{E}^{\nu_{\theta_n}\times\nu_\theta}\big[|(x^n(L(n))) - (x(L(n)))|\big] = 0. \end{aligned} \tag{6.168}$$

We can also show this if f is a Lipschitz function of the y component. Hence ν_θ is indeed continuous as a function of θ. □

- **Proof of Lemma 6.2.13**

Proof. Note that for all $N \in \mathbb{N}$ fixed, by Itô-calculus we find from (6.15) that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[|x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K|y_i^{[N],1}(t) - y_i^{[N],2}(t)|\right] \\
&= -\frac{2c}{N}\sum_{j\in[N]}\mathbb{E}\left[|x_{j,1}^{[N]}(t) - x_{j,2}^{[N]}(t)|\,1_{\left\{\operatorname{sgn}\left(x_{j,1}^{[N]}(t)-x_{j,2}^{[N]}(t)\right)\neq\operatorname{sgn}\left(x_i^{[N],1}(t)-x_i^{[N],2}(t)\right)\right\}}\right] \\
&\quad - 2Ke\,\mathbb{E}\Big[|x_i^{[N],1}(t) - x_i^{[N],2}(t)| \\
&\qquad + |y_i^{[N],1}(t) - y_i^{[N],2}(t)|\,1_{\left\{\operatorname{sgn}\left(x_i^{[N],1}(t)-x_i^{[N],2}(t)\right)\neq\operatorname{sgn}\left(y_i^{[N],1}(t)-y_i^{[N],2}(t)\right)\right\}}\Big].
\end{aligned} \tag{6.169}$$

Therefore, for each $N \in \mathbb{N}$,

$$t \mapsto \mathbb{E}\left[|x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K|y_i^{[N],1}(t) - y_i^{[N],2}(t)|\right] \text{ is decreasing.} \tag{6.170}$$

Fix any $t(N) \to \infty$. By the assumption in Lemma 6.2.13, the proofs of Lemma 6.2.9 and Lemma 6.2.11 imply that (6.157) holds for both $(X^{[N],1}, Y^{[N],1})$ and $(X^{[N],2}, Y^{[N],2})$. Using the construction in the proof of Lemma 6.2.11, we can construct *one* sequence $(l(N))_{N\in\mathbb{N}}$, satisfying $l(N) \le t(N)$, $\lim_{N\to\infty} l(N) = \infty$ and $\lim_{N\to\infty} l(N)/N = 0$, such that (6.161) with $\bar{L}(N)$ replaced by $l(N)$ holds for both the systems arising from $(X^{[N],1}, Y^{[N],1})$ and $(X^{[N],2}, Y^{[N],2})$.

Write

$$\begin{aligned}
&\mathbb{E}\left[|x_i^{[N],1}(l(N)) - x_i^{[N],2}(l(N))| + K|y_i^{[N],1}(l(N)) - y_i^{[N],2}(l(N))|\right] \\
\le&\mathbb{E}\left[|x_i^{[N],1}(l(N)) - x_i^{\mu,1}(l(N))| + K|y_i^{[N],1}(l(N)) - y_i^{\mu,1}(l(N))|\right] \\
&+\mathbb{E}\left[|x_i^{\mu,1}(l(N)) - x_i^{\mu,2}(l(N))| + K|y_i^{\mu,1}(l(N)) - y_i^{\mu,2}(l(N))|\right] \\
&+\mathbb{E}\left[|x_i^{\mu,2}(l(N)) - x_i^{[N],2}(l(N))| + K|y_i^{\mu,2}(l(N)) - y_i^{[N],2}(l(N))|\right].
\end{aligned} \tag{6.171}$$

Note that in the right-hand side of the inequality the first and the third term tend to zero by (6.161). The second term tends to zero because $\mu\{\bar{\Theta}_1 = \bar{\Theta}_2\} = 1$, and hence Lemma 6.2.5 can be applied. Therefore

$$\lim_{N\to\infty}\mathbb{E}\left[|x_i^{[N],1}(l(N)) - x_i^{[N],2}(l(N))| + K|y_i^{[N],1}(l(N)) - y_i^{[N],2}(l(N))|\right] = 0. \tag{6.172}$$

Using the monotonicity in (6.170), we get

$$\begin{aligned}
&\lim_{N\to\infty}\mathbb{E}\left[|x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K|y_i^{[N],1}(t(N)) - y_i^{[N],2}(t(N))|\right] \\
&\le \lim_{N\to\infty}\mathbb{E}\left[|x_i^{[N],1}(l(N)) - x_i^{[N],2}(l(N))| + K|y_i^{[N],1}(l(N)) - y_i^{[N],2}(l(N))|\right] = 0.
\end{aligned} \tag{6.173}$$

□

Combining the proofs of Proposition 6.2.4, Lemma 6.2.11 and Lemma 6.2.13, we obtain the following corollary. This corollary turns out to be important in Section 6.3.2 in the proof of Lemma 6.2.16 to obtain the limiting evolution of the 1-blocks.

Corollary 6.3.1. *Fix $s > 0$. Let μ_N be the measure obtained by periodic configuration of*

$$\mathcal{L}\big[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))\big], \tag{6.174}$$

and let μ be a weak limit point of the sequence $(\mu_N)_{N\in\mathbb{N}}$. Let

$$\bar{\Theta} = \lim_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} \frac{x_i^\mu + K y_i^\mu}{1+K} \text{ in } L^2(\mu). \tag{6.175}$$

and let $(X^{\nu_{\bar{\Theta}}}, Y^{\nu_{\bar{\Theta}}})$ be the infinite system evolving according to (6.82) and starting from its equilibrium measure. Consider the finite system $(X^{[N]}, Y^{[N]})$ as a system on $([0,1]\times[0,1))^{\mathbb{N}_0}$ obtained by periodic continuation. Construct $(X^{[N]}, Y^{[N]})$ and $(X^{\nu_{\bar{\Theta}}}, Y^{\nu_{\bar{\Theta}}})$ on one probability space. Then, for all $t \geq 0$,

$$\lim_{N\to\infty} \mathbb{E}\left[\left|x_i^{[N]}(Ns+t) - x_i^{\nu_{\bar{\Theta}}}(t)\right|\right] + K\,\mathbb{E}\left[\left|y_i^{[N]}(Ns+t) - y_i^{\nu_{\bar{\Theta}}}(t)\right|\right] = 0, \qquad \forall\, i \in [N]. \tag{6.176}$$

Proof. By Proposition 6.2.4 we have that $\lim_{k\to\infty} \mathcal{L}[X^{[N_k]}(N_k s) + Y^{[N_k]}(N_k s)] = \nu(s) = \nu_{\bar{\Theta}}$. Let ν_{N_k} be defined by periodic continuation of the configuration of $(X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s))$ and let $\nu = \lim_{k\to\infty} \nu_{N_k}$, so $\nu = \nu_{\bar{\Theta}}$. Construct the process $(X^{[N]}(t), Y^{[N]}(t))_{t\geq 0}$, $(X^{\nu_{N_k}}(t), Y^{\nu_{N_k}}(t))_{t\geq 0}$ and $(X^{\nu}(t), Y^{\nu}(t))_{t\geq 0}$ on one probability space and use for all processes the same Brownian motions. Then the couplings in the proofs of Lemma 6.2.9 and Lemma 6.2.11 imply (6.176). □

§6.3.2 Proof of step 2. Convergence of the estimator

In this section we prove the three lemmas stated in Step 2 of Section 6.2.2. Afterwards we prove Proposition 6.2.14 with the help of these lemmas.

• Proof of Lemma 6.2.15

Proof. Recall the definition of $\bar{\Theta}^{[N]}(t)$ in (6.43). It follows from the SSDE in (6.15) that

$$\mathrm{d}\bar{\Theta}^{[N]}(t) = \frac{1}{1+K}\frac{1}{N} \sum_{i\in[N]} \sqrt{g\big(x_i^{[N]}(t)\big)}\,\mathrm{d}w_i(t). \tag{6.177}$$

Hence we see that $t \mapsto \bar{\Theta}^{[N]}(t)$ is a continuous martingale. By Itô's formula we have

$$\begin{aligned} \mathbb{E}\big[\big(\bar{\Theta}^{[N]}(t)\big)^2\big] &= \mathbb{E}\big[\big(\bar{\Theta}^{[N]}(0)\big)^2\big] + \frac{1}{(1+K)^2}\int_0^t \mathrm{d}r\, \frac{1}{N^2} \sum_{i\in[N]} g\big(x_i^{[N]}(r)\big) \\ &\leq 1 + \frac{1}{N}\frac{\|g\|}{(1+K)^2}\, t. \end{aligned} \tag{6.178}$$

Since g is a bounded function, we get that $t \mapsto \bar{\Theta}^{[N]}(t)$ is square integrable. It follows that,

$$\Big(\big(\bar{\Theta}^{[N]}(Ns+t) - \bar{\Theta}^{[N]}(Ns)\big)^2\Big)_{t\geq 0} \tag{6.179}$$

is a sub-martingale. Therefore, defining the stopping time

$$S^N_\epsilon = \inf\left\{t \geq 0 : \left(\bar\Theta^{[N]}(Ns+t) - \bar\Theta^{[N]}(Ns)\right)^2 \geq \epsilon\right\} \wedge L(N), \tag{6.180}$$

we find, by the continuity of $t \mapsto \bar\Theta^{[N]}(Ns+t)$ and the optional sampling theorem, that

$$\mathbb{P}\left(S^N_\epsilon \in [Ns, Ns+L(N))\right) \leq \frac{1}{\epsilon^2}\mathbb{E}\left[\left(\bar\Theta^{[N]}(Ns+L(N)) - \bar\Theta^{[N]}(Ns)\right)^2\right]. \tag{6.181}$$

Combining (6.178) and (6.181), we find

$$\lim_{N\to\infty} \sup_{0\leq t\leq L(N)} \left|\bar\Theta^{[N]}(Ns+t) - \bar\Theta^{[N]}(Ns)\right| = 0 \text{ in probability.} \tag{6.182}$$

To obtain the increasing process, note that by Itô-calculus it follows from (6.177) that

$$\left\langle\bar\Theta^{[N]}(Ns)\right\rangle_{s\geq 0} = \frac{1}{(1+K)^2}\int_0^s \mathrm{d}r\, \frac{1}{N}\sum_{i\in[N]} g\left(x_i^{[N]}(Nr)\right). \tag{6.183}$$

We are left to show that the sequence of processes $((\bar\Theta^{[N]}(Ns))_{s\geq 0})_{N\in\mathbb{N}}$ is tight. Note that (6.183) implies that, for all $N \in \mathbb{N}$ and $s \geq 0$,

$$\left\langle\bar\Theta^{[N]}(Ns)\right\rangle \leq \frac{\|g\|}{(1+K)^2}\, s \tag{6.184}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}s}\left\langle\bar\Theta^{[N]}(Ns)\right\rangle \leq \frac{\|g\|}{(1+K)^2}. \tag{6.185}$$

Hence the sequence $(\langle\bar\Theta^{[N]}(Ns)\rangle_{s\geq 0})_{N\in\mathbb{N}}$ is equicontinuous. Therefore, by the Arzela-Ascoli theorem (see e.g. [7, Theorem 7.2]), for each $T \geq 0$ the set $(\langle\bar\Theta^{[N]}(Ns)\rangle_{0\leq s\leq T})_{N\in\mathbb{N}}$ is relatively compact in $C([0,T],\mathbb{R})$, the space of continuous functions from $[0,T] \to \mathbb{R}$. Therefore the set of laws $(\mathcal{L}\langle\bar\Theta^{[N]}(Ns)\rangle_{0\leq s\leq T})_{N\in\mathbb{N}}$ is tight in $\mathcal{P}(C([0,T],\mathbb{R}))$. Hence it follows that $(\mathcal{L}\langle\bar\Theta^{[N]}(Ns)\rangle_{s\geq 0})_{N\in\mathbb{N}}$ is tight in $\mathcal{P}(C([0,\infty),\mathbb{R}))$, the set of probability laws on $C([0,\infty),\mathbb{R})$.

Since $(\bar\Theta^{[N]}(Ns) - \bar\Theta^{[N]}(0))_{s\geq 0}$ is a stochastic integral, we can represent it as a time-transformed Brownian motion (see e.g. [62, Chapter 5]):

$$\left(\bar\Theta^{[N]}(Ns) - \bar\Theta^{[N]}(0)\right)_{s\geq 0} = w\left(\langle\bar\Theta^{[N]}(Ns)\rangle\right)_{s\geq 0}. \tag{6.186}$$

Let χ be a standard normal random variable. Then

$$\begin{aligned}\mathbb{E}\left[\left(w\left(\langle\bar\Theta^{[N]}(Ns)\rangle\right) - w\left(\langle\bar\Theta^{[N]}(Nr)\rangle\right)\right)^2\right] &\leq \mathbb{E}\left[\left(\langle\bar\Theta^{[N]}(Ns)\rangle - \langle\bar\Theta^{[N]}(Nr)\rangle\right)^2\right]\mathbb{E}\left[\chi^4\right]\\ &\leq (s-r)^2\,\frac{\|g\|^2}{(1+K)^4}\,\mathbb{E}\left[\chi^4\right].\end{aligned} \tag{6.187}$$

Hence it follows from Kolmogorov's criterion for weak compactness (see e.g. [62, Chapter XIII, Theorem 1.8]) that the sequence $(\mathcal{L}[(\bar\Theta^{[N]}(Ns))_{s\geq 0}])_{N\in\mathbb{N}}$ is tight in $\mathcal{P}(C([0,\infty),\mathbb{R}))$. □

• **Proof of Lemma 6.2.16**

Proof. For ease of notation we will suppress the subsequence notation and assume that

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\bar\Theta^{[N]}(Ns)\right)_{s>0}\right] = \mathcal{L}\big[(\bar\Theta(s))_{s>0}\big]. \tag{6.188}$$

The processes $(\bar\Theta^{[N]}(Ns))_{s\geq 0}$ are martingales, see (6.177), measurable w.r.t. the canonical filtration $(\mathcal{F}_s)_{s\geq 0}$ and so are there weak limit points. Therefore also the weak limit point $(\bar\Theta(s))_{s>0}$ is a martingale, (see [21, Section 3]). To obtain (6.93), we use the following strategy. Recall from the proof of Lemma 6.2.15 that the sequence $\{\langle\bar\Theta^{[N]}(Ns)\rangle_{s>0}\}_{N\in\mathbb{N}}$ is tight. Hence, in order to prove that

$$\lim_{N\to\infty} \mathcal{L}\left[\langle\bar\Theta^{[N]}(Ns)\rangle_{s>0}\right] = \mathcal{L}\left[\left(\int_0^s \mathrm{d}r \frac{1}{(1+K)^2}\,\mathbb{E}^{\nu_{\bar\Theta(r)}}[g(x_0)]\right)_{s>0}\right], \tag{6.189}$$

it is enough to prove that the finite-dimensional distributions of $(\langle\bar\Theta^{[N]}(Ns)\rangle_{s>0})_{N\in\mathbb{N}}$ converge to the finite-dimensional distribution of $(\int_0^s \mathrm{d}r\,\frac{1}{(1+K)^2}\mathbb{E}^{\nu_{\bar\Theta(r)}}[g(x_0)])_{s>0}$. We will prove a slightly stronger result, namely,

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\langle\bar\Theta^{[N]}(Ns)\rangle - \int_0^s \mathrm{d}r\,\frac{1}{(1+K)^2}\,\mathbb{E}^{\nu_{\bar\Theta(r)}}[g(x_0)]\right|\right] = 0. \tag{6.190}$$

Once we have obtained (6.190) and hence (6.189), by Skohorod's theorem we can construct the processes $(\langle\bar\Theta^{[N]}(Ns)\rangle_{s>0})_{N\in\mathbb{N}}$ on a single probability space, to obtain

$$\lim_{N\to\infty} \langle\bar\Theta^{[N]}(Ns)\rangle_{s\geq 0} = \left(\int_0^s \mathrm{d}r\,\frac{1}{(1+K)^2}\mathbb{E}^{\nu_{\bar\Theta(r)}}[g(x_0)]\right)_{s\geq 0} \quad a.s. \tag{6.191}$$

Using the continuity of Brownian motion, we get that (recall (6.186))

$$\begin{aligned}\lim_{N\to\infty} \left(\bar\Theta^{[N]}(Ns)\right)_{s>0} &= \lim_{N\to\infty}\left[w\left(\langle\bar\Theta^{[N]}(Ns)\rangle\right)_{s>0} + \bar\Theta^{[N]}(0)\right]\\ &= w\left(\int_0^s \mathrm{d}r\,\frac{1}{(1+K)^2}\,\mathbb{E}^{\nu_{\bar\Theta(r)}}[g(x_0)]\right)_{s>0} + \vartheta_0 \quad a.s.\end{aligned} \tag{6.192}$$

Therefore we can choose a version of $(\bar\Theta(s))_{s>0}$ such that

$$\lim_{N\to\infty} \left(\bar\Theta^{[N]}(Ns)\right)_{s>0} = \lim_{N\to\infty}\left(\bar\Theta(s)\right)_{s>0} \quad a.s. \tag{6.193}$$

and

$$\lim_{N\to\infty} \left(\bar\Theta^{[N]}(Ns), \langle\bar\Theta^{[N]}(Ns)\rangle\right)_{s>0} = \left(\bar\Theta(s), \langle\bar\Theta(s)\rangle\right)_{s>0} \quad a.s. \tag{6.194}$$

By the continuous mapping theorem, (6.93) follows. The martingale property follows from the fact that $\left(\bar\Theta^{[N]}(Ns)^2 - \langle\bar\Theta^{[N]}(Ns)\rangle\right)_{s>0}$ are martingales. Therefore, to finish the proof of Lemma 6.2.16 we are left to prove (6.190).

To prove (6.190), define the empirical measures on $[0,1]$ by

$$U^{[N]}(Ns) = \frac{1}{N}\sum_{i\in[N]} \delta_{x_i(Ns)}. \tag{6.195}$$

Note that we can write

$$\begin{aligned}
&\mathbb{E}\left[\left|\langle\bar{\Theta}^{[N]}(Ns)\rangle - \int_0^s \mathrm{d}r\,\frac{1}{(1+K)^2}\,\mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right|\right]\\
&= \frac{1}{(1+K)^2}\,\mathbb{E}\left[\left|\int_0^s \mathrm{d}r\;\mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \int_0^s \mathrm{d}r\;\mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right|\right]\\
&\le \frac{1}{(1+K)^2}\int_0^s \mathrm{d}r\,\mathbb{E}\left[\left|\mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right|\right].
\end{aligned} \tag{6.196}$$

Hence, to prove (6.190) it is enough to prove that, for all $r > 0$,

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right|\right] = 0 \tag{6.197}$$

and apply the dominated convergence theorem.

To prove (6.197), we will use the coupling arguments from Section 6.3.1, as well as ergodicity and invariance under the evolution of $\nu_{\bar{\Theta}}$. As before, let $z^{[N]}(t)$ denote the $[N]$-component system $(x_i^{[N]}(t), y_i^{[N]}(t))_{t\ge 0}$ evolving according to (6.15), viewed as a system on $\mathbb{N}_0$ obtained by periodic continuation and with initial law $\mathcal{L}[z^{[N]}(0)] = \mathcal{L}[x_i^{[N]}(Nr - L(N)), y_i^{[N]}(Nr - L(N))]$. Let $(z^{\mu_N}(t))_{t\ge 0}$ denote the infinite system $(x_i^{\mu_N}(t), y_i^{\mu_N}(t))_{t\ge 0}$ evolving according to (6.74) with initial law μ_N obtained by periodic continuation of the configuration of $(x_i^{[N]}(Nr-L(N)), y_i^{[N]}(Nr-L(N)))$, and let μ be a weak limit point of the sequence μ_N. Note that, for all $r > 0$, by Lemma 6.2.10 we have that $\lim_{N\to\infty} \bar{\Theta}^{[N]}(Nr) = \bar{\Theta}(r)$ for $\bar{\Theta}(r) = \lim_{n\to\infty} \frac{1}{n}\sum_{i\in[n]} \frac{x_i+Ky_i}{1+K}$ in $L_2(\mu)$. Let $\bar{L}(N)$ be the sequence constructed in Corollary 6.3.1. Then we can write

$$\begin{aligned}
&\mathbb{E}\left[\left|\mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \mathbb{E}^{\nu_{\bar{\Theta}}(r)}[g(x_0)]\right|\right]\\
&\le \mathbb{E}\left[\left|\frac{1}{N}\sum_{i\in[N]} g(x_i^{[N]}(Nr)) - \frac{1}{N}\sum_{i\in[N]} g\big(x_i^{\nu_{\bar{\Theta}}(r)}(\bar{L}(N))\big)\right|\right]\\
&\qquad + \mathbb{E}\left[\left|\frac{1}{N}\sum_{i\in[N]} g\big(x_i^{\nu_{\bar{\Theta}(r)}}(\bar{L}(N))\big) - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right|\right]\\
&\le (\mathrm{Lip}\,g)\,\mathbb{E}\left[\left|x_0^{[N]}(Nr) - x_0^{\nu_{\bar{\Theta}}(r)}(\bar{L}(N))\right|\right]\\
&\qquad + \mathbb{E}\left[\left|\frac{1}{N}\sum_{i\in[N]} g\big(x_i^{\nu_{\bar{\Theta}(r)}}(\bar{L}(N))\big) - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)]\right|\right],
\end{aligned} \tag{6.198}$$

where in the second inequality we use the Lipschitz property of g and the translation invariance of the system. Note that the first term tends to 0 as $N \to \infty$ by Corollary 6.3.1. Finally, note that by Lemma 6.2.5 $(x_i)_{i\in\mathbb{N}_0}$ is a sequence of bounded i.i.d. random variables under $\nu_{\bar{\Theta}(r)}$. Hence the last term tends to zero by the law of large numbers. □

- **Proof of Lemma 6.2.17**

Proof. Note that, since g is Lipschitz, the function

$$\theta \mapsto \mathbb{E}^{\nu_\theta}[g], \tag{6.199}$$

is Lipschitz by Lemma 6.2.6. Hence, by [72, Theorem 1], the SDE

$$\mathrm{d}\Phi(s) = \frac{1}{(1+K)} \sqrt{\mathbb{E}^{\nu_{\Phi(s)}}[g]}\, \mathrm{d}w(s) \tag{6.200}$$

has a pathwise unique solution (see įte[p315–317]KS91) and a unique solution in law (see [72, Proposition 1]). This implies that the martingale problem with generator

$$\frac{1}{(1+K)^2}\, \mathbb{E}^{\nu_\Phi}[g] \frac{\mathrm{d}}{\mathrm{d}\Phi^2} \tag{6.201}$$

has a unique solution. In particular, choosing $f(\Phi) = \Phi^2$, we see that the martingale problem implies that

$$\left(\Phi^2(s) - \frac{1}{(1+K)^2} \int_0^s \mathrm{d}u\, \mathbb{E}^{\nu_{\Phi(u)}}[g(x_0)]\right)_{s>0} \tag{6.202}$$

is a martingale.

Since $(\bar{\Theta}(s))_{s>0}$ is a continuous bounded martingale, it could be written as a time transformed Brownian motion. The uniqueness of the martingale problem in (6.94) now follows from the fact that the quadratic variation of a martingale is unique. □

- **Proof of Proposition 6.2.14**

Proof. Combining Lemma 6.2.16–6.2.17, all converging subsequences of $(\mathcal{L}[(\bar{\Theta}^{[N]}(Ns))_{s\geq 0}])_{N\in\mathbb{N}}$ converge to the same limit, which is the unique process satisfying the martingale problem in (6.94). □

§6.3.3 Proof of step 3. Convergence of the 1-blocks in the Meyer-Zheng topology

In this section we prove Proposition 6.2.18 stated in Step 3 of Section 6.2.2. The Lemmas 6.2.19, 6.2.20 and 6.2.21 are proven in Appendix B.2.

- **Proof of Proposition 6.2.18**

Proof. By Proposition 6.2.14 we have that

$$\lim_{N\to\infty} \mathcal{L}[\bar{\Theta}^{[N]}(Ns)_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \tag{6.203}$$

in the normal topology and therefore (see B.2 Lemma B.2.1)

$$\lim_{N\to\infty} \mathcal{L}[\bar{\Theta}^{[N]}(Ns)_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \text{ in Meyer-Zheng topology.} \tag{6.204}$$

By Lemma 6.2.7, for $s > 0$,

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{[N]}(Ns) - x_1^{[N]}(s)\right|\right] = 0 \tag{6.205}$$

and therefore, by Lemma 6.2.19,

$$\lim_{N\to\infty} d_P(\psi_{\bar{\Theta}^{[N]}}, \psi_{x_1^{[N]}}) = 0 \text{ in probability.} \tag{6.206}$$

To apply the above results to our model, we recall the following basic result (see [7, Chapter 1]), which also holds for the Meyer-Zheng topology.

Lemma 6.3.2. *Let X_n, Y_n be random variables. If*

$$\lim_{n\to\infty} \mathcal{L}[X_n] = \mathcal{L}[X] \tag{6.207}$$

and

$$\lim_{n\to\infty} d(X_n, Y_n) = 0 \textit{ in probability,} \tag{6.208}$$

then

$$\lim_{n\to\infty} \mathcal{L}[Y_n] = \mathcal{L}(X). \tag{6.209}$$

Applying Lemma (6.3.2) to our case, we obtain

$$\lim_{N\to\infty} \mathcal{L}[(x_1^{[N]}(s))_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \text{ in the Meyer-Zheng topology.} \tag{6.210}$$

The argument for

$$\lim_{N\to\infty} \mathcal{L}[(y_1^{[N]}(s))_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \text{ in the Meyer-Zheng topology} \tag{6.211}$$

follows in the same way. By Lemma 6.2.20, we obtain

$$\lim_{N\to\infty} \mathcal{L}[(x_1^{[N]}(s), y_1^{[N]}(s) - x_1^{[N]}(s))_{s>0}] = \mathcal{L}[(\bar{\Theta}(s), 0)_{s>0}] \text{ in the Meyer-Zheng topology.} \tag{6.212}$$

Applying Lemma 6.2.21 with $f(x, y) = f(x, y + x)$ and the continuous mapping theorem, we obtain

$$\lim_{N\to\infty} \mathcal{L}\left[\left(x_1^{[N]}(s), y_1^{[N]}(s)\right)_{s>0}\right] = \mathcal{L}[(\bar{\Theta}(s), \bar{\Theta}(s))_{s>0}] \text{ in the Meyer-Zheng topology.} \tag{6.213}$$

□

§6.3.4 Proof of step 4. Mean-field finite-systems scheme

- **Proof of Proposition 6.2.1**

Proof. Proposition 6.2.1(b) follows directly from Propostion 6.2.14. The proof of Proposition 6.2.1(a) follows from Proposition 6.2.1(b) and Proposition 6.2.18.

To prove Proposition 6.2.1(c) fix $t > 0$. Consider the processes $(X^{[N]}(sN+t), Y^{[N]}(sN+t))_{t\geq 0}$ as processes on $([0,1]^2)^{\mathbb{N}_0}$ by periodic continuation. Since $([0,1]^2)^{\mathbb{N}_0}$ is compact, the sequence $(X^{[N]}(sN+t), Y^{[N]}(sN+t))_{N\in\mathbb{N}}$ is tight and hence has a converging subsequence. Let $(\bar{\Theta}(s))_{s>0}$ be the limiting process obtained in Proposition (6.2.14). This has continuous paths and is the unique solution of a well-posed martingale problem, and hence is a Markov process. Denote by Q_s the time-s semigroup corresponding to $(\bar{\Theta}(s))_{s\geq 0}$. Combining Proposition 6.2.4, Proposition 6.2.14 and Lemma 6.2.15, we get that, for each converging subsequence,

$$\lim_{k\to\infty} \mathcal{L}\left[X^{[N_k]}(sN_k+t), Y^{[N_k]}(sN_k+t)\right] = \int Q_s\left(\theta, \mathrm{d}\theta'\right)\nu_{\theta'} = \nu(s), \tag{6.214}$$

and hence it follows that, for all $t \geq 0$,

$$\lim_{N\to\infty} \mathcal{L}\left[X^{[N]}(sN+t), Y^{[N]}(sN+t)\right] = \int Q_s\left(\theta, \mathrm{d}\theta'\right)\nu_{\theta'} = \nu(s). \tag{6.215}$$

Let $(X^{\nu(s)}(t), Y^{\nu(s)}(t))_{t\geq 0}$ be the infinite system defined in (6.18), starting from initial measure $\nu(s)$. Then it follows from Corollary 6.3.1 that we can construct the processes $(X^{[N]}(sN+t), Y^{[N]}(sN+t))_{t\geq 0}$ and $(X^{\nu(s)}(t), Y^{\nu(s)}(t))_{t\geq 0}$ on one probability space such that, for all $t \geq 0$,

$$\lim_{N\to\infty} \mathbb{E}\left[\left|z^{\nu(s)}_{(i,R_i)}(t) - z^{[N]}_{(i,R_i)}(sN+t)\right|\right] = 0 \qquad \forall\,(i,R_i) \in \mathbb{Z}\times\{A,D\}. \tag{6.216}$$

Hence we see that the finite-dimensional distributions of the process $(X^{[N]}(sN+t), Y^{[N]}(sN+t))_{t\geq 0}$ converge to the finite-dimensional distributions of te process $(X^{\nu(s)}(t), Y^{\nu(s)}(t))_{t\geq 0}$.

Since we want that

$$\lim_{N\to\infty} \mathcal{L}\left[\left(X^{[N]}(sN+t), Y^{[N]}(sN+t)\right)_{t\geq 0}\right] = \mathcal{L}\left[\left(X^{\nu(s)}(t), Y^{\nu(s)}(t)\right)_{t\geq 0}\right], \tag{6.217}$$

we are left to show the tightness of $\mathcal{L}[(X^{[N]}(sN+t), Y^{[N]}(sN+t))_{t\geq 0}]_{N\in\mathbb{N}}$ in the path space $\mathcal{C}([0,\infty),([0,1]\times[0,1))^{\mathbb{N}_0})$. Since $([0,1]^2)^{\mathbb{N}_0}$ is endowed with the product topology, it is enough to show for all sequence components $(x^{[N]}_i(t))_{t\geq 0}$ and $(y^{[N]}_i(t))_{t\geq 0}$ that they are tight in path space (see [7, Theorem 7.3]).

To prove that the components are tight, we use a tightness criterion for semimartingales by Joffe and Metivier, [49, Proposition 3.2.3]. To use this criterion, we have to show that for all $i \in [N]$ the components $(x^{[N]}_i(t), y^{[N]}_i(t))$ are $\mathcal{D}$-semimartingales as defined in [49, Definition 3.1.1]. To do this, let $\mathcal{C}^* \subset \mathcal{C}_b\left([0,1]^2\right)$ be the set of polynomials on $[0,1]^2$, and define the operator

$$G^{[N]}_{\dagger}\colon\, (\mathcal{C}^*\times[0,1]^2\times[0,\infty),\Omega) \to \mathbb{R} \tag{6.218}$$

by

$$\begin{aligned} G^{[N]}_{\dagger}(f,(x,y),t,\omega) = &\left[\frac{c}{N}\sum_{j\in[N]}[x^{[N]}_j(t,\omega)-x] + K\,[y-x]\right]\frac{\partial f}{\partial x} \\ &+ \frac{1}{2}g(x)\frac{\partial^2 f}{\partial x^2} + e\,[x-y]\frac{\partial f}{\partial y}. \end{aligned} \tag{6.219}$$

Chapter 6

We use the subscript $\dagger$ to emphasize that $G^{[N]}_\dagger$ is the operator of a $\mathcal{D}$-semi-martingale and not a generator. Below we check in 4 steps that the component processes $(x^{[N]}_i(t), y^{[N]}_i(t))_{t\geq 0}$ are indeed $\mathcal{D}$-semi-martingales.

(a) The functions

$$f_1(x_i, y_i) = x_i, \quad f_2(x_i, y_i) = y_i, \tag{6.220}$$

are in $\mathcal{C}^*$, and so are f_1^2, $f_1 f_2$, f_2^2.

(b) For every $((x,y),t,\omega) \in \big([0,1]^2 \times [0,\infty) \times \Omega\big)$, the mapping $f \mapsto G^{[N]}_\dagger(f,(x,y),t,\omega)$ is linear on $\mathcal{C}^*$ and $G^{[N]}_\dagger(f,\cdot,t,\omega) \in \mathcal{C}^*$.

(c) Let $(\mathcal{F}_s)_{s\geq 0}$ be the filtration generated by the Brownian motions $((w_i(s))_{s\geq 0})_{i\in[N]}$, and let $\mathcal{P}$ be the σ-algebra generated by the predictable sets, i.e., sets of the form $(s,t] \times F$ for $F \in \mathcal{F}_s$. Since the component processes $(x^{[N]}_j(t))_{t\geq 0}$ are continuous, $((x,y),t,\omega) \mapsto G^{[N]}_\dagger(f,(x,y),t,\omega)$ is $\mathcal{B}([0,1]^2) \otimes \mathcal{P}$ measurable for every $f \in \mathcal{C}^*$, where $\mathcal{P}$ is the σ-algebra generated by the sets of the form $(s,t] \times F$ for $F \in \mathcal{F}_s$.

(d) Applying Itô's formula to the SSDE in (6.15), we obtain, for every $f \in \mathcal{C}^*$,

$$\begin{aligned}
f\big(x^{[N]}_i(t), y^{[N]}_i(t)\big) &= f\big(x^{[N]}_i(0), y^{[N]}_i(0)\big) \\
&\quad + \int_0^t \mathrm{d}s\, \frac{c}{N} \sum_{j\in[N]} \big[x^{[N]}_j(s,\omega) - x^{[N]}_i(s)\big]\, \frac{\partial f}{\partial x}\big(x^{[N]}_i(t), y^{[N]}_i(t)\big) \\
&\quad + \frac{1}{2}\int_0^t \mathrm{d}w_i(s)\, \sqrt{g(x^{[N]}_i(s))}\, \frac{\partial f}{\partial x}\big(x^{[N]}_i(t), y^{[N]}_i(t)\big) \\
&\quad + \int_0^t \mathrm{d}s\, Ke\big[y^{[N]}_i(s) - x^{[N]}_i(s)\big]\, \frac{\partial f}{\partial x}\big(x^{[N]}_i(t), y^{[N]}_i(t)\big) \\
&\quad + \int_0^t \mathrm{d}s\, e\big[x^{[N]}_i(s) - y^{[N]}_i(s)\big]\, \frac{\partial f}{\partial y}\big(x^{[N]}_i(t), y^{[N]}_i(t)\big) \\
&\quad + \int_0^t \mathrm{d}s\, g(x^{[N]}_i(s))\, \frac{\partial^2 f}{\partial x^2}\big(x^{[N]}_i(t), y^{[N]}_i(t)\big).
\end{aligned} \tag{6.221}$$

Therefore

$$\begin{aligned}
M^{[N],f}(t,\omega) &= f\big(x^{[N]}_i(t,\omega), y^{[N]}_i(t,\omega)\big) - f\big(x^{[N]}_i(0,\omega), y^{[N]}_i(0,\omega)\big) \\
&\quad - \int_0^t \mathrm{d}s\, G^{[N]}_\dagger\big(f(x^{[N]}_i(s,\omega), y^{[N]}_i(s,\omega)), s, \omega\big)
\end{aligned} \tag{6.222}$$

is a square-integrable martingale on $(\Omega, (\mathcal{F}_s)_{s\geq 0}, \mathbb{P})$.

To check that the sequence of component processes $((x^{[N]}_i(t), y^{[N]}_i(t)))_{N\in\mathbb{N}}$ is tight, we need the local characteristics of the $\mathcal{D}$-semi-martingale, which are defined in [49, Definition 3.1.2] as (recall (6.220))

$$\begin{aligned}
b_1^{[N]}((x,y),t,\omega) &= G_\dagger^{[N]}(f_1,(x,y),t,\omega),\\
b_2^{[N]}((x,y),t,\omega) &= G_\dagger^{[N]}(f_2,(x,y),t,\omega),\\
a_{(1,1)}^{[N]}((x,y),t,\omega) &= G_\dagger^{[N]}(f_1 f_1,(x,y),t,\omega) - 2x\, b_1((x,y),t,\omega),\\
a_{(2,1)}^{[N]}((x,y),t,\omega) &= G_\dagger^{[N]}(f_1 f_2,(x,y),t,\omega) - x\, b_2((x,y),t,\omega) - y b_1((x,y),t,\omega),\\
a_{(1,2)}^{[N]}((x,y),t,\omega) &= a_{(2,1)}((x,y),t,\omega),\\
a_{(2,2)}^{[N]}((x,y),t,\omega) &= G_\dagger^{[N]}(f_2 f_2,(x,y),t,\omega) - 2y\, b_2((x,y),t,\omega).
\end{aligned} \tag{6.223}$$

Hence

$$\begin{aligned}
b_1^{[N]}((x,y),t,\omega) &= \frac{c}{N}\sum_{j\in[N]}[x_j^{[N]}(t,\omega)-x] + Ke[y-x],\\
b_2^{[N]}((x,y),t,\omega) &= e[x-y],\\
a_{(1,1)}^{[N]}((x,y),t,\omega) &= 2g(x),\\
a_{(1,2)}^{[N]}((x,y),t,\omega) &= a_{(2,1)}((x,y),t,\omega) = 0,\\
a_{(2,2)}^{[N]}((x,y),t,\omega) &= 0.
\end{aligned} \tag{6.224}$$

Here, $b_i^{[N]}$ and $a_{i,j}^{[N]}$, $i,j \in \{1,2\}$, are called the local coefficients of first and second order. We check that the hypotheses [49, H1, H2, H3 in Section 3.2.1] are satisfied.

H_1: For all $N \in \mathbb{N}$,

$$\sum_{i\in\{1,2\}} |b_i^{[N]}((x,y),t,\omega)|^2 + \sum_{i,j\in\{1,2\}} |a_{(i,j)}^{[N]}((x,y),t,\omega)|^2 \le 4(c+Ke+e)^2 + 2||g||^2. \tag{6.225}$$

Hence, choosing as positive adapted process the constant process 1 and letting the constant be equal to $4(c+Ke+e)^2+2||g||^2$, we see that H_1 is satisfied.

H_2: Since the component processes are bounded by 1, also H_2 is satisfied.

H_3: Since the increasing càdlàg function $(A^{[N]}(t))_{t\ge 0}$ in [49, Definition 3.1.1] is in our case $A^{[N]}(t) = t$, also H_3 is satisfied.

Since H_1, H_2, H_3 are met, [49, Proposition 3.2.3] implies that $((x_i^{[N]}(t), y_i^{[N]}(t))_{t>0})_{N\in\mathbb{N}}$ are tight in the space of càdlàg paths $\mathcal{D}((0,\infty],[0,1]^2)$. Hence (6.217) indeed holds. □

Chapter 6

CHAPTER 7

Two-colour mean-field system

In this chapter we extend the results obtained in Section 6.2.1 to a mean-field system where the seed-bank consists of two colours, one colour that interacts on the slow time scale and one colour that interacts on the fast time scale. To do so we follow the set-up used in Chapter 4.4. In particular, we highlight the role of the second colour. Section 7.1 builds up the setting and states the main scaling result: Proposition 7.1.2. Section 7.2 provides the proof of this proposition based on a series of lemmas, which are stated and proved first.

§7.1 Two-colour mean-field finite-systems scheme

Setup. In this section we consider a simplified version of our SSDE in (4.20) on the finite geographic space

$$[N] = \{0, 1, \dots, N-1\}, \qquad N \in \mathbb{N}. \tag{7.1}$$

The migration kernel $a^{\Omega_N}(\cdot,\cdot)$ is replaced by the migration kernel $a^{[N]}(i,j) = c_0 N^{-1}$ for all $i, j \in [N]$, where $c_0 \in (0,\infty)$ is a constant. The seed-bank consists of *two colours*, labeled 0 and 1. The exchange rates between the active and the colour-0 dormant population are given by $K_0 e_0, e_0$. The exchange rates between active and the colour-1 dormant population are given by $\frac{K_1 e_1}{N}$, $\frac{e_1}{N}$. The state space is

$$S = \mathfrak{s}^{[N]}, \qquad \mathfrak{s} = [0,1] \times [0,1]^2, \tag{7.2}$$

and the system, consisting of three components, is denoted by

$$\begin{aligned} &Z^{[N]}(t) = \big(X^{[N]}(t), (Y_0^{[N]}(t), Y_1^{[N]}(t))\big)_{t \geq 0}, \\ &\big(X^{[N]}(t), (Y_0^{[N]}(t), Y_1^{[N]}(t))\big) = \big(x_i(t), (y_{i,0}(t), y_{i,1}(t))\big)_{i \in [N]}. \end{aligned} \tag{7.3}$$

The components of $(Z^{[N]}(t))_{t \geq 0}$ evolve according to the SSDE

$$\begin{aligned} \mathrm{d}x_i^{[N]}(t) &= \frac{c_0}{N} \sum_{j \in [N]} [x_j^{[N]}(t) - x_i^{[N]}(t)]\,\mathrm{d}t + \sqrt{g(x_i^{[N]}(t))}\,\mathrm{d}w_i(t) \\ &\qquad + K_0 e_0\,[y_{i,0}^{[N]}(t) - x_i^{[N]}(t)]\,\mathrm{d}t + \frac{K_1 e_1}{N}\,[y_{i,1}^{[N]}(t) - x_i^{[N]}(t)]\,\mathrm{d}t, \\ \mathrm{d}y_{i,0}^{[N]}(t) &= e_0\,[x_i^{[N]}(t) - y_{i,0}^{[N]}(t)]\,\mathrm{d}t, \\ \mathrm{d}y_{i,1}^{[N]}(t) &= \frac{e_1}{N}\,[x_i^{[N]}(t) - y_{i,1}^{[N]}(t)]\,\mathrm{d}t, \qquad i \in [N], \end{aligned} \tag{7.4}$$

which is a special case of (4.20). The initial state is $\mu(0) = \mu^{\otimes[N]}$ for some $\mu \in \mathcal{P}([0,1]^3)$. The SSDE in (7.4) has a unique weak solution coming from a well-posed martingale problem [67, Theorem 3.1]. By [67, Theorem 3.2], (7.4) has a unique strong solution for every deterministic initial state $Z^{[N]}(0)$. Therefore the solutino of (7.4) is Feller and Markov for any initial law. The SSDE in (7.4) can alternatively be written as

$$\begin{aligned}
\mathrm{d}x_i^{[N]}(t) &= c_0 \left[\frac{1}{N}\sum_{j\in[N]} x_j^{[N]}(t) - x_i^{[N]}(t)\right] \mathrm{d}t + \sqrt{g\big(x_i^{[N]}(t)\big)}\,\mathrm{d}w_i(t)\\
&\qquad + K_0 e_0\,[y_{i,0}^{[N]}(t) - x_i^{[N]}(t)]\,\mathrm{d}t + \frac{K_1 e_1}{N}\,[y_{i,1}^{[N]}(t) - x_i^{[N]}(t)]\,\mathrm{d}t, \qquad (7.5)\\
\mathrm{d}y_{i,0}^{[N]}(t) &= e_0\,[x_i^{[N]}(t) - y_{i,0}^{[N]}(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,1}^{[N]}(t) &= \frac{e_1}{N}\,[x_i^{[N]}(t) - y_{i,1}^{[N]}(t)]\,\mathrm{d}t, \qquad i\in[N].
\end{aligned}$$

So the migration term for a single colony can be interpreted as a drift towards the average of the active population. We are interested in $\mathcal{L}[(Z^{[N]}(t)))_{t\geq 0}]$ in the limit as $N\to\infty$, on time scales t and Ns. Heuristically, analysing the SSDE in (7.5), we can foresee the following results, which are made precise in Proposition 7.1.2.

• **On time scale** $1 = N^0$ (space-time scale 0), in the limit as $N\to\infty$ the colour-1 dormant population $(Y_1^{[N]}(t))_{t\geq 0}$ in (7.4) converges to a constant process, since the single components $y_{i,1}$ do not move on time scale t. The components of $(X^{[N]}(t), Y_0^{[N]}(t))_{t\geq 0}$ converge to i.i.d. copies of the single-colony McKean-Vlasov process in (6.1), where in the corresponding SSDE the parameters c, e, K are replaced by c_0, e_0, K_0 and $E=1$. So on time scale t we only see the colour-0 dormant population interacting with the active population, and the colour-1 dormant population is not yet coming into play. Therefore the colour-0 dormant population is the *effective seed-bank* on time scale 1, and the process

$$z_0^{\mathrm{eff},[N]}(t) = (x_0^{[N]}(t), y_{0,0}^{[N]}(t))_{t\geq 0} \qquad (7.6)$$

is called the effective process on level 0. Note that the active population has a drift towards $\frac{1}{N}\sum_{j\in[N]} x_j(t)$, which in the McKean-Vlasov limit is replaced by $\mathbb{E}[x(t)]$ given by (4.111).

• **On time scale** N (space-time scale 1), we look at the averages

$$\begin{aligned}
(z_1^{[N]}(s))_{s>0} &= \Big(x_1^{[N]}(s), (y_{0,1}^{[N]}(s), y_{1,1}^{[N]}(s))\Big)_{s>0}\\
&= \left(\frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Ns), \left(\frac{1}{N}\sum_{i\in[N]} y_{i,0}^{[N]}(Ns), \frac{1}{N}\sum_{i\in[N]} y_{i,1}^{[N]}(Ns)\right)\right)_{s>0}. \qquad (7.7)
\end{aligned}$$

Again the lower index 1 indicates that the average is the analogue of the 1-block average defined in (4.22). Using (7.4), we see that the dynamics of the system in (7.7) is given by the SSDE

$$\begin{aligned}
\mathrm{d}x_1^{[N]}(s) &= \sqrt{\frac{1}{N}\sum_{i\in[N]} g(x_i^{[N]}(Ns))}\,\mathrm{d}w(s) + NK_0e_0\left[y_{0,1}^{[N]}(s) - x_1^{[N]}(s)\right]\mathrm{d}s\\
&\qquad + K_1e_1\left[y_{1,1}^{[N]}(s) - x_1^{[N]}(s)\right]\mathrm{d}s, \qquad (7.8)\\
\mathrm{d}y_{0,1}^{[N]}(s) &= Ne_0\left[x_1^{[N]}(s) - y_{0,1}^{[N]}(s)\right]\mathrm{d}s,\\
\mathrm{d}y_{1,1}^{[N]}(s) &= e_1\left[x_1^{[N]}(s) - y_{1,1}^{[N]}(s)\right]\mathrm{d}s.
\end{aligned}$$

Thus, as in the mean-field system with one-colour, on time scale N infinite rates appear in the interaction of the active population with the colour-0 dormant population. Therefore in the limit as $N \to \infty$ the path becomes rougher and rougher at rarer and rarer times. Using the *Meyer-Zheng topology* we can prove that $\lim_{N\to\infty} y_{0,1}^{[N]}(s) = \lim_{N\to\infty} x_1^{[N]}(s)$ most of the time. On the other hand, on time scale N, $x_1^{[N]}(s)$ has a non-trivial interaction with $y_{1,1}^{[N]}(s)$, and therefore we say that on time scale N the colour-1 dormant population is the *effective seed-bank*. Note that for the evolution of the average $\frac{x_1^{[N]}(s)+K_0y_{0,1}^{[N]}(s)}{1+K_0}$ the rates with a factor N in front cancel out. We will use the quantity $\frac{x_1^{[N]}(s)+K_0y_{0,1}^{[N]}(s)}{1+K_0}$ to obtain results in the classical path-space topology. We call

$$(z_1^{[N],\mathrm{eff}}(s))_{s>0} = \left(\frac{x_1^{[N]}(s) + K_0y_{0,1}^{[N]}(s)}{1+K_0}, y_{1,1}(s)\right)_{s>0} \qquad (7.9)$$

the effective process on space-time scale 1. We will call space-time scale 1 also level 1.

Scaling limit. To describe the limiting dynamics of the system in (7.4), we need the infinite-dimensional process

$$(Z(t))_{t\geq 0} = \big((z_i(t))_{t\geq 0}\big)_{i\in\mathbb{N}_0} = \big((x_i(t),(y_{i,0}(t),y_{i,1}(t)))_{t\geq 0}\big)_{i\in\mathbb{N}_0} \qquad (7.10)$$

with state space $([0,1]^3)^{\mathbb{N}_0}$ that evolves according to

$$\begin{aligned}
\mathrm{d}x_i(t) &= c_0[\theta - x_i(t)]\,\mathrm{d}t + \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t) + K_0e_0\,[y_{i,0}(t) - x_i(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,0}(t) &= e_0\,[x_i(t) - y_{i,0}(t)]\,\mathrm{d}t, \qquad (7.11)\\
y_{i,1}(t) &= y_{i,1}, \qquad i\in\mathbb{N}_0.
\end{aligned}$$

Here, $\theta\in[0,1]$ and $y_{i,1}\in[0,1]$ for all $i\in\mathbb{N}_0$. We will also need the limiting effective process

$$(Z^{\mathrm{eff}}(t))_{t\geq 0} = \big((z_i^{\mathrm{eff}}(t))_{t\geq 0}\big)_{i\in\mathbb{N}_0} = \big((x_i^{\mathrm{eff}}(t),y_{i,0}^{\mathrm{eff}}(t))_{t\geq 0}\big)_{i\in\mathbb{N}_0} \qquad (7.12)$$

with state space $([0,1]^2)^{\mathbb{N}_0}$ that evolves according to

$$\begin{aligned}
\mathrm{d}x_i^{\mathrm{eff}}(t) &= c_0[\theta - x_i^{\mathrm{eff}}(t)]\,\mathrm{d}t + \sqrt{g\big(x_i^{\mathrm{eff}}(t)\big)}\,\mathrm{d}w(t) + K_0e_0\,[y_{i,0}^{\mathrm{eff}}(t) - x_i^{\mathrm{eff}}(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,0}^{\mathrm{eff}}(t) &= e_0\,[x_i^{\mathrm{eff}}(t) - y_{i,0}^{\mathrm{eff}}(t)]\,\mathrm{d}t, \qquad i\in\mathbb{N}_0.
\end{aligned} \qquad (7.13)$$

Like for the one-colour mean-field finite-systems scheme, we need the following list of ingredients to formally state our multi-scaling properties:

(a) For positive times $t > 0$, we define the so-called *estimators* for the finite system by:

$$\begin{aligned} \bar{\Theta}^{(1),[N]}(t) &= \frac{1}{N} \sum_{i\in[N]} \frac{x_i^{[N]}(t) + K_0 y_{i,0}^{[N]}(t)}{1+K_0}, \\ \Theta_x^{(1),[N]}(t) &= \frac{1}{N} \sum_{i\in[N]} x_i^{[N]}(t), \\ \Theta_{y_0}^{(1),[N]}(t) &= \frac{1}{N} \sum_{i\in[N]} y_{i,0}^{[N]}(t), \\ \Theta_{y_1}^{(1),[N]}(t) &= \frac{1}{N} \sum_{i\in[N]} y_{i,1}^{[N]}(t). \end{aligned} \tag{7.14}$$

We abbreviate

$$\begin{aligned} \mathbf{\Theta}^{(1),[N]}(t) &= \left(\Theta_x^{(1),[N]}(t), \Theta_{y_0}^{(1),[N]}(t), \Theta_{y_1}^{(1),[N]}(t)\right), \\ \mathbf{\Theta}^{\mathrm{eff},(1),[N]}(t) &= \left(\bar{\Theta}^{(1),[N]}(t), \Theta_{y_1}^{(1),[N]}(t)\right). \end{aligned} \tag{7.15}$$

We refer to $(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(t))_{t\geq 0}$ as the *effective estimator process* and to $(\mathbf{\Theta}^{(1),[N]}(t))_{t\geq 0}$ as the *estimator process*.

(b) The *time scale* Ns is such that $\mathcal{L}[\bar{\Theta}^{[N]}(Ns - L(N)) - \bar{\Theta}^{[N]}(Ns)] = \delta_0$ for all $L(N)$ satisfying $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, but not for $L(N) = N$. In words, Ns is the time scale on which $\bar{\Theta}^{[N]}(\cdot)$ starts evolving, i.e., $(\bar{\Theta}^{[N]}(Ns))_{s>0}$ is no longer a fixed process. When we scale time by Ns, we will use s as a time index, which indicates the "fast time scale". The "slow time scale" will be indicated by t. Thus, the time scales for the two-colour mean-field system are the same as the time scales for the one-colour mean-field system.

Remark 7.1.1 (Notation). The upper index 1 in $\bar{\Theta}^{(1)}$ and $\Theta_{y_1}^{(1)}$ is used to indicate that we are working with a system of level 1, so the system that lives on space-time scale 1. This can later be easily generalized to levels 2 and k. ■

(c) The *invariant measure* (i.e., the equilibrium measure) for the evolution of a single colony in (7.11), written

$$\Gamma_{\theta,\theta,y_1}, \tag{7.16}$$

and the *invariant measure* of the infinite system in (7.11), written $\nu_{\theta,\theta,\mathbf{y}_1} = \Gamma_{\theta,\theta,y_1}^{\otimes\mathbb{N}_0}$ with $\theta \in [0,1]$ and $\mathbf{y}_1 \in [0,1]^{\mathbb{N}_0}$ a random variable. The existence of the invariant measure ν_θ and the convergence of $\mathcal{L}[Z(t)_{t\geq 0}]$ towards ν_θ will be shown in the proof of Proposition 7.1.2.

(d) The invariant measure of the effective process in (7.13),

$$\Gamma_\theta^{\mathrm{eff}}, \tag{7.17}$$

and the invariant measure for the full process, $\nu_\theta^{\mathrm{eff}} = (\Gamma_\theta^{\mathrm{eff}})^{\otimes\mathbb{N}_0}$.

(e) The renormalisation transformation $\mathcal{F}\colon \mathcal{G} \to \mathcal{G}$,

$$(\mathcal{F}g)(\theta) = \int_{([0,1]^2)^{\mathbb{N}_0}} g(x_0)\, \nu_\theta^{\mathrm{eff}}(\mathrm{d}x_0, \mathrm{d}y_{0,0}), \quad \theta \in [0,1], \tag{7.18}$$

where $\Gamma_\theta^{\mathrm{eff}}$ is the equilibrium measure of (7.16). Note that this is the same transformation as defined in (4.75), but for the truncated system. Since $\nu_\theta^{\mathrm{eff}}$ is a product measure, we can write

$$(\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x)\, \Gamma_\theta^{\mathrm{eff}}(\mathrm{d}x, \mathrm{d}y_0), \quad \theta \in [0,1], \tag{7.19}$$

(f) The limiting 1-block process $(z_1(s))_{s>0} = (x_1(s), (y_{0,1}(s), y_{1,1}(s)))_{s>0}$ evolving according to

$$\begin{aligned} \mathrm{d}x_1(s) &= \frac{1}{1+K_0}\left[\sqrt{(\mathcal{F}g)\big(x_1(s)\big)}\,\mathrm{d}w(s) + K_1 e_1\,[y_{1,1}(s) - x_1(s)]\,\mathrm{d}s\right], \\ y_{0,1}(s) &= x_1(s), \\ \mathrm{d}y_{1,1}(s) &= e_1\,[x_1(s) - y_{1,1}(s)]\;\mathrm{d}s, \end{aligned} \tag{7.20}$$

where $\mathcal{F}g$ is defined in (7.20). The effective process $(z_1^{\mathrm{eff}}(s))_{s>0} = (x_1^{\mathrm{eff}}(s), y_{1,1}^{\mathrm{eff}}(s))_{s>0}$ on space-time scale 1,

$$\begin{aligned} \mathrm{d}x_1^{\mathrm{eff}}(s) &= \frac{1}{1+K_0}\left[\sqrt{(\mathcal{F}g)(x_1^{\mathrm{eff}}(s))}\,\mathrm{d}w(s) + K_1 e_1\,[y_{1,1}^{\mathrm{eff}}(s) - x_1^{\mathrm{eff}}(s)]\,\mathrm{d}s\right], \\ \mathrm{d}y_{1,1}^{\mathrm{eff}}(s) &= e_1\,[x_1^{\mathrm{eff}}(s) - y_{1,1}^{\mathrm{eff}}(s)]\,\mathrm{d}s. \end{aligned} \tag{7.21}$$

We are now ready to state the scaling limit for the evolution of the averages in (7.7), which we refer to as the *mean-field finite-systems scheme with two colours.*

Proposition 7.1.2 (Mean-field: two-colour finite-systems scheme). *Suppose that* $\mathcal{L}[Z^{[N]}(0)] = \mu^{\otimes[N]}$ *for some* $\mu \in \mathcal{P}\left([0,1]\times[0,1]^2\right)$. *Let*

$$\vartheta_0 = \mathbb{E}^\mu\left[\frac{x + K_0 y_0}{1+K_0}\right], \qquad \theta_{y_1} = \mathbb{E}^\mu\,[y_1]\,. \tag{7.22}$$

(a) For the effective estimator process defined in (7.15),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\right)_{s>0}\right] = \mathcal{L}\left[\left(z_1^{\mathrm{eff}}(s)\right)_{s>0}\right], \tag{7.23}$$

where the limit is determined by the unique solution of the SSDE (7.21), *with initial state*

$$z_1^{\mathrm{eff}}(0) = \left(x_1^{\mathrm{eff}}(0), y_1^{\mathrm{eff}}(0)\right) = (\vartheta_0, \theta_{y_1})\,. \tag{7.24}$$

(b) Assume for the 1*-dormant single components that*

$$\lim_{N\to\infty} \mathcal{L}\left[Y_1^{[N]}(Ns)\middle|\mathbf{\Theta}^{(1),[N]}(Ns)\right] = P^{z_1(s)}_{Y_1(s)}. \tag{7.25}$$

Define

$$\Gamma^{\mathrm{eff}}_{(\vartheta_0,\theta_{y_1})}(s) = \int_{[0,1]^2} S_s\big((\vartheta_0,\theta_{y_1}),\mathrm{d}(u_x,u_y)\big)\,\Gamma^{\mathrm{eff}}_{u_x} \in \mathcal{P}([0,1]^2), \tag{7.26}$$

where $S_s((\vartheta_0,\theta_{y_1}),\cdot)$ is the time-s marginal law of the process $(z_1^{\mathrm{eff}}(s))_{s>0}$ starting from $(\theta_0,\theta_{y_1}) \in [0,1]^2$ and $\Gamma^{\mathrm{eff}}_{u_x}$ is the equilibrium distribution of the system in (7.13) *with $\theta = u_x$ (note that $\Gamma^{\mathrm{eff}}_{\vartheta_0,\theta_{y_1}}(0) = \Gamma^{\mathrm{eff}}_{\vartheta_0}$). Let $(z^{\mathrm{eff},\Gamma_{(\vartheta_0,\theta_{y_1})}(s)}(t))_{t\geq 0}$ be the process with initial law $z^{\mathrm{eff},\Gamma_{(\vartheta_0,\theta_{y_1})}(s)}(0)$ drawn according to $\Gamma^{\mathrm{eff}}_{(\vartheta_0,\theta_{y_1})}(s)$ (which is a mixture of random processes in equilibrium) that, conditional on $x_1^{\mathrm{eff}}(s) = \theta$, evolves according to* (7.13)*. Then, for every $s \in (0,\infty)$,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(z_0^{\mathrm{eff},[N]}(Ns+t)\right)_{t\geq 0}\right] = \mathcal{L}\left[(z^{\Gamma^{\mathrm{eff}}_{(\vartheta_0,\theta_{y_1})}(s)}(t))_{t\geq 0}\right]. \tag{7.27}$$

(c) *For the averages in* (7.7)*,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(z_1^{[N]}(s)\right)_{s>0}\right] = \mathcal{L}\left[(z_1(s))_{s>0}\right] \quad \text{in the Meyer-Zheng topology,} \tag{7.28}$$

where the limit process is the unique solution of the SSDE in (7.20) *with initial state*

$$z_1(0) = (x_1(0), y_{0,1}(0), y_{1,1}(0)) = (\vartheta_0,\vartheta_0,\theta_{y_1}). \tag{7.29}$$

(d) *Assume 7.25 and define*

$$\nu(s) = \int_{[0,1]^3} S_s\big((\vartheta_0,\vartheta_0,\theta_{y_1}),\mathrm{d}(u_x,u_x,u_{y_1})\big) \int_{[0,1]^{\mathbb{N}_0}} P^{(u_x,u_x,u_{y_1})}_{Y_1(s)}(\mathrm{d}\boldsymbol{y}_1)\,\nu_{u_x,\boldsymbol{y}_1}, \tag{7.30}$$

where $S_s((\vartheta_0,\vartheta_0,\theta_{y_1}),\cdot)$ is the time-s marginal law of the process $(z_1(s))_{s>0}$ in (7.20)*, starting from $(\vartheta_0,\vartheta_0,\theta_{y_1}) \in [0,1]^3$, and $\nu_{u_x,\boldsymbol{y}_1}$ is the equilibrium distribution of the system in* (7.11) *with $\theta = u_x$ and $(y_{i,1})_{i\in\mathbb{N}_0} = \boldsymbol{y}_1$, (note that $\nu(0) = \nu_{\vartheta_0,(y_{i,1}(0))_{i\in\mathbb{N}_0}}$). Let $(z^{\nu(s)}(t))_{t\geq 0}$ be the process on $([0,1]^3)^{\mathbb{N}_0}$ with initial measure $z^{\nu(s)}(0)$ drawn according to $\nu(s)$ (which is a mixture of random processes in equilibrium) that conditional on $x_1(s) = \theta$ and $Y_1(s) = \boldsymbol{y}_1$ evolves according to* (7.11) *with $\theta = u_x$ and $(y_{i,1})_{i\in\mathbb{N}_0} = \boldsymbol{y}_1$. Then, for every $s \in (0,\infty)$,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(Z^{[N]}(Ns+t)\right)_{t\geq 0}\right] = \mathcal{L}\left[(z^{\nu(s)}(t))_{t\geq 0}\right]. \tag{7.31}$$

Remark 7.1.3 (Law of 1-dormant single components). Note that

$$\left(\mathcal{L}\left[Y_1^{[N]}(Ns)\,\middle|\,\left(\bar{\Theta}^{[N]}(Ns),\Theta^{[N]}_{y_1}(Ns)\right)\right]\right)_{N\in\mathbb{N}_0} \tag{7.32}$$

is a tight sequence of measures. Hence there exist weak limit points. In Section 8 we will see that if there is a higher layer in the hierarchy, then we can show that all weak limit points of (7.32) are the same and we can identify the limit. For Theorems 4.4.2 and 4.4.4 we do not need this assumption, since there will alwyas be multiple higher levels. ■

§7.2 Proof of the two-colour mean-field finite-systems scheme

The proof of Proposition 7.1.2, the finite-systems scheme with one level and two colours, follows the strategy used in Section 6.3 for the proof of Proposition 6.2.1. Like for the one-colour finite-systems scheme, we denote the slow time scale by t and the fast time scale by s. The proof consists of the following 6 steps:

1 Tightness of the effective estimator processes defined in (7.15).

$$\big((\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s>0}\big)_{N\in\mathbb{N}} \tag{7.33}$$

2 Stability property of $(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns+t))_{t>0}$, i.e., for $L(N)$ satisfying $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, and all $\epsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left[\sup_{0\leq t\leq L(N)} \left|\bar{\Theta}^{(1),[N]}(Ns) - \bar{\Theta}^{(1),[N]}(Ns-t)\right| > \epsilon\right] = 0. \tag{7.34}$$

and

$$\lim_{N\to\infty} \mathbb{P}\left[\sup_{0\leq t\leq L(N)} \left|\Theta_{y_1}^{(1),[N]}(Ns) - \Theta_{y_1}^{(1),[N]}(Ns-t)\right| > \epsilon\right] = 0. \tag{7.35}$$

3 Equilibrium of the infinite system and the one-dimensional distribution of the effective single components $(Z(Ns+t))_{t>0}$, analogous to Proposition 6.2.4.

4 Limiting evolution of the effective processes $((\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s>0})_{N\in\mathbb{N}}$.

5 Evolution of the 1-blocks in the Meyer-Zheng topology.

6 Proof of Proposition 7.1.2.

Step 1: Tightness of the 1-block estimators.

Lemma 7.2.1 (Tightness of the 1-block estimator). *Let*

$$(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s>0} \tag{7.36}$$

be defined as in (7.14)*. Then* $(\mathcal{L}[(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s>0}])_{N\in\mathbb{N}}$ *is a tight sequence of probability measures on* $\mathcal{C}((0,\infty),[0,1]^2)$*.*

Proof. To prove tightness of $((\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s>0})_{N\in\mathbb{N}}$, we will prove for all $\epsilon > 0$ that the set of measures $((\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s\geq\epsilon})_{N\in\mathbb{N}}$ is tight. To do so, fix $\epsilon > 0$. We will again use [49, Proposition 3.2.3]. From (7.4) we find that the 1-block averages

CHAPTER 7

$(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Ns))_{s>0}$ evolve according to

$$\begin{aligned}
\mathrm{d}\bar{\Theta}^{(1),[N]}(Ns) &= \frac{1}{1+K_0}\Bigg[\sqrt{\frac{1}{N}\sum_{i\in[N]} g\big(x_i^{[N]}(Ns)\big)}\,\mathrm{d}w_i(s) \\
&\qquad + K_1 e_1\left[\Theta_{y_1}^{(1),[N]} - \frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Ns)\right]\mathrm{d}s\Bigg], \\
\mathrm{d}\Theta_{y_1}^{(1),[N]}(Ns) &= e_1\left[\frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Ns) - \Theta_{y_1}^{(1),[N]}(Ns)\right]\mathrm{d}s.
\end{aligned} \tag{7.37}$$

To use [49, Proposition 3.2.3], we define $\mathcal{C}^*$ as the set of polynomials on $([0,1]^2)$. Note that $(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Ns))_{s\geq\epsilon}$ is a semi-martingale. Applying Itô's formula, we get

$$\begin{aligned}
&f\left(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\right) \\
&= f\left(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(N\epsilon)\right) \\
&\quad + \int_\epsilon^s \mathrm{d}w_i(r)\,\frac{1}{1+K_0}\sqrt{\frac{1}{N}\sum_{i\in[N]} g\big(x_i^{[N]}(Nr)\big)}\,\frac{\partial f}{\partial x}\left(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Nr)\right) \\
&\quad + \int_\epsilon^s \mathrm{d}r\,\frac{K_1 e_1}{1+K_0}\left[\Theta_{y_1}^{(1),[N]}(Nr) - \frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Nr)\right]\frac{\partial f}{\partial x}\left(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Nr)\right) \\
&\quad + \int_\epsilon^s \mathrm{d}r\, e_1\left[\frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Nr) - \Theta_{y_1}^{(1),[N]}(Nr)\right]\frac{\partial f}{\partial y}\left(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Nr)\right) \\
&\quad + \int_\epsilon^s \mathrm{d}r\,\frac{1}{2(1+K_0)^2}\frac{1}{N}\sum_{i\in[N]} g\big(x_i^{[N]}(Nr)\big)\,\frac{\partial^2 f}{\partial x^2}\left(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Nr)\right)
\end{aligned} \tag{7.38}$$

for all $f\in\mathcal{C}^*$. Hence, if we define the operator

$$\begin{aligned}
&G_\dagger^{(1),[N]}\colon\ (\mathcal{C}^*,[0,1]^2,[\epsilon,\infty),\Omega)\to\mathbb{R}, \\
&G_\dagger^{(1),[N]}(f,(x,y),s,\omega) = \frac{K_1 e_1}{1+K_0}\left[y - \frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Ns,\omega)\right]\frac{\partial f}{\partial x} \\
&\qquad + e_1\left[\frac{1}{N}\sum_{i\in[N]} x_i^{[N]}(Ns,\omega) - y\right]\frac{\partial f}{\partial y} \\
&\qquad + \frac{1}{2(1+K_0)^2}\frac{1}{N}\sum_{i\in[N]} g(x_i^{[N]}(Ns,\omega))\,\frac{\partial^2 f}{\partial x^2},
\end{aligned} \tag{7.39}$$

then we see that the process $(\boldsymbol{\Theta}^{\mathrm{eff},(1),[N]}(Ns))_{s\geq\epsilon}$ is a $\mathcal{D}$-semi-martingale for all $\epsilon>0$. For all $\epsilon>0$ the conditions H_1, H_2, H_3 are satisfied as before. Therefore we conclude from [49, Proposition 3.2.3] that the sequence $((\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns))_{s\geq\epsilon})_{N\in\mathbb{N}}$ is tight. Since this is true for all $\epsilon>0$, we conclude that $(\mathcal{L}[(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns))_{s>0}])_{N\in\mathbb{N}}$ is tight. □

Step 2: Stability of the 1-block estimators.

Lemma 7.2.2 (Stability property of the 1-block estimator). *Let* $\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(t)$ *be defined as in* (7.14). *For any* $L(N)$ *satisfying* $\lim_{N\to\infty}L(N)=\infty$ *and* $\lim_{N\to\infty}L(N)/N=0$,

$$\lim_{N\to\infty}\sup_{0\leq t\leq L(N)}\left|\bar\Theta^{(1),[N]}(Ns)-\bar\Theta^{(1),[N]}(Ns-t)\right|=0 \text{ in probability} \tag{7.40}$$

and

$$\lim_{N\to\infty}\sup_{0\leq t\leq L(N)}\left|\Theta^{(1),[N]}_{y_1}(Ns)-\Theta^{(1),[N]}_{y_1}(Ns-t)\right|=0 \text{ in probability.} \tag{7.41}$$

Proof. Fix $\epsilon>0$. From the SSDE (7.4) we obtain that, for N large enough,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq t\leq L(N)}\left|\bar\Theta^{(1),[N]}(Ns)-\bar\Theta^{(1),[N]}(Ns-t)\right|>\epsilon\right)\\
&=\mathbb{P}\Bigg(\sup_{0\leq t\leq L(N)}\frac{1}{1+K_0}\Bigg|\int_{Ns-t}^{Ns}\mathrm{d}r\,\frac{K_1e_1}{N}\left[\Theta^{(1),[N]}_{y_1}(r)-\frac{1}{N}\sum_{i\in[N]}x_i^{[N]}(r)\right]\\
&\qquad\qquad+\int_{Ns-t}^{Ns}\mathrm{d}w_i(r)\,\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N]}(r))}\Bigg|>\epsilon\Bigg]\\
&\leq\mathbb{P}\left(\left|\frac{L(N)2K_1e_1}{N(1+K_0)}\right|+\sup_{0\leq t\leq L(N)}\left|\frac{1}{1+K_0}\int_{Ns-t}^{Ns}\mathrm{d}w_i(r)\,\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N]}(r))}\right|>\epsilon\right)\\
&=\mathbb{P}\left(\sup_{0\leq t\leq L(N)}\left|\frac{1}{1+K_0}\int_{Ns-t}^{Ns}\mathrm{d}w_i(r)\,\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N]}(r))}\right|>\epsilon-\frac{L(N)2K_1e_1}{N(1+K_0)}\right)\\
&\leq\mathbb{P}\left(\sup_{0\leq t\leq L(N)}\left|\frac{1}{1+K_0}\int_{Ns-t}^{Ns}\mathrm{d}w_i(r)\,\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N]}(r))}\right|>\frac{\epsilon}{2}\right).
\end{aligned} \tag{7.42}$$

Applying the same optional stopping argument as used in the proof of Lemma 6.2.15, we find (7.40). For (7.41), note that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq t\leq L(N)}\left|\Theta^{(1),[N]}_{y_1}(Ns)-\Theta^{(1),[N]}_{y_2}(Ns-t)\right|>\epsilon\right)\\
&=\mathbb{P}\left(\sup_{0\leq t\leq L(N)}\frac{1}{1+K_0}\left|\int_{Ns-t}^{Ns}\mathrm{d}r\,\frac{e_1}{N}\left[\Theta^{(1),[N]}_{y_1}(r)-\frac{1}{N}\sum_{i\in[N]}x_i^{[N]}(r)\right]\right|>\epsilon\right) \qquad (7.43)\\
&\leq\mathbb{P}\left(\frac{2e_1L(N)}{(1+K_0)N}>\epsilon\right).
\end{aligned}$$

Let $N \to \infty$ to obtain (7.41). $\square$

Step 3: Equilibrium for the infinite system. To derive the equilibrium of the single components in the infinite system, we derive the following analoque of Proposition 6.2.4. Recall that the finite system is denoted by $Z^{[N_k]}$ in (7.3), and recall the list of ingredients in Section 7.1.

Proposition 7.2.3 (Equilibrium for the infinite 2-colour system). *Let $(N_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{N}$. Fix $s > 0$. Let $L(N)$ satisfy $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, and suppose that*

$$\begin{aligned}
&\lim_{k\to\infty} \mathcal{L}\left[\mathbf{\Theta}^{\mathrm{eff},(1),[N_k]}(N_k s)\right] = P_{\mathbf{\Theta}^{\mathrm{eff}}(s)},\\
&\lim_{k\to\infty} \mathcal{L}\left[Y_1^{[N_k]}(N_k s)\middle|\mathbf{\Theta}^{\mathrm{eff},(1),[N_k]}(N_k s)\right] = P_{Y_1(s)}^{\mathbf{\Theta}^{\mathrm{eff},(1)}(s)},\\
&\lim_{k\to\infty} \mathcal{L}\left[\sup_{0\le t\le L(N_k)} \left|\bar{\Theta}^{[N_k]}(N_k s) - \bar{\Theta}^{[N_k]}(N_k s - t)\right| + \left|\Theta_{y_1}^{[N_k]}(N_k s) - \Theta_{y_1}{}^{[N_k]}(N_k s - t)\right|\right]\\
&\qquad\qquad\qquad = \delta_0,\\
&\lim_{k\to\infty} \mathcal{L}\big[Z^{[N_k]}(N_k s)\big] = \nu(s).
\end{aligned} \tag{7.44}$$

Then $\nu(s)$ is of the form

$$\nu(s) = \int_{[0,1]^2} P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\mathrm{d}\theta, \mathrm{d}\theta_y) \int_{[0,1]^{\mathbb{N}_0}} P_{Y_1(s)}^{(\theta,\theta_y)}(\mathrm{d}\boldsymbol{y}_1)\, \nu_{\theta,\boldsymbol{y}_1}, \tag{7.45}$$

where $\boldsymbol{y}_1 = (y_{i,1})_{i\in\mathbb{N}_0}$ is a sequence with elements in $[0,1]$, and $\nu_{\theta,\boldsymbol{y}_1}$ is the equilibrium measure of the process in (7.10) *evolving according to* (7.11) *with $(y_{i,1})_{i\in\mathbb{N}_0}$ given by the sequence $\boldsymbol{y}_1 =$.*

Preparation for the proof of Proposition 7.2.3. The proof of Proposition 7.2.3 follows the same line of argument as used in the proof of Proposition 6.2.4. We need lemmas that are similar to Lemmas 6.2.5-6.2.11, but this time in the setting of the two-colour hierarchical mean-field finite-systems scheme. Afterwards we prove Proposition 7.2.3.

Lemma 7.2.4 (Convergence for the infinite system). *Let μ be an exchangeable probability measure on $([0,1]^3)^{\mathbb{N}_0}$. Then for the system $(Z(t))_{t\ge0}$ given by* (7.10) *with $\mathcal{L}[Z(0)] = \mu$,*

$$\lim_{t\to\infty} \mathcal{L}[Z(t)] = \nu_{\theta,\boldsymbol{y}_1}, \tag{7.46}$$

where $\nu_{\theta,\boldsymbol{y}_1}$ is of the form

$$\nu_{\theta,\boldsymbol{y}_1} = \prod_{i\in\mathbb{N}_0} \Gamma_{\theta,y_{i,1}} \tag{7.47}$$

with $\Gamma_{\theta,y_{i,1}}$ the equilibrium of the ith single-component process in (7.11).

Proof. For each component of the infinite system in (7.10) the 1-dormant single component process $(y_{i,1}(t))_{t\ge0}$ does not move on time scale t. Hence, given the states of 1-dormant single components, we can use a similar argument as in the proof of Proposition 6.1.2 (see Section 6.1.3) to show that the single components converge to an equilibrium measure $\Gamma_{\theta,y_{i,1}}$. Since the single components do not interact, the claim in Lemma 7.2.4 follows. $\square$

The second lemma establishes the continuity of the equilibrium with respect to θ, its center of drift.

Lemma 7.2.5 (Continuity of the equilibrium). *Let $\mathcal{P}(([0,1]^3)^{\mathbb{N}_0})$ denote the space of probability measures on $([0,1]^3)^{\mathbb{N}_0}$. The mapping*

$$\begin{aligned} [0,1]\times[0,1]^{\mathbb{N}_0} &\to \mathcal{P}(([0,1]^3)^{\mathbb{N}_0}) \\ (\theta,\boldsymbol{y}_1) &\mapsto \nu_{\theta,\boldsymbol{y}_1} \end{aligned} \tag{7.48}$$

is continuous. Furthermore, if h is a Lipschitz function on $[0,1]$, then also $\mathcal{F}h$ defined by

$$(\mathcal{F}h)(\theta)=\mathbb{E}^{\nu_{\theta,\boldsymbol{y}_1}}[h(\cdot)]=\int_{([0,1]^3)^{\mathbb{N}_0}}\nu_{\theta,\boldsymbol{y}_1}(\mathrm{d}z)\,h(x_0) \tag{7.49}$$

is a Lipschitz function on $[0,1]$, whose values are independent of $\boldsymbol{y}_1$.

Proof. Lemma 7.2.5 follows from the proof of Lemma 7.2.9. $\square$

The third lemma characterises the speed at which the estimators $(\Theta_x^{[N]}(t))_{t\geq 0}$ and $(\Theta_y^{[N]}(t))_{t\geq 0}$ converge to each other when $N\to\infty$ and $t\to\infty$.

Lemma 7.2.6 (Comparison of empirical averages). *Let $(\Theta_x^{(1),[N]}(t))_{t\geq 0}$ and $(\Theta_{y_0}^{(1),[N]}(t))_{t\geq 0}$ be defined as in (7.14). Then*

$$\begin{aligned} \mathbb{E}\left[\left|\Theta_x^{(1),[N]}(t)-\Theta_{y_0}^{(1),[N]}(t)\right|\right] &\leq \sqrt{\mathbb{E}\left[\left(\Theta_x^{(1),[N]}(0)-\Theta_{y_0}^{(1),[N]}(0)\right)^2\right]}\,\mathrm{e}^{-(K_0e_0+e_0)t} \\ &\quad+\sqrt{\frac{2}{K_0e_0+e_0}\left[\frac{||g||}{N}+\frac{4K_1e_1}{N}\right]}. \end{aligned} \tag{7.50}$$

Proof. From (7.4) it follows via Itô-calculus that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\left(\Theta_x^{(1),[N]}(t)-\Theta_{y_0}^{(1),[N]}(t)\right)^2\right] &= -2(K_0e_0+e_0)\,\mathbb{E}\left[\left(\Theta_x^{(1),[N]}(t)-\Theta_{y_0}^{(1),[N]}(t)\right)^2\right] \\ &\quad+h^{[N]}(t), \end{aligned} \tag{7.51}$$

where

$$\begin{aligned} h^{[N]}(t) &= \mathbb{E}\left[\frac{2K_1e_1}{N}\left(\Theta_x^{(1),[N]}(t)-\Theta_{y_0}^{(1),[N]}(t)\right)\left[\Theta_{y_1}^{(1),[N]}(t)-\Theta_x^{(1),[N]}(t)\right]\right] \\ &\quad+\frac{2}{N^2}\sum_{i\in[N]}\mathbb{E}\big[g\big(x_i^{[N]}(r)\big)\big]. \end{aligned} \tag{7.52}$$

Hence

$$\mathbb{E}\left[\left(\Theta_x^{(1),[N]}(t)-\Theta_{y_0}^{(1),[N]}(t)\right)^2\right]=\mathbb{E}\left[\left(\Theta_x^{(1),[N]}(0)-\Theta_{y_0}^{(1),[N]}(0)\right)^2\right]\mathrm{e}^{-2(K_0e_0+e_0)t} + \int_0^t \mathrm{d}r\,\mathrm{e}^{-2(K_0e_0+e_0)(t-r)}h^{[N]}(r). \tag{7.53}$$

Take the square root on both sides and use Jensen's inequality to get (7.50). □

Like for the mean-field system with one colour, we need to compare the finite system in (7.3) with an infinite system. To derive the analogue of Lemma 6.2.9, let $L(N)$ satisfy $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N=0$. Define the measure μ_N on $([0,1]^3)^{\mathbb{N}_0}$ by continuing the configuration of

$$Z^{[N]}(Ns-L(N))=\left(X^{[N]}(Ns-L(N)),\left(Y_0^{[N]}(Ns-L(N)),Y_1^{[N]}(Ns-L(N))\right)\right) \tag{7.54}$$

periodically to $([0,1]^3)^{\mathbb{N}_0}$. Let

$$\bar{\Theta}^{(1),[N]}=\frac{1}{N}\sum_{i\in[N]}\frac{x_i^{[N]}(Ns-L(N))+K_0y_{i,0}^{[N]}(Ns-L(N))}{1+K_0}. \tag{7.55}$$

Let

$$(Z^{\mu_N}(t))_{t\geq 0}=\left(X^{\mu_N}(t),(Y_0^{\mu_N}(t),Y_1^{\mu_N}(t))\right)_{t\geq 0} \tag{7.56}$$

be the infinite system evolving according to

$$\begin{aligned}
\mathrm{d}x_i^{\mu_N}(t)&=c_0\,[\bar{\Theta}^{(1),[N]}-x_i^{\mu_N}(t)]\,\mathrm{d}t+\sqrt{g\big(x_i^{\mu_N}(t)\big)}\,\mathrm{d}w_i(t)+K_0e_0\,[y_{i,0}^{\mu_N}(t)-x_i^{\mu_N}(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,0}^{\mu_N}(t)&=e_0\,[x_i^{\mu_N}(t)-y_{i,0}^{\mu_N}(t)]\,\mathrm{d}t,\\
y_{i,1}^{\mu_N}(t)&=y_{i,1}^{\mu_N}(0),\qquad i\in\mathbb{N}_0,
\end{aligned} \tag{7.57}$$

starting from initial distribution μ_N. Then the following lemma is the equivalent of Lemma 6.2.9 for the two-colour mean-field system.

Lemma 7.2.7 (Comparison of finite and infinite systems). *Fix $s>0$, and let $L(N)$ satisfy $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N$. Suppose that*

$$\lim_{N\to\infty}\sup_{0\leq t\leq L(N)}\left|\bar{\Theta}^{(1),[N]}(Ns)-\bar{\Theta}^{(1),[N]}(Ns-t)\right|=0 \quad \textit{in probability.} \tag{7.58}$$

Then, for all $t\geq 0$,

$$\lim_{k\to\infty}\left|\mathbb{E}\left[f\big(Z^{\mu_N}(t)\big)-f\big(Z^{[N]}(Ns-L(N)+t)\big)\right]\right|=0 \qquad \forall f\in\mathcal{C}\big(([0,1]^3)^{\mathbb{N}_0},\mathbb{R}\big). \tag{7.59}$$

Proof. We proceed as in the proof of Lemma 6.2.9. We rewrite the SSDE in (7.4) as

$$
\begin{aligned}
\mathrm{d}x_i^{[N]}(t) &= c_0\left[\Theta^{(1),[N]} - x_i^{[N]}(t)\right]\mathrm{d}t \\
&\quad + c_0\left[\bar{\Theta}^{(1),[N]}(t) - \Theta^{(1),[N]}\right]\mathrm{d}t + c_0\left[\Theta_x^{(1),[N]}(t) - \bar{\Theta}^{(1),[N]}(t)\right]\mathrm{d}t \\
&\quad + \sqrt{g\big(x_i^{[N]}(t)\big)}\,\mathrm{d}w_i(t) \\
&\quad + K_0 e_0\left[y_{i,0}^{[N]}(t) - x_i^{[N]}(t)\right]\mathrm{d}t + \frac{K_1 e_1}{N}\left[y_{i,1}^{[N]}(t) - x_i^{[N]}(t)\right]\mathrm{d}t, \\
\mathrm{d}y_{i,0}^{[N]}(t) &= e_0\left[x_i^{[N]}(t) - y_{i,0}^{[N]}(t)\right]\mathrm{d}t, \\
\mathrm{d}y_{i,1}^{[N]}(t) &= \frac{e_1}{N}\left[x_i^{[N]}(t) - y_{i,1}^{[N]}(t)\right]\mathrm{d}t, \qquad i \in [N].
\end{aligned}
\tag{7.60}
$$

As before, we consider the finite system in (7.60) as a system on $([0,1]^3)^{\mathbb{N}_0}$ by periodic continuation, and we couple the finite system in (7.60) and the infinite system in (7.59) via there Brownian motions. We denote the coupled process by $\tilde{z}(t) = (\tilde{z}_i(t))_{i\in\mathbb{N}_0} = (\tilde{z}_i^{[N]}(t), \tilde{z}_i^{\mu_N}(t))_{i\in\mathbb{N}_0}$, where $\tilde{z}_i^{[N]}(t) = (\tilde{x}_i^{[N]}(t), \tilde{y}_{i,0}^{[N]}(t), \tilde{y}_{i,1}^{[N]}(t))$ and $\tilde{z}_i^{\mu_N}(t) = (\tilde{x}_i^{\mu_N}(t), \tilde{y}_{i,0}^{\mu_N}(t), \tilde{y}_{i,1}^{\mu_N}(t))$. We define

$$
\begin{aligned}
\Delta_{i,0}^{[N]}(t) &= \tilde{x}_i^{[N]}(t) - \tilde{x}_i^{\mu_N}(t), \\
\delta_{i,0}^{[N]}(t) &= \tilde{y}_{i,0}^{[N]}(t) - \tilde{y}_{i,0}^{\mu_N}(t), \\
\delta_{i,1}^{[N]}(t) &= \tilde{y}_{i,1}^{[N]}(t) - \tilde{y}_{i,1}^{\mu_N}(t).
\end{aligned}
\tag{7.61}
$$

As in the proof of Lemma 6.2.9, we have to show that, for all $t \geq 0$,

$$
\lim_{N\to\infty} \mathbb{E}[|\Delta_i^{[N]}(t)|] = 0, \qquad \lim_{N\to\infty} \mathbb{E}[|\delta_{i,0}^{[N]}(t)|] = 0, \qquad \lim_{N\to\infty} \mathbb{E}[|\delta_{i,1}^{[N]}(t)|] = 0. \tag{7.62}
$$

To prove the third limit in (7.63), note that, by (7.57), (7.60) and the choice of the initial measure in the coupling,

$$
y_{i,1}^{[N]}(t) = y_{i,1}^{[N]}(0) + \frac{e_1}{N}\int_0^t \mathrm{d}r\,\left[x_i^{[N]}(r) - y_{i,1}^{[N]}(r)\right] = y_{i,1}^{\mu_N}(t) + \frac{e_1}{N}\int_0^t \mathrm{d}r\,\left[x_i^{[N]}(r) - y_{i,1}^{[N]}(r)\right]. \tag{7.63}
$$

Hence

$$
\lim_{N\to\infty} \mathbb{E}[|\delta_{i,1}^{[N]}(L(N))|] = 0. \tag{7.64}
$$

To prove the first two limits in (7.63), we argue as in the proof of Lemma 6.2.9, but we need to add extra drift terms towards the first seed-bank. Using Itô-calculus, we obtain

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[|\Delta_i^{[N]}(t)| + K|\delta_{i,0}^{[N]}(t)|] \\
&= -c\,\mathbb{E}[\Delta_i^{[N]}(t)] \\
&\quad - 2K_0 e_0\,\mathbb{E}\left[[|\Delta_i^{[N]}(t)| + |\delta_i^{[N]}(t)|]\,1_{\{\operatorname{sgn}\Delta_i^{[N]}(t)\neq\operatorname{sgn}\delta_{i,0}^{[N]}(t)\}}\right] \\
&\quad + c\operatorname{sgn}\Delta_i^{[N]}(t)\left[\bar{\Theta}^{(1),[N]}(t) - \bar{\Theta}^{(1),[N]}\right] \\
&\quad + c\operatorname{sgn}\Delta_i^{[N]}(t)\left[\bar{\Theta}_x^{(1),[N]}(t) - \bar{\Theta}^{(1),[N]}(t)\right] \\
&\quad + \frac{K_1 e_1}{N}\operatorname{sgn}\Delta_i^{[N]}(t)\left[\delta_{i,1}^{[N]}(t) - \Delta_i^{[N]}(t)\right].
\end{aligned}
\tag{7.65}
$$

This can be rewritten as

$$
\begin{aligned}
0 &\leq \mathbb{E}[|\Delta_i^{[N]}(t)| + K_0|\delta_{i,0}^{[N]}(t)|] \\
&\leq \mathbb{E}[|\Delta_i^{[N]}(0)| + K|\delta_{i,0}^{[N]}(0)|] - c\int_0^t \mathrm{d}r\,\mathbb{E}[\Delta_i^{[N]}(r)] \\
&\quad - 2K_0e_0\int_0^t \mathrm{d}r\,\mathbb{E}\left[[|\Delta_i^{[N]}(r)| + |\delta_{i,0}^{[N]}(r)|]\,1_{\{\mathrm{sgn}\,\Delta_i^{[N]}(t)\neq \mathrm{sgn}\,\delta_{i,0}^{[N]}(t)\}}\right] \\
&\quad + c\int_0^t \mathrm{d}r\,|\bar\Theta^{(1),[N]}(r) - \Theta^{(1),[N]}| \\
&\quad + c\int_0^t \mathrm{d}r\,|\bar\Theta_x^{(1),[N]}(r) - \bar\Theta^{(1),[N]}(r)| \\
&\quad + \frac{K_1e_1}{N}\int_0^t \mathrm{d}r\,\left|\delta_{i,1}^{[N]}(r) - \Delta_i^{[N]}(r)\right|.
\end{aligned} \tag{7.66}
$$

By the construction of the measure μ_N, we have

$$\lim_{N\to\infty}\mathbb{E}[|\Delta_i^{[N]}(0)| + K_0|\delta_i^{[N]}(0)|] = 0. \tag{7.67}$$

Therefore, for all $t \geq 0$,

$$\lim_{N\to\infty}\mathbb{E}[|\Delta_i^{[N]}(t)| + K_0|\delta_{i,0}^{[N]}(t)|] = 0. \tag{7.68}$$

Combine this with (7.64) and use that Lipschitz functions are dense in the set of bounded continuous functions. Then, as in the proof of Lemma 6.2.9, we get the claim in (7.59). □

Before we can prove that the infinite system $(X^{\mu_N}(t), Y_0^{\mu_N}(t), Y_1^{\mu_N}(t))_{t\geq 0}$ converges to a limiting system as $N\to\infty$, we need the following regularity property for the estimators $(\bar\Theta^{[N]}, \Theta_{y_1}^{[N]})$.

Lemma 7.2.8 (Stability of the estimator for the conserved quantity). *Define μ_N as in Lemma 7.2.7. Let $(x_i^N, y_{i,0}^N, y_{i,1}^N)_{i\in[N]}$ be distributed according to the exchangeable probability measure μ_N on $([0,1]^3)^{\mathbb{N}_0}$ restricted to $([0,1]^3)^{[N]}$. Suppose that $\lim_{N\to\infty}\mu_N = \mu$ for some exchangeable probability measure μ on $([0,1]^3)^{\mathbb{N}_0}$. Define the random variable ϕ on $(\mu, ([0,1]^3)^{\mathbb{N}_0})$ by putting*

$$
\begin{aligned}
&\phi = (\phi_1, \phi_2), \\
&\phi_1 = \lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]}\frac{x_i + Ky_{i,0}}{1+K}, \qquad \phi_2 = \lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]} y_{i,1},
\end{aligned} \tag{7.69}
$$

and the random variable $\phi^{[N]}$ on $(\mu_N, ([0,1]^3)^{\mathbb{N}_0})$ by putting

$$
\begin{aligned}
&\phi^{[N]} = (\phi_1^{[N]}, \phi_2^{[N]}) \\
&\phi_1^{[N]} = \frac{1}{N}\sum_{i\in[N]}\frac{x_i^N + Ky_{i,0}^N}{1+K}, \qquad \phi_2^{[N]} = \frac{1}{N}\sum_{i\in[N]} y_{i,1}^N.
\end{aligned} \tag{7.70}
$$

Then

$$\lim_{N\to\infty}\mathcal{L}[\phi^{[N]}] = \mathcal{L}[\phi]. \tag{7.71}$$

Proof. We can use a similar argument as in the proof of Lemma 6.2.10. Define

$$D^{[N]}(Z) = \left(\frac{1}{N} \sum_{j \in [N]} \frac{x_j + K_0 y_{j,0}}{1 + K_0}, \frac{1}{N} \sum_{j \in [N]} y_{i,1} \right). \tag{7.72}$$

Then we can proceed as in the proof of Lemma 6.2.10, using Fourier analysis for both components of $D^{[N]}(Z)$ separately. □

In the fifth and final lemma we state the convergence of $\mathcal{L}[(X^{\mu_N}(t), Y_0^{\mu_N}(t), Y_1^{\mu_N}(t))]$ to the law of a limiting system as $N \to \infty$.

Lemma 7.2.9 (Uniformity of the ergodic theorem for the infinite system). *Let μ_N be defined as in* (7.56)*. Since $(\mu_N)_{N\in\mathbb{N}}$ is tight, it has convergent subsequences. Let $(N_k)_{k\in\mathbb{N}}$ be a subsequence such that $\mu = \lim_{k\to\infty} \mu_{N_k}$. Define*

$$\Theta = \lim_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} \frac{x_i^\mu + K y_{i,0}^\mu}{1+K} \qquad \text{in } L_2(\mu). \tag{7.73}$$

Let $Z^\mu(t) = \big(X^\mu(t), Y_0^\mu(t), Y_1^\mu(t)\big)_{t\geq 0}$ be the infinite system evolving according to

$$\begin{aligned} \mathrm{d}x_i^\mu(t) &= c\,[\Theta - x_i^\mu(t)]\,\mathrm{d}t + \sqrt{g(x_i^\mu(t))}\,\mathrm{d}w_i(t) + Ke\,[y_{i,1}^\mu(t) - x_i^\mu(t)]\,\mathrm{d}t,\\ \mathrm{d}y_{i,0}^\mu(t) &= e\,[x_i^\mu(t) - y_{i,1}^\mu(t)]\,\mathrm{d}t,\\ \mathrm{d}y_{i,1}^\mu(t) &= y_{i,1}^\mu(0), \qquad i\in\mathbb{N}_0. \end{aligned} \tag{7.74}$$

and let $Z^{\mu_{N_k}}(t) = (X^{\mu_{N_k}}(t), Y_0^{\mu_{N_k}}(t), Y_1^{\mu_{N_k}}(t))_{t\geq 0}$ be the infinite system defined in (7.56)*. Then*

(a) *For all $t \geq 0$,*

$$\lim_{k\to\infty} \big|\mathbb{E}\big[f\big(Z^{\mu_{N_k}}(t)\big)\big] - \mathbb{E}\big[f\big(Z^\mu(t)\big)\big]\big| = 0 \qquad \forall\, f \in \mathcal{C}\big(([0,1]^2)^{\mathbb{N}_0}, \mathbb{R}\big). \tag{7.75}$$

(b) *There exists a sequence $\bar{L}(N)$ satisfying $\lim_{N\to\infty} \bar{L}(N) = \infty$ and $\lim_{N\to\infty} \bar{L}(N)/N = 0$ such that*

$$\begin{aligned} \lim_{k\to\infty} &\big|\mathbb{E}\big[f\big(Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big) - f\big(Z^{\mu_{N_k}}(\bar{L}(N_k))\big)\big|\big]\\ &+ \big|\mathbb{E}\big[f\big(Z^{\mu_{N_k}}(\bar{L}(N_k))\big)\big] - \mathbb{E}\big[f\big(Z^\mu(\bar{L}(N_k))\big)\big]\big| = 0 \qquad \forall\, f \in \mathcal{C}\big(([0,1]^2)^{\mathbb{N}_0}, \mathbb{R}\big). \end{aligned} \tag{7.76}$$

Proof. As in the proof of Lemma 6.2.11, we can construct $(z_i^{\mu_N})_{i\in\mathbb{N}_0}$ and $(z_i^\mu)_{i\in\mathbb{N}_0}$ on one probability space. Then

$$\lim_{N\to\infty} y_{i,1}^{\mu_N}(0) = y_{i,1}^\mu(0) \text{ a.s.} \tag{7.77}$$

and

$$\lim_{N\to\infty} \mathbb{E}[|\bar{\Theta}^{[N]} - \Theta|] = 0. \tag{7.78}$$

Via a similar coupling as in Lemma (7.2.7), it follows via Itô-calculus that (7.75) holds. Combining (7.64), (7.68), (7.77) and (7.78), we obtain, via a similar construction as in the proof of Lemma 6.2.11, a sequence $\bar{L}(N)$ such that

$$\begin{aligned}&\lim_{N\to\infty}\mathbb{E}[|\Delta_i^N(\bar{L}(N))|+K_0|\delta_{i,0}^N(\bar{L}(N))|]+K_1|\delta_{i,1}^N(\bar{L}(N))|]\\&+\mathbb{E}[|\Delta_i^{\mu_N}(\bar{L}(N))|+K_0|\delta_{i,0}^{\mu_N}(\bar{L}(N))|]+K_1|\delta_{i,1}^{\mu_N}(\bar{L}(N))|]=0.\end{aligned}\tag{7.79}$$

As in the proof of Lemma 6.2.11, we can again use Lipschitz functions to conclude (7.76). □

Lemma 7.2.10 (Coupling of finite systems). *Let*

$$Z^{[N],1}=(X^{[N],1},Y_0^{[N],1},Y_1^{[N],1})\tag{7.80}$$

be the finite system evolving according to (7.4) *starting from an exchangeable initial measure. Let* $\mu^{[N],1}$ *be the measure obtained by periodic continuation of the configuration of* $Z^{[N],1}(0)$*. Similarly, let*

$$Z^{[N],2}=(X^{[N],2},Y_0^{[N],2},Y_1^{[N],2})\tag{7.81}$$

be the finite system evolving according to (7.4) *starting from an exchangeable initial measure. Let* $\mu^{[N],2}$ *be the measure obtained by periodic continuation of the configuration of* $Z^{[N],2}(0)$*. Let* $\tilde{\mu}$ *be any weak limit point of the sequence of measures* $(\mu^{[N],1}\times\mu^{[N],2})_{N\in\mathbb{N}}$*. Define the random variables* $\bar{\Theta}^{[N],1}$ *and* $\bar{\Theta}^{[N],2}$ *on* $(([0,1]^3)^{\mathbb{N}_0}\times([0,1]^3)^{\mathbb{N}_0},\mu^{[N],1}\times\mu^{[N],2})$ *and* $\bar{\Theta}_1$ *and* $\bar{\Theta}_2$ *on* $(([0,1]^3)^{\mathbb{N}_0}\times([0,1]^3)^{\mathbb{N}_0},\bar{\mu})$ *by*

$$\begin{aligned}&\bar{\Theta}^{[N],1}=\frac{1}{N}\sum_{i\in[N]}\frac{x_i^{[N],1}+K_0y_{i,0}^{[N],1}}{1+K_0}, &&\bar{\Theta}^{[N],2}=\frac{1}{N}\sum_{i\in[N]}\frac{x_i^{[N],2}+K_0y_{i,0}^{[N],2}}{1+K_0},\\&\bar{\Theta}^1=\lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]}\frac{x_i^1+K_0y_{i,0}^1}{1+K_0}, &&\bar{\Theta}^2=\lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]}\frac{x_i^2+K_0y_{i,0}^2}{1+K_0},\end{aligned}\tag{7.82}$$

and let $(\bar{\Theta}^{(1),[N],1}(t))_{t\geq0}$ *and* $(\bar{\Theta}^{(1),[N],2}(t))_{t\geq0}$ *be defined as in* (7.14) *for* $Z^{[N],1}$*, respectively,* $Z^{[N],2}$*. Suppose that*

$$\lim_{N\to\infty}\sup_{0\leq t\leq L(N)}\left(\left|\bar{\Theta}^{[N],k}(0)-\bar{\Theta}^{[N],k}(t)\right|\right)=0\quad\textit{in probability},\qquad k\in\{1,2\},\tag{7.83}$$

and suppose that $\tilde{\mu}(\{\bar{\Theta}_1=\bar{\Theta}_2,\,Y_1^1=Y_1^2\})=1$*. Then, for any* $t(N)\to\infty$*,*

$$\begin{aligned}&\lim_{N\to\infty}\mathbb{E}\big[|x_i^{[N],1}(t(N))-x_i^{[N],2}(t(N))|+K_0|y_{i,0}^{[N],1}(t(N))-y_{i,0}^{[N],2}(t(N))|\\&\quad+K_1|y_{i,1}^{[N],1}(t(N))-y_{i,1}^{[N],2}(t(N))|\big]=0.\end{aligned}\tag{7.84}$$

Proof. Via standard Itô-calculus we obtain from (7.4) that

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[|x_i^{[N],1}(t)-x_i^{[N],2}(t)|+K_0|y_{i,0}^{[N],1}(t)-y_{i,0}^{[N],2}(t)|+K_1|y_{i,1}^{[N],1}(t)-y_{i,1}^{[N],2}(t)|\right]\\
&=-\frac{2c}{N}\sum_{j\in[N]}\mathbb{E}\left[|x_j^{[N],1}(t)-x_j^{[N],2}(t)|1_{\{\mathrm{sgn}\,(x_j^{[N],1}(t)-x_j^{[N],2}(t))\neq\mathrm{sgn}\,(x_i^{[N],1}(t)-x_i^{[N],2}(t))\}}\right]\\
&\quad-2K_0e_0\,\mathbb{E}\Bigg[|x_i^{[N],1}(t)-x_i^{[N],2}(t)|+K|y_{i,0}^{[N],1}(t)-y_{i,0}^{[N],2}(t)|\\
&\qquad\times 1_{\{\mathrm{sgn}\,(x_i^{[N],1}(t)-x_i^{[N],2}(t))\neq\mathrm{sgn}\,(y_{i,0}^{[N],1}(t)-y_{i,0}^{[N],2}(t))\}}\Bigg]\\
&\quad-2\frac{K_1e_1}{N}\,\mathbb{E}\Bigg[|x_i^{[N],1}(t)-x_i^{[N],2}(t)|+K_1|y_{i,0}^{[N],1}(t)-y_{i,0}^{[N],2}(t)|\\
&\qquad\times 1_{\{\mathrm{sgn}\,(x_i^{[N],1}(t)-x_i^{[N],2}(t))\neq\mathrm{sgn}\,(y_{i,1}^{[N],1}(t)-y_{i,1}^{[N],2}(t))\}}\Bigg].
\end{aligned}
\tag{7.85}
$$

Therefore, for all $N\in\mathbb{N}$,

$$
t\mapsto\mathbb{E}\left[|x_i^{[N],1}(t)-x_i^{[N],2}(t)|+K_0|y_{i,0}^{[N],1}(t)-y_{i,0}^{[N],2}(t)|+K_1|y_{i,1}^{[N],1}(t)-y_{i,1}^{[N],2}(t)|\right]
\tag{7.86}
$$

is a decreasing function. Hence we can use the same strategy as in the proof of Lemma 6.2.13 to finish the proof. □

• Proof of Proposition 7.2.3

Proof. We follow a similar argument as in the proof of Proposition 6.2.4. Let $L(N)$ satisfy $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N=0$. Let μ_N be the measure on $([0,1]^3)^{\mathbb{N}_0}$ obtained by periodic continuation of $\mathcal{L}[Z^{[N]}(Ns-L(N))]$. Note that $([0,1]^3)^{\mathbb{N}_0}$ is compact. Hence, letting $(N_k)_{k\in\mathbb{N}}$ be the subsequence in Proposition 7.2.3, we can pass to a possibly further subsequence and obtain

$$
\lim_{k\to\infty}\mu_{N_k}=\mu. \tag{7.87}
$$

Since we assumed that $\mathcal{L}[Z^{[N]}(0)]$ is exchangeable and the dynamics preserve exchangeability, the measures μ_{N_k} are translation invariant and also the limiting law μ is translation invariant.

Let $\phi=(\phi_1,\phi_2)$ be defined as in (7.69) in Lemma 7.2.8. Then we can condition on $\phi=(\phi_1,\phi_2)$ and write

$$
\mu=\int_{[0,1]^2}\mu_\rho\,\mathrm{d}\Lambda(\rho), \tag{7.88}
$$

where $\Lambda(\cdot)=\mathcal{L}[\phi]=\mathcal{L}[(\phi_1,\phi_2)]$ and $\rho=(\rho_1,\rho_2)$. By assumption we know that

$$
\lim_{k\to\infty}\mathcal{L}\left[\mathbf{\Theta}^{\mathrm{eff},(1),[N_k]}(N_ks)\right]=P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\cdot) \tag{7.89}
$$

and

$$\lim_{k\to\infty} \mathcal{L}\left[\sup_{0\le t\le L(N_k)} \left|\bar{\Theta}^{[N_k]}(N_k s) - \bar{\Theta}^{[N_k]}(N_k s - t)\right| + \left|\Theta_{y_1}^{[N_k]}(N_k s) - {\Theta_{y_1}}^{[N_k]}(N_k s - t)\right|\right] \\ = \delta_0. \tag{7.90}$$

Hence

$$\lim_{k\to\infty} \mathcal{L}\left[\mathbf{\Theta}^{\mathrm{eff},(1),[N_k]}(N_k s - L(N_k))\right] = P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\cdot). \tag{7.91}$$

Recall that

$$\Lambda(\cdot) = \mathcal{L}\left[\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i\in[n]}\frac{x_i + K y_{i,0}}{1+K}, \frac{1}{n}\sum_{i\in[n]} y_{i,1}\right)\right] \quad \text{on } (\mu, ([0,1]^2)^{\mathbb{N}_0}). \tag{7.92}$$

By Lemma 6.2.10, if

$$\phi^{N_k} = (\phi_1^{N_k}, \phi_2^{N_k}) = \left(\frac{1}{N_k}\sum_{i\in[N_k]}\frac{x_i + K y_{i,0}}{1+K}, \frac{1}{N_k}\sum_{i\in[N_k]} y_{i,1}^{[N_k]}\right) \text{ on } (\mu_{N_k}, ([0,1]^3)^{\mathbb{N}_0}), \tag{7.93}$$

then $\lim_{k\to\infty}\mathcal{L}[\phi^{N_k}] = \mathcal{L}[\phi]$. Taking the subsequence $(\mu_{N_k})_{k\in\mathbb{N}}$, we get $\Lambda(\cdot) = P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\cdot)$, and hence

$$\mu = \int_{[0,1]} \mu_\rho \, \mathrm{d}P_s(\rho). \tag{7.94}$$

Let $\bar{L}(N)$ be the sequence constructed in Lemma 7.2.9[b]. By construction we can require that $\bar{L}(N) \le L(N)$ for all $N \in \mathbb{N}$. Write

$$\begin{aligned} &\mathcal{L}\big[Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big] \\ &= \mathcal{L}\big[Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big] - \mathcal{L}\big[Z^{\mu_{N_k}}(\bar{L}(N_k))\big], \\ &\quad + \mathcal{L}\big[Z^{\mu_{N_k}}(\bar{L}(N_k))\big] - \mathcal{L}\big[Z^{\mu}(\bar{L}(N_k))\big] \\ &\quad + \mathcal{L}\big[Z^{\mu}(\bar{L}(N_k))\big]. \end{aligned} \tag{7.95}$$

By Lemma 7.2.9 the first and second differences tend to zero as $k \to \infty$. Hence

$$\lim_{k\to\infty} \mathcal{L}\big[Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big] = \mathcal{L}\big[Z^{\mu}(\bar{L}(N_k))\big]. \tag{7.96}$$

By (7.88),

$$\mathcal{L}\big[Z^{\mu}(\bar{L}(N_k))\big] = \int_{[0,1]^2} \mathcal{L}\big[Z^{\mu_\rho}(\bar{L}(N_k))\big]\, P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\mathrm{d}\rho). \tag{7.97}$$

For the infinite system $(Z^{\mu_\rho}(t))_{t\ge0} = \big(X^{\mu_\rho}(t), Y_0^{\mu_\rho}(t), Y_1^{\mu_\rho}(t)\big)_{t\ge0}$ we have

$$Y_1^{\mu}(t) = Y_1^{\mu}(0)\; a.s. \tag{7.98}$$

and hence, since $\lim_{k\to\infty}\bar{L}(N_k)/N_k = 0$ by (7.44),

$$\lim_{k\to\infty} \mathcal{L}[Y_1^{\mu_\rho}(\bar{L}(N_k))] = \mathcal{L}[Y_1^{\mu_\rho}(0)] \;\forall\, \rho \in [0,1]. \tag{7.99}$$

Therefore

$$\lim_{k\to\infty} \mathcal{L}[Y_1^{\mu_\rho}(\bar{L}(N_k))] = P^\rho_{Y_1(s)}(\cdot) \tag{7.100}$$

and

$$\begin{aligned}\mathcal{L}\big[X^{\mu_\rho}(\bar{L}(N_k)), Y_0^{\mu_\rho}(\bar{L}(N_k)), Y_1^{\mu_\rho}(\bar{L}(N_k))\big] \\ = \int \mathcal{L}\big[X_1^{\mu_\rho}(\bar{L}(N_k)), Y_0^{\mu_\rho}(\bar{L}(N_k)), \mathbf{y}_1\big] \mathrm{d}P^\rho_{Y_1(s)}(\mathrm{d}\mathbf{y}_1).\end{aligned} \tag{7.101}$$

Hence, since $\lim_{k\to\infty} \bar{L}(N_k) = \infty$, by Lemma 6.2.5 we have

$$\begin{aligned}\lim_{k\to\infty} \mathcal{L}\big[Z^{\mu_\rho}(\bar{L}(N_k))\big] &= \lim_{k\to\infty} \mathcal{L}\big[X_1^{\mu_\rho}(\bar{L}(N_k)), Y_0^{\mu_\rho}(\bar{L}(N_k)), Y_1^{\mu_\rho}(\bar{L}(N_k))\big] \\ &= \int \nu_{\rho,\mathbf{y}_1} P^\rho_{Y_1(s)}(\mathrm{d}\mathbf{y}_1).\end{aligned} \tag{7.102}$$

Therefore, by (6.109), (7.97) and Lemma 6.2.6,

$$\lim_{k\to\infty} \mathcal{L}\big[Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))\big] = \int_{[0,1]} P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\mathrm{d}\rho) \int \nu_{\rho,\mathbf{y}_1} P^\rho_{Y_1(s)}(\mathrm{d}\mathbf{y}_1). \tag{7.103}$$

To finish the proof, we proceed as in the proof of Proposition 6.2.4 and invoke Lemma 7.2.10. Let $Z^{[N],1} = (X^{[N],1}, Y_0^{[N],1}, Y_1^{[N],1})$ be the finite system starting from

$$\mathcal{L}\big[Z^{[N]}(Ns - L(N))\big] = \mathcal{L}\big[X^{[N]}(Ns - L(N)), Y_0^{[N]}(Ns - L(N)), Y_1^{[N]}(Ns - L(N))\big]. \tag{7.104}$$

Let $(\bar{L}(N))_{N\in\mathbb{N}}$ be the sequence constructed in Lemma 7.2.9. Let $Z^{[N],2} = (X^{[N],2}, Y_0^{[N],2}, Y_1^{[N],2})$ be the finite system starting from

$$\mathcal{L}\big[X^{[N]}(Ns - \bar{L}(N))\big] = \mathcal{L}\left[X^{[N]}(Ns - \bar{L}(N)), Y_0^{[N]}(Ns - \bar{L}(N)), Y_1^{[N]}(Ns - \bar{L}(N))\right]. \tag{7.105}$$

Choose for $t(N)$ in Lemma 7.2.10 the sequence $\bar{L}(N)$. Let $\mu^{[N],1}$ be defined by the periodic continuation of the configuration of $Z^{[N]}(Ns - L(N))$ and $\mu^{[N],2}$ be defined by periodic continuation of the configuration of $Z^{[N]}(Ns - \bar{L}(N))$. Define Θ_1 and Θ_2 according to (6.85), where under $\mu^{[N],2}$ we replace $L(N)$ by $\bar{L}(N)$. Then, by the assumptions in (7.44),

$$\begin{aligned}\lim_{k\to\infty} |\Theta^{(1),[N_k],1} - \Theta^{(1),[N_k],2}| &= \lim_{k\to\infty} |\Theta^{N_k}(N_k s - L(N_k)) - \Theta^{N_k}(N_k s - \bar{L}(N_k))| \\ &= 0 \text{ in probability.}\end{aligned} \tag{7.106}$$

Using 7.63 we see that also, for all $i \in [N]$,

$$\begin{aligned}\lim_{k\to\infty} |y_{i,1}^{[N_k],1}(0) - y_{i,1}^{[N_k],2}(0)| &= \lim_{k\to\infty} |y_{i,1}^{[N_k]}(N_k s - L(N_k)) - y_{i,1}^{[N_k]}(N_k s - \bar{L}(N_k))| \\ &= 0 \text{ in probability.}\end{aligned} \tag{7.107}$$

Therefore, if μ is any weak limit point of the sequence $\left(\mu^{[N_k],1} \times \mu^{[N_k],2}\right)_{k\in\mathbb{N}}$, then

$$\mu(\{\Theta_1 = \Theta_2,\, Y_1^1 = Y_1^2\}) = 1. \tag{7.108}$$

Hence, by possibly passing to a further subsequence, we can now apply Lemma 7.2.10 to obtain, for all i,

$$\begin{aligned}\lim_{k\to\infty} \mathbb{E}\Big[&|x_i^{[N_k],1}(\bar{L}(N_k)) - x_i^{[N_k],2}(\bar{L}(N_k))| \\ &+ K_0\,|y_{i,0}^{[N_k],1}(\bar{L}(N_k)) - y_{i,0}^{[N_k],2}(\bar{L}(N_k))| \\ &+ K_1\,|y_{i,1}^{[N_k],1}(\bar{L}(N_k)) - y_{i,1}^{[N_k],2}(\bar{L}(N_k))|\Big] = 0.\end{aligned} \tag{7.109}$$

Hence

$$\lim_{N\to\infty} \Big(\mathcal{L}[Z^{[N],1}(\bar{L}(N_k))] - \mathcal{L}[Z^{[N],2}(\bar{L}(N_k))]\Big) = \delta_0 \tag{7.110}$$

and therefore

$$\lim_{k\to\infty} \mathcal{L}\big(Z^{[N_k]}(N_k s)\big) = \int_{[0,1]} P_{\mathbf{\Theta}^{\mathrm{eff}}(s)}(\mathrm{d}\rho) \int \nu_{\rho,\mathbf{y}_1} P^{\rho}_{Y_1(s)}(\mathrm{d}\mathbf{y}_1). \tag{7.111}$$

This concludes the proof of Proposition 7.2.3. □

Like for the one-colour mean-field system, Proposition 7.2.3 and Lemmas 7.2.4–7.2.10 give rise to the following corollary, which will be important to derive the evolution of the 1-blocks on time scale Ns.

Corollary 7.2.11. *Fix $s > 0$. Let μ_N be the measure obtained by periodic continuation of*

$$Z^{[N]}(Ns-L(N)) = (X^{[N]}(Ns-L(N)), Y_0^{[N]}(Ns-L(N)), Y_1^{[N]}(Ns-L(N))), \tag{7.112}$$

and let μ be a weak limit point of the sequence $(\mu_N)_{N\in\mathbb{N}}$. Let

$$\Theta = \lim_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} \frac{x_i^{\mu} + K y_i^{\mu}}{1+K} \qquad \text{in } L_2(\mu), \tag{7.113}$$

and let $(Z^{\nu_\Theta}(t))_{t>0} = (X^{\nu_\Theta}(t), Y_0^{\nu_\Theta}(t), Y_1^{\nu_\Theta}(t))_{t>0}$ be the infinite system evolving according to (7.74) starting from its equilibrium measure. Consider the finite system $Z^{[N]}$ as a system on $([0,1]^3)^{\mathbb{N}_0}$ by periodic continuation. Construct $(Z^{[N]}(t))_{t>0}$ and $(Z^{\nu_\Theta}(t))_{t>0}$ on one probability space. Then, for all $t \geq 0$,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\left[\left|x_i^{[N]}(Ns+t) - x_i^{\nu_\Theta}(t)\right|\right] + K_0\,\mathbb{E}\left[\left|y_{i,0}^{[N]}(Ns+t) - y_{i,0}^{\nu_\Theta}(t)\right|\right]& \\ + K_1\,\mathbb{E}\left[\left|y_{i,1}^{[N]}(Ns+t) - y_{i,1}^{\nu_\Theta}(t)\right|\right] = 0 \qquad \forall\, i \in [N].&\end{aligned} \tag{7.114}$$

Proof. Proceed as in the proof of Corollary 6.3.1, but use the setup of the two-colour mean-field system and therefore replace Proposition 6.2.4, Lemma 6.2.11 and Lemma 6.2.13 by, respectively Proposition 7.2.3, Lemma 7.2.9 and Lemma 7.2.10. □

Step 4: Limiting evolution of the 1-blocks.

Lemma 7.2.12 (Limiting evolution of the 1-blocks). *Let* $(z_1^{\text{eff}}(s))_{s>0}$ *be the process defined in* (7.21) *with initial state*

$$z_1^{\text{eff}}(0) = (\vartheta_0, \theta_{y_1}). \tag{7.115}$$

Then

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns)\right)_{s>0}\right] = \mathcal{L}\left[(z_1^{\text{eff}}(s))_{s>0}\right]. \tag{7.116}$$

Proof. By [72], the SSDE in (7.21) has a unique strong solution. Therefore the process $(z_1^{\text{eff}}(s))_{s>0}$ is Markov. Its generator G is given by

$$G = \frac{K_1 e_1}{1+K_0}\,(y-x)\,\frac{\partial}{\partial x} + e_1(x-y)\,\frac{\partial}{\partial y} + \frac{1}{(1+K_0)^2}\,(\mathcal{F}g)(x)\,\frac{\partial^2}{\partial x^2}, \tag{7.117}$$

and hence $(z_1^{\text{eff}}(s))_{s\geq 0}$ solves the martingale problem for G. We will use [49, Theorem 3.3.1], to prove that (7.116) holds.

Define

$$(\vartheta_0^N, \vartheta_{y_1}^N) = \left(\bar{\Theta}^{(1),[N]}(0), \Theta_{y_1}^{(1),[N]}(0)\right). \tag{7.118}$$

Since we start from an i.i.d. law, by the law of large numbers we have that

$$\lim_{N\to\infty} \mathbf{\Theta}^{\text{eff},(1),[N]}(0) = \lim_{N\to\infty} (\vartheta_0^N, \vartheta_{y_1}^N) = (\vartheta_0, \theta_{y_1}) \qquad a.s. \tag{7.119}$$

By the SSDE in (7.37) and an optional sampling argument, we have, for all $N \in \mathbb{N}$,

$$\lim_{s\downarrow 0} \left(\bar{\Theta}^{(1),[N]}(Ns), \Theta^{(1),[N]}(Ns)\right) = (\vartheta_0^N, \vartheta_{y_1}^N) \quad \text{a.s.} \tag{7.120}$$

Therefore we can continuously extend the process $(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns))_{s>0}$ to 0 and, in particular,

$$\lim_{N\to\infty} \mathcal{L}\left[\mathbf{\Theta}^{\text{eff},(1),[N]}(0)\right] = \mathcal{L}\left[z_1^{\text{eff}}(0)\right]. \tag{7.121}$$

Since we already showed that the processes

$$\left(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns)\right)_{s>0} \tag{7.122}$$

are $\mathcal{D}$-semimartingales, and are trivially bounded, we are left to show that

$$\lim_{N\to\infty} \int_0^s \mathrm{d}r\, \mathbb{E}\left[\left|G_\dagger^{(1),[N]}\big(f, \mathbf{\Theta}^{\text{eff},(1),[N]}(Nr), r, \cdot\big) - (Gf)\big(\mathbf{\Theta}^{\text{eff},(1),[N]}(Nr)\big)\right|\right] = 0. \tag{7.123}$$

Here, $G_\dagger^{(1),[N]}$ is the operator defined in (7.39). Since we are working on the space $\mathcal{C}^*$ of polynomials on $[0,1]^2$, all derivatives of $f \in \mathcal{C}^*$ are bounded. Hence, by dominated convergence, it is enough to prove that, for all $s > 0$,

$$\lim_{N\to\infty} \mathbb{E}^{[N]}\left[\left|G_\dagger^{(1),[N]}\big(f, \mathbf{\Theta}^{\text{eff},(1),[N]}(Ns), s, \cdot\big) - (Gf)\big(\mathbf{\Theta}^{\text{eff},(1),[N]}(Ns)\big)\right|\right] = 0. \tag{7.124}$$

Note that

$$
\begin{aligned}
&\mathbb{E}\Big[\Big|G_{\dagger}^{(1),[N]}\big(f,\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns),s,\cdot\big)-(Gf)\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\Big|\Big]\\
&=\mathbb{E}\Bigg[\Bigg|\frac{K_1e_1}{1+K_0}\Big[\Theta_{y_1}^{(1),[N]}(Ns)-\frac{1}{N}\sum_{i\in[N]}x_i(Ns,\omega)\Big]\frac{\partial f}{\partial x}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\\
&\qquad+e_1\Big[\frac{1}{N}\sum_{i\in[N]}x_i(Ns,\omega)-\Theta_{y_1}^{(1),[N]}(Ns)\Big]\frac{\partial f}{\partial y}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\\
&\qquad+\frac{1}{(1+K_0)^2}\frac{1}{N}\sum_{i\in[N]}g(x_i(Ns,\omega))\frac{\partial^2 f}{\partial x^2}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\\
&\qquad-\frac{K_1e_1}{1+K_0}\big[\Theta_{y_1}^{(1),[N]}(Ns)-\bar{\Theta}^{(1),[N]}(Ns)\big]\frac{\partial f}{\partial x}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\\
&\qquad-e_1\big[\bar{\Theta}^{(1),[N]}(Ns)-\Theta_{y_1}^{(1),[N]}(Ns)\big]\frac{\partial f}{\partial y}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\\
&\qquad-\frac{1}{(1+K_0)^2}(\mathcal{F}g)(\bar{\Theta}^{(1),[N]}(Ns))\frac{\partial^2 f}{\partial x^2}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\Bigg|\Bigg].
\end{aligned}
\tag{7.125}
$$

Hence

$$
\begin{aligned}
&\lim_{N\to\infty}\mathbb{E}\Big[\Big|G_{\dagger}^{(1),[N]}\big(f,\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns),s,\cdot\big)-(Gf)\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\Big|\Big]\\
&\leq\lim_{N\to\infty}\mathbb{E}\left[\frac{K_1e_1}{1+K_0}\left|\bar{\Theta}^{(1),[N]}(Ns)-\frac{1}{N}\sum_{i\in[N]}x_i(Ns,\omega)\right|\left|\frac{\partial f}{\partial x}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\right|\right]\\
&+\lim_{N\to\infty}\mathbb{E}\left[e_1\left|\frac{1}{N}\sum_{i\in[N]}x_i(Ns,\omega)-\bar{\Theta}^{(1),[N]}(Ns)\right|\left|\frac{\partial f}{\partial y}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\right|\right]\\
&+\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{(1+K_0)^2}\left|\frac{1}{N}\sum_{i\in[N]}g(x_i(Ns,\omega))-(\mathcal{F}g)\big(\bar{\Theta}^{(1),[N]}(Ns)\big)\right|\left|\frac{\partial^2 f}{\partial x^2}\big(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\big)\right|\right].
\end{aligned}
\tag{7.126}
$$

Note that each of the derivatives is bounded by aconstant because we work on $\mathcal{C}^*$. The first and the second term tend to zero by Lemma 7.2.6. For the third term we can use a similar argument as used in (6.198), since we showed Lemmas 7.2.4–7.2.10 for the single components in the mean-field system with two colours. □

Step 5: Evolution of the averages in the Meyer-Zheng topology. In this section we prove the following proposition

Proposition 7.2.13 (Convergence in the Meyer-Zheng topology). *Suppose that the effective estimator process defined in* (7.15) *satisfies*

$$
\lim_{N\to\infty}\mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{eff},(1),[N]}(Ns)\right)_{s>0}\right]=\mathcal{L}\left[\left(z_1^{\mathrm{eff}}(s)\right)_{s>0}\right].
\tag{7.127}
$$

Then for the averages in (7.7),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(z_1^{[N]}(s)\right)_{s>0}\right] = \mathcal{L}\left[(z_1(s))_{s>0}\right] \quad \text{in the Meyer-Zheng topology,} \tag{7.128}$$

where the limiting process $(z_1(s))_{s>0}$ *is defined as in* (7.20).

To prove Proposition 7.2.13 we need the following characterisation of continuous functions in the Meyer-Zheng topology

Lemma 7.2.14 (Convergence of marginals in the Meyer-Zheng topology). *Let* (E,d) *be a Polish space with metric* d*. Suppose that* $(X_n(s),Y_n(s))_{s>0}$ *is a stochastic process with state space* E^2*. If*

$$\lim_{n\to\infty} \mathcal{L}\left[(X_n(s),Y_n(s))_{s>0}\right] = \mathcal{L}\left[(X(s),Y(s))_{s>0}\right] \text{ in the Meyer-Zheng topology,} \tag{7.129}$$

then the marginals also converge in the Meyer-Zheng topology, i.e.,

$$\begin{aligned} &\lim_{n\to\infty} \mathcal{L}\left[(X_n(s))_{s>0}\right] = \mathcal{L}\left[(X(s))_{s>0}\right] \text{ in the Meyer-Zheng topology,}\\ &\lim_{n\to\infty} \mathcal{L}\left[(Y_n(s))_{s>0}\right] = \mathcal{L}\left[(Y(s))_{s>0}\right] \text{ in the Meyer-Zheng topology.} \end{aligned} \tag{7.130}$$

The proof of Lemma 7.2.14 is given in Appendix B.2.3.

Proof of Proposition 7.2.13. By Lemma 7.2.6, we have that, for all $s>0$,

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\bar\Theta^{[N]}(Ns) - x_1^{[N]}(s)\right|\right] = 0 \tag{7.131}$$

and

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\bar\Theta^{[N]}(Ns) - y_{0,1}^{[N]}(s)\right|\right] = 0. \tag{7.132}$$

Applying Lemmas 6.2.19, 6.2.20 and 6.2.21, like in the proof of Proposition 6.2.18, we obtain

$$\begin{aligned} &\lim_{N\to\infty} \mathcal{L}\left[\left(x_1^{[N]}(s), y_{0,1}^{[N]}(s), \bar\Theta^{[N]}(Ns), \Theta_{y_{1,1}}^{[N]}(Ns)\right)_{s>0}\right]\\ &= \mathcal{L}\left[\left(x_1^{\mathrm{eff}}(s), x_1^{\mathrm{eff}}(s), x_1^{\mathrm{eff}}(s), y_1^{\mathrm{eff}}(s)\right)_{s>0}\right] \text{ in the Meyer-Zheng topology.} \end{aligned} \tag{7.133}$$

Applying Lemma 7.2.14, we get the claim. □

Step 6: Proof of the two-colour mean-field finite-systems scheme.

Proof. The proof of Proposition 7.1.2(a) follows directly from Lemma 7.2.12. The proof of Proposition 7.1.2(b) is a consequence of Proposition 7.1.2(d). The proof of Prosition 7.1.2(c) follows from Proposition 7.1.2(a) by applying Proposition 7.2.13. The proof of Proposition (7.1.2)(d) follows by the same argument as used in the proof Proposition 6.2.1(c) in Section 6.3.4. In this argument we have to replace the two-component system $Z^{[N]}(Ns+t) = (X^{[N]}(Ns+t), Y^{[N]}(Ns+t))_{t\geq 0}$ by the three-component system $Z^{[N]}(Ns+t) = (X^{[N]}(Ns+t), Y_0^{[N]}(Ns+t), Y_1^{[N]}(Ns+t))_{t\geq 0}$

and use the infinite system defined in 7.11 instead of the infinite system defined in (6.42). We now use the two-dimensional transition kernel in (7.30), which controls the transition probabilities of the two-dimensional process $(\bar{\Theta}^{(1)}(s), \Theta_{y_1}^{(1)}(s))_{s>0}$, instead of the one-dimensional transition kernel in (6.59). $\square$

CHAPTER 8

Two-level three-colour mean-field system

To get a proper understanding of how the migration comes into play on different space-time scales, we next look at a two-level mean-field system where the geographic space consists of two layers and the seed-bank consist of three layers, corresponding three colours 0, 1, 2. In Section 8.1 we give the set-up of the two-level three-colour mean-field model. In Section 8.2 we give a scheme to prove the analysis of the two-level three-colour mean-field model. Finally, in Section 8.3 we prove the steps of the scheme given in Section 8.2.

§8.1 Two-level three-colour mean-field finite-systems scheme

We consider a restricted version of the SSDE in (4.20) on the finite geographic space

$$[N^2] = \{0, 1, \dots, N^2 - 1\}, \qquad N \in \mathbb{N}. \tag{8.1}$$

This space should be interpreted as grouping the N-blocks consisting of N colonies together, i.e.

$$[N^2] = \bigcup_{l=0}^{N-1} \{Nl, Nl+1, \cdots, Nl+N-1\}. \tag{8.2}$$

With this interpretation we can use the metric $d_{[N^2]}$ that is induced by the metric d_{Ω_N} on the hierarchical group Ω_N (recall (4.4)). The migration kernel $a^{\Omega_N}(\cdot,\cdot)$ is restricted to $[N^2]$ by setting all migration rates outside the 2-block equal to 0, i.e., $c_k = 0$ for all $k \geq 2$. Hence the migration kernel is given by

$$a^{[N^2]}(i,j) = 1_{\{d_{[N^2]}(i,j) \leq 1\}} \frac{c_0}{N} + \frac{c_1}{N^3}, \tag{8.3}$$

where $c_0, c_1 \in (0,\infty)$ are constants. The seed-bank of the restricted system consists of *three colours*, labeled 0 1 and 2, with exchange rates given by $K_0 e_0, e_0$, $\frac{K_1 e_1}{N}$, $\frac{e_1}{N}$ and $\frac{K_2 e_2}{N^2}$, $\frac{e_2}{N^2}$ respectively. The state space of the restricted system is

$$S = \mathfrak{s}^{[N^2]}, \qquad \mathfrak{s} = [0,1] \times [0,1]^3, \tag{8.4}$$

and the restricted system is denoted by

$$
\begin{aligned}
(Z^{[N^2]}(t))_{t\geq 0} &= \Big(X^{[N^2]}(t), \Big(Y_0^{[N^2]}(t), Y_1^{[N^2]}(t), Y_2^{[N^2]}(t)\Big)\Big)_{t\geq 0}, \\
\Big(X^{[N^2]}(t), \Big(Y_0^{[N^2]}(t), Y_1^{[N^2]}(t), Y_2^{[N^2]}(t)\Big)\Big) &= \Big(x_i^{[N^2]}(t), \Big(y_{i,0}^{[N^2]}(t), y_{i,1}^{[N^2]}(t), y_{i,2}^{[N^2]}(t)\Big)\Big)_{i\in[N^2]}.
\end{aligned}
\tag{8.5}
$$

The components of the restricted system $(Z^{[N^2]}(t))_{t\geq 0}$ evolve according to the SSDE

$$
\begin{aligned}
\mathrm{d}x_i^{[N^2]}(t) &= \frac{c_0}{N} \sum_{j\in[N^2]} 1_{\{d_{[N^2]}(i,j)\leq 1\}} [x_j^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t \\
&\quad + \frac{c_1}{N^3} \sum_{j\in[N^2]} [x_j^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t + \sqrt{g(x_i^{[N^2]}(t))}\,\mathrm{d}w_i(t) \\
&\quad + K_0 e_0\,[y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t \\
&\quad + \frac{K_1 e_1}{N}\,[y_{i,1}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t \\
&\quad + \frac{K_2 e_2}{N^2}\,[y_{i,2}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t, \\
\mathrm{d}y_{i,0}^{[N^2]}(t) &= e_0\,[x_i^{[N^2]}(t) - y_{i,0}^{[N^2]}(t)]\,\mathrm{d}t, \\
\mathrm{d}y_{i,1}^{[N^2]}(t) &= \frac{e_1}{N}\,[x_i^{[N^2]}(t) - y_{i,1}^{[N^2]}(t)]\,\mathrm{d}t, \\
\mathrm{d}y_{i,2}^{[N^2]}(t) &= \frac{e_2}{N^2}\,[x_i^{[N^2]}(t) - y_{i,2}^{[N^2]}(t)]\,\mathrm{d}t, \qquad i\in[N^2],
\end{aligned}
\tag{8.6}
$$

which is a special case of (4.20). By [67, Theorem 3.1], the SSDE in (8.6) is the unique solution. It is important to note that we can write the SSDE also

$$
\begin{aligned}
\mathrm{d}x_i^{[N^2]}(t) &= c_0 \left[\frac{1}{N} \sum_{j\in[N]_i} x_j^{[N^2]}(t) - x_i^{[N^2]}(t)\right] \mathrm{d}t \\
&\quad + \frac{c_1}{N} \left[\frac{1}{N^2} \sum_{j\in[N^2]} x_j^{[N^2]}(t) - x_i^{[N^2]}(t)\right] \mathrm{d}t + \sqrt{g(x_i^{[N^2]}(t))}\,\mathrm{d}w_i(t) \\
&\quad + K_0 e_0\,[y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t + \frac{K_1 e_1}{N}\,[y_{i,1}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t \\
&\quad + \frac{K_2 e_2}{N^2}\,[y_{i,2}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t, \\
\mathrm{d}y_{i,0}^{[N^2]}(t) &= e_0\,[x_i^{[N^2]}(t) - y_{i,0}^{[N^2]}(t)]\,\mathrm{d}t, \\
\mathrm{d}y_{i,1}^{[N^2]}(t) &= \frac{e_1}{N}\,[x_i^{[N^2]}(t) - y_{i,1}^{[N^2]}(t)]\,\mathrm{d}t, \\
\mathrm{d}y_{i,2}^{[N^2]}(t) &= \frac{e_2}{N^2}\,[x_i^{[N^2]}(t) - y_{i,2}^{[N^2]}(t)]\,\mathrm{d}t, \qquad i\in[N^2],
\end{aligned}
\tag{8.7}
$$

where $[N]_i$ denotes the set of colonies in the 1-block around site $i\in[N^2]$. Therefore the migration term for a single colony in the two-level mean-field system can be interpreted as a drift towards the 1-block average of the active population at rate c_0 and a drift towards the 2-block average of the active population at rate $\frac{c_1}{N}$. We are interested in (8.7) on time scales N^0, N and N^2. On time scale N^0 we will look at the single colonies, i.e., space-time scale 0. On time scale N we will look at the 1-block averages, i.e., space-time scale 1 and on time scale N^2 we will look at the 2-block averages, i.e., space-time scale 2. In the sequel we will focus on site 0, the 1-block around site 0 and the 2-block around site 0. We will suppress this site from the notation, but instead use subscripts 0, 1, 2 to indicate when we look at a single colony, a 1-block average or a 2-block average. We will use the convention that in the subscript of a dormant population the first subscript denotes the colour and the second subscript denotes the level of the block, so $y_{0,1}$ is the 1-block average around site 0 of the dormant population with colour 0, while $y_{1,0}$ is the 1-dormant single colony at site 0. Heuristically, we can read off the following results from the SSDE in (8.7).

• **On time scale** $1 = N^0$ (i.e., space-time scale 0) in the limit as $N \to \infty$, the colour-1 dormant population and the colour-2 dormant population do not yet move. Hence

$$\left(y_{1,0}^{[N^2]}(t_0), y_{2,0}^{[N^2]}(t_0)\right)_{t_0 \geq 0}, \tag{8.8}$$

converges as $N \to \infty$ to the constant processes on time scale t_0. Therefore the colour 1-dormant population and the colour 2-dormant population are both slow seed-banks on space-time scale 0. The components $((x_0^{[N^2]}(t_0), y_{0,0}^{[N^2]}(t_0)))_{t_0 \geq 0}$ converge to i.i.d. copies of the single-colony McKean-Vlasov process in (6.1), where in the corresponding SSDE the parameters e, K, c are replaced by c_0, e_0, K_0 and $E = 1$. So, on time scale 1 we only see the colour 0-dormant population evolve. Therefore the colour-0 dormant population is the *effective seed-bank* on time scale t_0. The process

$$(z_0^{[N^2]}(t_0))_{t_0 \geq 0} = (x_0^{[N^2]}(t_0), y_{0,0}^{[N^2]}(t_0))_{t_0 \geq 0}, \tag{8.9}$$

will be called *the single colony effective process.*

• **On time scale** N^1 (i.e., space-time scale 1), we look at the averages

$$\begin{aligned}(z_1^{[N^2]}(t_1))_{t_1>0} &= \left(x_1^{[N^2]}(t_1), \left(y_{0,1}^{[N^2]}(t_1), y_{1,1}^{[N^2]}(t_1), y_{2,1}^{[N^2]}(t_1)\right)\right)_{t_1>0}\\ &= \left(\frac{1}{N}\sum_{i\in[N]} x_i^{[N^2]}(Nt_1), \left(\frac{1}{N}\sum_{i\in[N]} y_{i,0}^{[N^2]}(Nt_1), \frac{1}{N}\sum_{i\in[N]} y_{i,1}^{[N^2]}(Nt_1), \frac{1}{N}\sum_{i\in[N]} y_{i,2}^{[N^2]}(Nt_1)\right)\right)_{t_1>0}.\end{aligned} \tag{8.10}$$

(Recall Remark 4.2.4 to appreciate the notation.) We use the lower index 1 to indicate that the average is the analogue of the 1-block average defined in (4.22). Using (8.6), we see that the dynamics of the system in (8.10) is given by the SSDE

$$
\begin{aligned}
\mathrm{d}x_1^{[N^2]}(t_1) &= c_1 \left[\frac{1}{N^2} \sum_{j\in[N^2]} x_j(Nt_1) - x_1(t_1) \right] \mathrm{d}t_1 + \sqrt{\frac{1}{N} \sum_{i\in[N]} g(x_i(Nt_1))}\, \mathrm{d}w(t_1) \\
&\quad + NK_0e_0 \left[y_{0,1}^{[N^2]}(t_1) - x_1^{[N^2]}(t_1) \right] \mathrm{d}t_1 \\
&\quad + K_1e_1 \left[y_{1,1}^{[N^2]}(t_1) - x_1^{[N^2]}(t_1) \right] \mathrm{d}t_1 \\
&\quad + \frac{K_2e_2}{N} \left[y_{2,1}^{[N^2]}(t_1) - x_1^{[N^2]}(t_1) \right] \mathrm{d}t_1, \\
\mathrm{d}y_{0,1}^{[N^2]}(t_1) &= Ne_0 \left[x_1^{[N^2]}(t_1) - y_{0,1}^{[N^2]}(t_1) \right] \mathrm{d}t_1, \\
\mathrm{d}y_{1,1}^{[N^2]}(t_1) &= e_1 \left[x_1^{[N^2]}(t_1) - y_{1,1}^{[N^2]}(t_1) \right] \mathrm{d}t_1, \\
\mathrm{d}y_{2,1}^{[N^2]}(t_1) &= \frac{e_2}{N} \left[x_1^{[N^2]}(t_1) - y_{1,1}^{[N^2]}(t_1) \right] \mathrm{d}t_1.
\end{aligned}
\tag{8.11}
$$

In the limit $N \to \infty$ we expect that the colour 2-dormant population does not move, since it only interacts with the active population at rate $\frac{e_2}{N}$. Therefore we expect $(y_{2,1}^{[N^2]}(t))_{t>0}$ to converge to a constant process and hence we say that the colour 2-dormant population behaves like a *slow seed-bank.* The colour 1-dormant population, however, has a non-trivial interaction with the active population and therefore is the *effective seed-bank* on space-time scale 1. The colour 0-dormant population has, in the limit as $N \to \infty$, an infinitely strong interaction with the active population. Therefore we expect that, in the limit as $N \to \infty$, its path becomes rougher and rougher at rarer and rarer times. We will need to use the *Meyer-Zheng topology* to prove that

$$
\lim_{N\to\infty} y_{0,1}^{[N^2]}(t_1) = \lim_{N\to\infty} x_1^{[N^2]}(t_1) \text{ for most } t_1. \tag{8.12}
$$

Therefore the colour 0-dormant population equalizes with the active population, due to its infinitely strong interaction with the active population. Hence at space-time scale 1, the colour 0-dormant population behaves like a *fast seed-bank.* If we look at the active population, then we see that it feels a drift towards the 2-block average of the active population, and resamples at a rate that is the 1-block average of the resampling rates in the single colonies. Furthermore, in the limit as $N \to \infty$, it feels an infinitely fast drift towards the colour 0-dormant population, has a non-trivial interaction with the colour 1-dormant population, and its interaction with the colour 2-dormant population cancels out. As long as we focus on the combination

$$
\frac{x_1^{[N^2]}(t_1) + K_0 y_{0,1}^{[N^2]}(t_1)}{1+K_0},
$$

we see that the colour-0 terms with the factor N in front cancel out. This will allow us to do most of the analysis in the path space topology, without using the Meyer-Zheng topology. The process

$$
\left(\frac{x_1^{[N^2]}(t_1) + K_0 y_{0,1}^{[N^2]}(t_1)}{1+K_0}, y_{1,1}^{[N^2]}(t_1) \right)_{t_1>0}
$$

will therefore be called the *effective process.*

• **On time scale** N^2 (i.e., space-time scale 2) we look at the equivalent of the 2-block averages in (4.22),

$$\begin{aligned}
&\left(x_2^{[N^2]}(t_2),(y_{0,2}^{[N^2]}(t_2),y_{1,2}^{[N^2]}(t_2),y_{2,2}^{[N^2]}(t_2))\right)_{t_2>0}\\
&=\Bigg(\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(N^2t_2),\\
&\qquad\left(\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,0}^{[N^2]}(N^2t_2),\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,1}^{[N^2]}(N^2t_2),\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,2}^{[N^2]}(N^2t_2)\right)\Bigg)_{t_2>0},
\end{aligned}\tag{8.13}$$

which evolves according to the SSDE

$$\begin{aligned}
\mathrm{d}x_2^{[N^2]}(t_2) &= \sqrt{\frac{1}{N^2}\sum_{i\in[N^2]}g(x_i(N^2t_2))}\,\mathrm{d}w(t_2)\\
&\qquad + N^2K_0e_0\left[y_{0,2}^{[N^2]}(t_2)-x_2^{[N^2]}(t_2)\right]\mathrm{d}t_2\\
&\qquad + NK_1e_1\left[y_{1,2}^{[N^2]}(t_2)-x_2^{[N^2]}(t_2)\right]\mathrm{d}t_2\\
&\qquad + K_2e_2\left[y_{2,2}^{[N^2]}(t_2)-x_2^{[N^2]}(t_2)\right]\mathrm{d}t_2,\\
\mathrm{d}y_{0,2}^{[N^2]}(t_2) &= N^2e_0\left[x_2^{[N^2]}(t_2)-y_{0,2}^{[N^2]}(t_2)\right]\mathrm{d}t_2,\\
\mathrm{d}y_{1,2}^{[N^2]}(t_2) &= Ne_1\left[x_2^{[N^2]}(t_2)-y_{1,2}^{[N^2]}(t_2)\right]\mathrm{d}t_2,\\
\mathrm{d}y_{2,2}^{[N^2]}(t_2) &= e_2\left[x_2^{[N^2]}(t_2)-y_{2,2}^{[N^2]}(t_2)\right]\mathrm{d}t_2.
\end{aligned}\tag{8.14}$$

In this case we see that migration in the active component cancels out and the resampling rate is given by the average over the complete population. In the limit as $N\to\infty$, we see that the active population interacts at an infinitely fast rate with the 0-dormant population as well as with the colour 1-dormant population. Hence both the colour 0 and the colour 1 seed-banks are fast seed-banks and we expect equalisation of the active population and the colour 0-dormant population and the colour 1-dormant population in Meyer-Zheng topology. The active population, in the limit as $N\to\infty$, has a non-trivial interaction with the colour 2-dormant population, and hence the colour 2-dormant population is the *effective seed-bank* on time scale N^2. Looking at the quantity

$$\frac{x_2^{[N^2]}(t_2)+K_0y_{0,2}^{[N^2]}(t_2)+K_1y_{1,2}^{[N^2]}(t_2)}{1+K_0+K_1},\tag{8.15}$$

for which we find

$$\mathrm{d}\left[\frac{x_2^{[N^2]}(t_2)+K_0y_{0,2}^{[N^2]}(t_2)+K_1y_{1,2}^{[N^2]}(t_2)}{1+K_0+K_1}\right]$$
$$=\frac{1}{1+K_0+K_1}\sqrt{\frac{1}{N^2}\sum_{i\in[N^2]}g(x_i(N^2t_2))}\,\mathrm{d}w(t_2)+K_2e_2\left[y_{2,2}^{[N^2]}(t_2)-x_2^{[N^2]}(t_2)\right]\mathrm{d}t_2, \tag{8.16}$$

we see that the infinite rates cancel out. We will call

$$\left(\frac{x_2^{[N^2]}(t_2)+K_0y_{0,2}^{[N^2]}(t_2)+K_1y_{1,2}^{[N^2]}(t_2)}{1+K_0+K_1},y_{2,2}^{[N^2]}(t_2)\right)_{t_2>0} \tag{8.17}$$

the effective process. Using the effective process we can analyse our system in path space.

► **Scaling limit.** Let $(z_0(t))_{t\geq 0}=(x_0(t),(y_{0,0}(t),y_{1,0}(t),y_{2,0}(t)))_{t\geq 0}$ be the process evolving according to

$$\begin{aligned}\mathrm{d}x_0(t)&=c_0\left[\theta-x_0(t)\right]\mathrm{d}t+\sqrt{g(x_0(t))}\,\mathrm{d}w(t)\\&\qquad+K_0e_0\left[y_{0,0}(t)-x_0(t)\right]\mathrm{d}t,\\\mathrm{d}y_{0,0}(t)&=e_0\left[x_0(t)-y_{0,0}(t)\right]\mathrm{d}t,\\y_{1,0}(t)&=y_{1,0},\\y_{2,0}(t)&=y_{2,0},\end{aligned} \tag{8.18}$$

where $\theta\in[0,1]$, $y_{1,0}\in[0,1]$ and $y_{2,0}\in[0,1]$. The process $(z_0(t))_{t\geq 0})$ will be the limiting process for the single colonies. The corresponding single colony effective processes are given by

$$\begin{aligned}\mathrm{d}x_0^{\mathrm{eff}}(t)&=c_0[\theta-x_0^{\mathrm{eff}}(t)]\,\mathrm{d}t+\sqrt{g(x_0^{\mathrm{eff}}(t))}\,\mathrm{d}w(t)+K_0e_0\left[y_{0,0}^{\mathrm{eff}}(t)-x_0^{\mathrm{eff}}(t)\right]\mathrm{d}t,\\\mathrm{d}y_{0,0}^{\mathrm{eff}}(t)&=e_0\left[x_0^{\mathrm{eff}}(t)-y_{0,0}^{\mathrm{eff}}(t)\right]\mathrm{d}t,\qquad i\in\mathbb{N}_0,\end{aligned} \tag{8.19}$$

where $\theta\in[0,1]$. By [72], (8.18) and (8.19) have a unique strong solution. Like for the one-colour mean-field finite-systems scheme, we need the following list of ingredients to formally state the multi-scale analysis:

(a) For positive times $t>0$, we define the following 1-*block estimators* for the finite system:

$$\begin{aligned}
\bar{\Theta}^{(1),[N^2]}(t) &= \frac{1}{N}\sum_{i\in[N]} \frac{x_i^{[N^2]}(t)+K_0 y_{i,0}^{[N^2]}(t)}{1+K_0},\\
\Theta_x^{(1),[N^2]}(t) &= \frac{1}{N}\sum_{i\in[N]} x_i^{[N^2]}(t),\\
\Theta_{y_0}^{(1),[N^2]}(t) &= \frac{1}{N}\sum_{i\in[N]} y_{i,0}^{[N^2]}(t),\\
\Theta_{y_1}^{(1),[N^2]}(t) &= \frac{1}{N}\sum_{i\in[N]} y_{i,1}^{[N^2]}(t),\\
\Theta_{y_2}^{(1),[N^2]}(t) &= \frac{1}{N}\sum_{i\in[N]} y_{i,2}^{[N^2]}(t).
\end{aligned} \tag{8.20}$$

We abbreviate

$$\begin{aligned}
\mathbf{\Theta}^{(1),[N^2]}(t) &= \left(\Theta_x^{(1),[N^2]}(t), \left(\Theta_{y_0}^{(1),[N^2]}(t), \Theta_{y_1}^{(1),[N^2]}(t), \Theta_{y_2}^{(1),[N^2]}(t)\right)\right),\\
\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(t) &= \left(\bar{\Theta}^{(1),[N^2]}(t), \Theta_{y_1}^{(1),[N^2]}(t), \Theta_{y_2}^{(1),[N^2]}(t)\right),\\
\mathbf{\Theta}^{\mathrm{eff},(1),[N^2]}(t) &= \left(\bar{\Theta}^{(1),[N^2]}(t), \Theta_{y_1}^{(1),[N^2]}(t)\right).
\end{aligned} \tag{8.21}$$

We call $(\mathbf{\Theta}^{(1),[N^2]}(t))_{t>0}$ the 1-*block estimator process*, $(\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(t))_{t>0}$ the *auxiliary* 1-*block estimator process* and $(\mathbf{\Theta}^{\mathrm{eff},(1),[N^2]}(t))_{t>0}$ the *effective* 1-*block estimator process*. The auxiliary 1-block estimator will be useful in the proofs. For $t>0$, we define the following 2-*block estimators* for the finite system:

$$\begin{aligned}
\bar{\Theta}^{(2),[N^2]}(t) &= \frac{1}{N^2}\sum_{i\in[N^2]} \frac{x_i^{[N^2]}(t)+K_0 y_{i,0}^{[N^2]}(t)+K_1 y_{i,1}^{[N^2]}(t)}{1+K_0+K_1},\\
\Theta_x^{(2),[N^2]}(t) &= \frac{1}{N^2}\sum_{i\in[N^2]} x_i^{[N^2]}(t),\\
\Theta_{y_0}^{(2),[N^2]}(t) &= \frac{1}{N^2}\sum_{i\in[N^2]} y_{i,0}^{[N^2]}(t),\\
\Theta_{y_1}^{(2),[N^2]}(t) &= \frac{1}{N^2}\sum_{i\in[N^2]} y_{i,1}^{[N^2]}(t),\\
\Theta_{y_2}^{(2),[N^2]}(t) &= \frac{1}{N^2}\sum_{i\in[N^2]} y_{i,2}^{[N^2]}(t).
\end{aligned} \tag{8.22}$$

We abbreviate

$$\begin{aligned}
\mathbf{\Theta}^{(2),[N^2]}(t) &= \left(\Theta_x^{(2),[N^2]}(t), \left(\Theta_{y_0}^{(2),[N^2]}(t), \Theta_{y_1}^{(2),[N^2]}(t), \Theta_{y_2}^{(2),[N^2]}(t)\right)\right),\\
\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(t) &= \left(\bar{\Theta}^{(2),[N^2]}(t), \Theta_{y_1}^{(2),[N^2]}(t)\right).
\end{aligned} \tag{8.23}$$

We call $(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N^2]}(t))_{t>0}$ the *effective* 2-*block estimator process* and $(\boldsymbol{\Theta}^{(2),[N^2]}(t))_{t>0}$ as the 2-*block estimator process.*

(b) The *time scale* N for which $\mathcal{L}[\bar{\Theta}^{(1),[N^2]}(Nt_1 - L(N)) - \bar{\Theta}^{(1),[N^2]}(Nt_1)] = \delta_0$ for all $L(N)$ such that $L(N) \to \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, but not for $L(N) = N$. In words, N is the time scale on which $\bar{\Theta}^{(1),[N^2]}(\cdot)$ starts evolving, i.e., $(\bar{\Theta}^{(1),[N^2]}(Nt_1))_{t_1>0}$ is no longer a fixed process. When we use time scale N, we will use t_1 as a time index, which indicates the "faster time scale". For the "slow time scale" we use t_0 as time index.

The *time scale* N^2 for which $\mathcal{L}[\bar{\Theta}^{(2),[N^2]}(N^2t_2 - L(N)N) - \bar{\Theta}^{(2),[N^2]}(N^2t_2)] = \delta_0$ for all $L(N)$ such that $L(N) \to \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, but not for $L(N) = N$. In words, N^2 is the time scale on which $\bar{\Theta}^{(2),[N^2]}(\cdot)$ starts evolving, i.e., $(\bar{\Theta}^{(2),[N^2]}(N^2t_2))_{t_2>0}$, is no longer a fixed process. When we use time scale N^2, we will use t_2 as a time index, which indicates the "fastest time scale".

(c) The *invariant measure* for the evolution of a single colony in (8.18), written

$$\Gamma^{(0)}_{\theta,\mathbf{y}_0}, \qquad \mathbf{y}_0 = (\theta, y_{1,0}, y_{2,0}), \tag{8.24}$$

and the invariant measure of the level-0 effective process evolving according to (8.19), written

$$\Gamma^{\mathrm{eff},(0)}_{\theta}. \tag{8.25}$$

(d) The renormalisation transformation $\mathcal{F}\colon \mathcal{G} \to \mathcal{G}$,

$$(\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x)\, \Gamma^{\mathrm{eff},(0)}_{\theta}(\mathrm{d}x, \mathrm{d}y_0), \quad \theta \in [0,1], \tag{8.26}$$

where $\Gamma^{\mathrm{eff},(0)}_{\theta}$ is the equilibrium measure in (8.25). Note that this is the same transformation as defined in (4.75), but defined for the truncated system. Later we will study iterates of the renormalisation transformation. Therefore we will write $\mathcal{F}^{(1)}g = \mathcal{F}g$, to indicate that we apply the renormalisation transformation only once.

(e) The limiting 1-block process is given by

$$(z_1(t))_{t>0}) = (x_1(t), (y_{0,1}(t), y_{1,1}(t), y_{2,1}(t)))_{t>0} \tag{8.27}$$

and evolves according to

$$\begin{aligned} \mathrm{d}x_1(t) &= \frac{1}{1+K_0}\Bigg[c_1[\theta - x_1(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(1)}g)(x_1(t))}\,\mathrm{d}w(t) \\ &\qquad + K_1 e_1\,[y_{1,1}(t) - x_1(t)]\,\mathrm{d}t\Bigg], \\ y_{0,1}(t) &= x_1(t), \\ \mathrm{d}y_{1,1}(t) &= e_1\,[x_1(t) - y_{1,1}(t)]\,\mathrm{d}t, \\ y_{2,1}(t) &= y_{2,1}, \end{aligned} \tag{8.28}$$

where $\theta \in [0,1]$, and $y_{2,1} \in [0,1]$, and $\mathcal{F}^{(1)}$ is the renormalisation transformation defined in (8.26). The limiting 1-block process for the auxiliary estimator process is given by $(z_1^{\mathrm{aux}}(t))_{t>0} = (x_1^{\mathrm{aux}}(t), y_{1,1}^{\mathrm{aux}}(t), y_{2,1}^{\mathrm{aux}}(t))_{t>0}$ and evolves according to

$$\begin{aligned}
\mathrm{d}x_1^{\mathrm{aux}}(t) &= \frac{1}{1+K_0}\Bigg[c_1[\theta - x_1^{\mathrm{aux}}(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(1)}g)(x_1^{\mathrm{aux}}(t))}\,\mathrm{d}w(t) \\
&\qquad + K_1 e_1\,[y_{1,1}^{\mathrm{aux}}(t) - x_1^{\mathrm{aux}}(t)]\,\mathrm{d}t\Bigg], \\
\mathrm{d}y_{1,1}^{\mathrm{aux}}(t) &= e_1\,[x_1^{\mathrm{aux}}(t) - y_{1,1}^{\mathrm{aux}}(t)]\,\mathrm{d}t, \\
y_{2,1}^{\mathrm{aux}}(t) &= y_{2,1},
\end{aligned} \tag{8.29}$$

for $\theta \in [0,1]$. The auxiliary estimator process turns out to be important in the next section. The effective limiting 1-block process is given by $(z_1^{\mathrm{eff}}(t))_{t>0}) = (x_1^{\mathrm{eff}}(t), y_{1,1}^{\mathrm{eff}}(t))_{t>0}$ and evolves according to

$$\begin{aligned}
\mathrm{d}x_1^{\mathrm{eff}}(t) &= \frac{1}{1+K_0}\Bigg[c_1[\theta - x_1^{\mathrm{eff}}(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(1)}g)(x_1^{\mathrm{eff}}(t))}\,\mathrm{d}w(t) \\
&\qquad + K_1 e_1\,[y_{1,1}^{\mathrm{eff}}(t) - x_1^{\mathrm{eff}}(t)]\,\mathrm{d}t\Bigg], \\
\mathrm{d}y_{1,1}^{\mathrm{eff}}(t) &= e_1\,[x_1^{\mathrm{eff}}(t) - y_{1,1}^{\mathrm{eff}}(t)]\,\mathrm{d}t,
\end{aligned} \tag{8.30}$$

for $\theta \in [0,1]$. By [72], (8.28), (8.29) and (8.30) have a unique strong solution.

(f) The *invariant measure* of the infinite system in (8.28), written

$$\Gamma^{(1)}_{\theta,y_1}, \qquad y_1 = (\theta, \theta, y_{2,1}), \tag{8.31}$$

and the invariant measures of the level-1 limiting estimator process evolving according to (8.29) and the level-1 effective process evolving according to (8.30),

$$\Gamma^{\mathrm{aux},(1)}_{\theta}, \Gamma^{\mathrm{eff},(1)}_{\theta}. \tag{8.32}$$

(g) The first iteration of the renormalisation transformation,

$$(\mathcal{F}^{(2)}g)(\theta) = \int_{[0,1]^2} (\mathcal{F}g)(x)\,\Gamma^{\mathrm{eff},(1)}_{\theta}(\mathrm{d}x, \mathrm{d}y_1), \quad \theta \in [0,1]. \tag{8.33}$$

Hence

$$(\mathcal{F}^{(2)}g)(\theta) = \int_{[0,1]^2} \Gamma^{\mathrm{eff},(1)}_{\theta}(\mathrm{d}u, \mathrm{d}v) \int_{[0,1]^2} g(x)\,\Gamma^{\mathrm{eff},(0)}_{u}(\mathrm{d}x, \mathrm{d}y). \tag{8.34}$$

(h) The limiting 2-block process $(z_2(t))_{t>0} = (x_2(t), (y_{0,2}(t), y_{1,2}(t), y_{2,2}(t)))_{t>0}$

CHAPTER 8

evolves according to

$$
\begin{aligned}
\mathrm{d}x_2(t) &= \frac{1}{1+K_0+K_1}\left[\sqrt{(\mathcal{F}^{(2)}g)\big(x_2(t)\big)}\,\mathrm{d}w(t) + K_2e_2\left[y_{2,2}(t)-x_2(t)\right]\mathrm{d}t\right],\\
y_{0,2}(t) &= x_2(t),\\
\mathrm{d}y_{1,2}(t) &= x_2(t),\\
\mathrm{d}y_{2,2}(t) &= e_2\left[x_2(t)-y_{2,2}(t)\right]\,\mathrm{d}t,
\end{aligned}
\tag{8.35}
$$

where $\mathcal{F}^{(2)}g$ is defined as in (8.33). The limiting effective 2-block process on space-time scale 2 is $(z_2^{\text{eff}}(t))_{t>0}) = (x_2^{\text{eff}}(t), y_{2,2}^{\text{eff}}(t))_{t>0}$ and evolves according to

$$
\begin{aligned}
\mathrm{d}x_2^{\text{eff}}(t) &= \frac{1}{1+K_0+K_1}\left[\sqrt{(\mathcal{F}^{(2)}g)(x_2^{\text{eff}}(t))}\,\mathrm{d}w(t) + K_2e_2\left[y_{2,2}^{\text{eff}}(t)-x_2^{\text{eff}}(t)\right]\mathrm{d}t\right],\\
\mathrm{d}y_{2,2}^{\text{eff}}(t) &= e_2\left[x_2^{\text{eff}}(t)-y_{2,2}^{\text{eff}}(t)\right]\mathrm{d}t.
\end{aligned}
\tag{8.36}
$$

We are now ready to state the scaling limit for the evolution of the averages in (7.7).

Proposition 8.1.1 (Two-level three-colour finite-systems scheme). *Suppose that $\mu(0) = \mu^{\otimes[N^2]}$ for some $\mu \in \mathcal{P}\left([0,1]\times[0,1]^2\right)$. Let*

$$
\begin{aligned}
&\vartheta_0 = \mathbb{E}^{\mu}\left[\frac{x+K_0y_0}{1+K_0}\right], \quad \vartheta_1 = \mathbb{E}^{\mu}\left[\frac{x+K_0y_0+K_1y_1}{1+K_0+K_1}\right],\\
&\theta_{y_1} = \mathbb{E}^{\mu}\left[y_1\right], \quad \theta_{y_2} = \mathbb{E}^{\mu}\left[y_2\right].
\end{aligned}
\tag{8.37}
$$

and recall the limiting process $(z_2(t))_{t>0}$ in (8.35) *and the limiting process $(z_1(t))_{t>0}$ in* (8.28)*. Assume for the* 2*-dormant* 1*-blocks that*

$$
\lim_{N\to\infty}\mathcal{L}\left[Y_{2,1}^{[N^2]}(Nt_2)\middle|\boldsymbol{\Theta}^{(2),[N^2]}(N^2t_2)\right] = P^{z_2(t_2)}, \tag{8.38}
$$

and for the 2*-dormant* 0*-blocks (= single colonies) that*

$$
\lim_{N\to\infty}\mathcal{L}\left[Y_{2,0}^{[N^2]}(Nt_2+Nt_1)\middle|\boldsymbol{\Theta}^{\text{eff},(1),[N^2]}(N^2t_2+Nt_1)\right] = P^{z_1(t_1)}. \tag{8.39}
$$

Then the following hold:

(a) For the effective 2*-block estimator process defined in* (8.23),

$$
\lim_{N\to\infty}\mathcal{L}\left[\left(\boldsymbol{\Theta}^{\text{eff},(2),[N^2]}(N^2t_2)\right)_{t_2>0}\right] = \mathcal{L}\left[\left(z_2^{\text{eff}}(t_2)\right)_{t_2>0}\right], \tag{8.40}
$$

where the limit is determined by the unique solution of the SSDE (8.36) *with initial state*

$$
z_2^{\text{eff}}(0) = \left(x_2^{\text{eff}}(0), y_2^{\text{eff}}(0)\right) = (\vartheta_1, \theta_{y_2}). \tag{8.41}
$$

(b) *For the effective* 1*-block estimator process defined in* (8.21),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{eff},(1),[N^2]}(N^2t_2+Nt_1)\right)_{t_1>0}\right] = \mathcal{L}\left[(z_1^{\mathrm{eff}}(t_1))_{t_1>0}\right], \tag{8.42}$$

where, conditional on $x_2^{\mathrm{eff}}(t_2)=u$, *the limit process is the unique solution of the SSDE in* (8.30) *with* θ *replaced by* u *and with initial measure* $\Gamma_u^{\mathrm{eff},(1)}$.

(c) *For the single colony effective process defined in* (8.9),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(z_0^{\mathrm{eff},[N^2]}(N^2t_2+Nt_1+t_0)\right)_{t_0\geq 0}\right] = \mathcal{L}\left[(z_0^{\mathrm{eff}}(t_0))_{t_0\geq 0}\right], \tag{8.43}$$

where, conditional on $x_1^{\mathrm{eff}}(t_1)=v$, *the limit process is the unique solution of the SSDE in* (8.19) *with* θ *replaced by* v *and with initial measure* $\Gamma_v^{\mathrm{eff},(0)}$.

(d) *For the* 2*-block estimator process defined in* (8.23),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(2),[N^2]}(N^2t_2)\right)_{t_2>0}\right] = \mathcal{L}\left[(z_2(t_2))_{t_2>0}\right] \tag{8.44}$$

in the Meyer-Zheng topology,

where the limit process is the unique solution of the SSDE in (8.35) *with initial state*

$$z_2(0) = (\vartheta_1,(\vartheta_1,\vartheta_1,\theta_{y_2})). \tag{8.45}$$

(e) *Fix* $t_2>0$. *Assume* (8.38). *Define*

$$\begin{aligned}\Gamma^{(1)}(t_2) = &\int_{[0,1]^4} S_{t_2}^{[2]}\big((\vartheta_1,(\vartheta_1,\vartheta_1,\theta_{y_2})),\mathrm{d}(u_x,u_x,u_x,u_{y_{2,2}})\big)\\ &\int_{[0,1]} P^{(u_x,u_x,u_x,u_{y_{2,2}})}(\mathrm{d}y_{2,1})\,\Gamma^{(1)}_{(u_x,(u_x,u_x,y_{2,1}))} \in \mathcal{P}([0,1]^4),\end{aligned} \tag{8.46}$$

where $\Gamma^{(1)}_{(u_x,(u_x,u_x,y_{2,1}))}$ *is the equilibrium measure in* (8.31) *and* $S_{t_2}^{[2]}((\vartheta_1,(\vartheta_1,\vartheta_1,\theta_{y_2})),\cdot)$ *is the time-*t_2 *law of the limiting process* $(z_2(t_2))_{t_2>0}$ *in* (8.44) *starting from* $(\vartheta_1,(\vartheta_1,\vartheta_1,\theta_{y_2}))\in[0,1]\times[0,1]^3$.

Let $(z^{\Gamma^{(1)}(t_2)}(t_1))_{t_1\geq 0}$ *be the random process that conditioned on* $z_2(t_2)=(\theta,(\theta,\theta,y_{2,2}))$ *moves according to* (8.28) *with* $\theta=\theta$ *and* $y_{2,1}(0)=y_{2,1}$ *and with* $z^{\Gamma^{(1)}(t_2)}(0)$ *be drawn according to* $\Gamma^{(1)}(t_2)$ *(which is a mixture of random processes in equilibrium). Then for the* 1*-block estimator process defined in* (8.21),

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(1),[N^2]}(N^2t_2+Nt_1)\right)_{t_1>0}\right] = \mathcal{L}\left[(z^{\Gamma^{(1)}(t_2)}(t_1))_{t_1>0}\right] \tag{8.47}$$

in the Meyer-Zheng topology.

(f) *Let $z_1(t_1)$ be the limiting process obtained in (e). Assume* (8.39)*. Define, for $t_2 \in (0,\infty)$,*

$$\Gamma^{(0)}(t_2) = \int_{[0,1]^4} \Gamma^{(1)}(t_2)(\mathrm{d}z_1) \int_{[0,1]} P^{z_1}(\mathrm{d}y_{2,0})\, \Gamma^{(0)}_{(x_1,(x_1,y_{1,1}.y_{2,0}))}, \tag{8.48}$$

where $\Gamma^{(1)}(t_2)$ is as defined in (8.46)*. Let $(z^{\Gamma^{(0)}(t_2)}(t_0))_{t_0\geq 0}$ be the random process in* (8.18) *with $z^{\Gamma^{(0)}(t_2)}(0)$ drawn according to $\Gamma^{(0)}(t_2)$ which is a mixture of random processes in equilibrium. Then*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(z_0^{[N^2]}(N^2t_2+Nt_1+t_0)\right)_{t_0\geq 0}\right] = \mathcal{L}\left[(z^{\Gamma^{(0)}(t_2)}(t_0))_{t_0\geq 0}\right]. \tag{8.49}$$

Remark 8.1.2. Note that Proposition 8.1.1(f) does not depend on the choice of t_1, because $\Gamma^{(1)}(t_2)$ is already a mixture of equilibrium measures of the 1-block process. ■

Remark 8.1.3. Note that in Propostion 8.1.1(f) $\Gamma^{(0)}_{(x_1,(x_1,y_{1,1},y_{2,0}))}$ is the equilibrium measure of (8.18) (see also (8.24)), where $y_{1,0} = y_{1,1}$. This means that all colour 1-dormant single colonies equal the current state of the colour 1-dormant 1-block. We say that given the state of the 1-dormant 1-block, the 1-dormant single colonies become deterministic. This effect occurs once a slow seed-bank, in this case the colour 1 seed-bank, is already in equilbrium on the space-time scale where it is effective, in this case space-time-scale 1. Since we start at times N^2t_2, the 1-dormant 1-blocks are already in equilibrium. This will turn out to be the reason that the single colour 1-dormant colonies are equal to the current value of the 1-dormant 1-block averages. Note that at time N^2t_2 the 2-dormant 2-blocks do not yet have reached equilibrium. Hence the colour 2-dormant 1-blocks and the colour 2-dormant single colonies do not equal the instantaneous value of the 2-dormant 2-block averages. In the Section 8.3.8 we will treat this effect in detail. ■

§8.2 Scheme for the two-level three-colour mean-field analysis.

In this section we give a scheme to prove Proposition 8.1.1. The proof of the steps in the scheme will be written in Section 8.3. To analyse the two-level hierarchical mean-field system we use the results obtained in Sections 6.2.2, 6.3 and 7.2.

The scheme for the two-level three-colour hierarchical mean-field system comes in 11 steps. Recall the estimators defined in (8.20) and (8.22).

1 Tightness of the effective 2-block estimator processes

$$\left(\left(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2)\right)_{t_2>0}\right)_{N\in\mathbb{N}}. \tag{8.50}$$

2 Stability property of the 2-block estimators, i.e., for $L(N)$ such that $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N=0$,

$$\lim_{N\to\infty} \sup_{0\leq t\leq L(N)} \left|\bar{\Theta}^{(2),[N^2]}(N^2t_2) - \bar{\Theta}^{(2),[N^2]}(N^2t_2-Nt)\right| = 0 \text{ in probability} \tag{8.51}$$

and

$$\lim_{N\to\infty} \sup_{0\le t\le L(N)} \left|\Theta_{y_2}^{(2),[N^2]}(N^2t_2) - \Theta_{y_2}^{(2),[N^2]}(N^2t_2 - Nt)\right| = 0 \text{ in probability.} \tag{8.52}$$

3 Tightness of the effective 1-block estimator process (recall (8.21)),

$$\left(\left(\boldsymbol{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2 + Nt_1)\right)_{t_1>0}\right)_{N\in\mathbb{N}}. \tag{8.53}$$

4 Stability property of $(\boldsymbol{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2 + Nt_1))_{t_1>0}$, i.e., for $L(N)$ such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, for all $\epsilon > 0$,

$$\lim_{N\to\infty} \sup_{0\le t\le L(N)} \left|\bar{\Theta}^{(1),[N^2]}(N^2t_2 + Nt_1) - \bar{\Theta}^{(1),[N^2]}(N^2t_2 + Nt_1 - t)\right| = 0$$
$$\text{in probability,} \tag{8.54}$$

$$\lim_{N\to\infty} \sup_{0\le t\le L(N)} \left|\Theta_{y_1}^{(1),[N^2]}(N^2t_2 + Nt_1) - \Theta_{y_1}^{(1),[N]}(N^2t_2 + Nt_1 - t)\right| = 0 \tag{8.55}$$
$$\text{in probability,}$$

$$\lim_{N\to\infty} \sup_{0\le t\le L(N)} \left|\Theta_{y_2}^{(1),[N^2]}(N^2t_2 + Nt_1) - \Theta_{y_2}^{(1),[N]}(N^2t_2 + Nt_1 - t)\right| = 0 \tag{8.56}$$
$$\text{in probability.}$$

5 Recall that there are N 1-blocks in $[N^2]$. Since tightness of components implies tightness of the process, step 3 implies that the full 1-block process

$$\left(\left(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N^2]}(N^2t_2 + Nt_1)\right)_{t_1>0,\, i\in[N]}\right)_{N\in\mathbb{N}} \tag{8.57}$$

is tight. From the tightness in steps 1 and 3 we can construct a subsequence $(N_k)_{k\in\mathbb{N}}$ along which

$$\begin{aligned} &\lim_{k\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2t_2)\right)_{t_2>0}\right], \\ &\lim_{k\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N_k^2]}(N_k^2t_2 + N_kt_1)\right)_{t_1>0,\, i\in[N_k]}\right] \end{aligned} \tag{8.58}$$

both exists. Define the measure

$$\nu^{(0)}(t_2) = \prod_{i\in\mathbb{N}_0} \Gamma_i^{(0)}(t_2). \tag{8.59}$$

Show that along the same subsequence the single components converge to the infinite system, i.e.,

$$\lim_{k\to\infty} \mathcal{L}\Big[\Big(Z^{[N_k^2]}(N_k^2t_2 + N_kt_1 + t_0)\Big)_{t_0\ge0}\Big] = \mathcal{L}\left[(Z^{\nu^{(0)}(t_2)}(t_0))_{t_0\ge0}\right]. \tag{8.60}$$

Here, $(Z^{\nu^{(0)}(t_2)}(t_0))_{t_0 \geq 0}$ is the process starting from $\nu^{(0)}(t_2)$ with components evolving according to (8.18), where θ is now a random variable that inherits its law from

$$\lim_{k\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{i\in[N_k^2]}], \tag{8.61}$$

and, similarly, the laws of $y_{1,0}$ and $y_{2,0}$ in the limiting process $(Z^{\nu^{(0)}(t_2)}(t_0))_{t_0 \geq 0}$ are determined by

$$\lim_{k\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{i\in[N_k^2]}]. \tag{8.62}$$

6 Use the limiting evolution of the single colonies obtained in step 5 to identify the limiting 1-block process along the same subsequence, i.e., identify the limit

$$\lim_{k\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right)_{t_1>0, i\in[N_k]}\right]. \tag{8.63}$$

7 Identify the limit $\lim_{k\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2 t_2))_{t_2>0}]$ with the help of the limiting evolution of the single colonies obtained in step 5 and the limiting evolution of the full 1-block process obtained in step 6.

8 Prove that the 1-dormant single colonies at time $N^2 t_2 + N t_1$ equal, in the limit as $N \to \infty$, the 1-dormant 1-block averages. The proof of this step shows how the evolution of the slow seed-banks must be analysed.

9 Show that the convergence in step 8, step 7 and step 5 actually holds along each subsequence. Therefore we obtain the limiting evolution of the single colonies, the auxiliary 1-block process and the effective 2-block process.

10 Use the Meyer-Zheng topology to describe the limiting evolution of

$$\begin{aligned}\Big(&\Theta_x^{(1),[N^2]}(N^2 t_2 + N t_1), \Theta_{y_0}^{(1),[N^2]}(N^2 t_2 + N t_1), \Theta_{y_1}^{(1),[N^2]}(N^2 t_2 + N t_1),\\ &\Theta_{y_2}^{(1),[N^2]}(N^2 t_2 + N t_1)\Big)_{t_1>0}\end{aligned} \tag{8.64}$$

and

$$\left(\Theta_x^{(2),[N^2]}(N^2 t_2), \Theta_{y_0}^{(2),[N^2]}(N^2 t_2), \Theta_{y_1}^{(2),[N^2]}(N^2 t_2), \Theta_{y_2}^{(2),[N^2]}(N^2 t_2)\right)_{t_2>0}. \tag{8.65}$$

11 Combine the above steps to complete the proof of Proposition 8.1.1.

§8.3 Proof of two-level three-colour mean-field finite-systems scheme

In this section we prove the steps in the scheme given in Section 8.2.

§8.3.1 Tightness of the 2-block estimators

In this section we prove step 1 of the scheme.

Lemma 8.3.1 (Tightness of the 2-block estimator). *Let*

$$\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2) = (\bar{\Theta}^{(2),[N^2]}(N^2t_2), \Theta_{y_2}^{(2),[N^2]}(N^2t_2)) \tag{8.66}$$

be defined as in (8.22)*. Then* $(\mathcal{L}[(\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2))_{t_2>0}])_{N\in\mathbb{N}}$ *is a tight sequence of probability measures on* $\mathcal{C}((0,\infty),[0,1]^2)$*.*

Proof. To prove the tightness of the 1-blocks, we use [49, Proposition 3.2.3]. From (8.6) we find that $(\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(t))_{t>0}$ evolves according to

$$\begin{aligned}
\mathrm{d}\bar{\Theta}^{(2),[N^2]}(t) &= \frac{1}{1+K_0+K_1}\frac{1}{N^2}\sum_{i\in[N^2]}\sqrt{g(x_i^{[N^2]}(t))}\,\mathrm{d}w_i(t)\\
&\quad+\frac{1}{1+K_0+K_1}\frac{K_2e_2}{N^2}\left[\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,2}^{[N^2]}(t)-\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(t)\right]\mathrm{d}t,\\
\mathrm{d}\Theta_{y_2}^{(2),[N^2]}(t) &= \frac{e_2}{N^2}\left[\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(t)-\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,2}^{[N^2]}(t)\right]\mathrm{d}t.
\end{aligned} \tag{8.67}$$

Therefore the process $(\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2))_{t_2>0}$ evolves according to

$$\begin{aligned}
\mathrm{d}\bar{\Theta}^{(2),[N^2]}(N^2t_2) &= \frac{1}{1+K_0+K_1}\sqrt{\frac{1}{N^2}\sum_{i\in[N^2]}g(x_i^{[N^2]}(N^2t_2))\,\mathrm{d}w_i(t_2)}\\
&\quad+\frac{1}{1+K_0+K_1}K_2e_2\left[\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,2}^{[N^2]}(N^2t_2)-\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(N^2t_2)\right]\mathrm{d}t_2,\\
\mathrm{d}\Theta_{y_2}^{(2),[N^2]}(N^2t_2) &= e_2\left[\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(N^2t_2)-\frac{1}{N^2}\sum_{i\in[N^2]}y_{i,2}^{[N^2]}(N^2t_2)\right]\mathrm{d}t_2.
\end{aligned} \tag{8.68}$$

To use [49, Proposition 3.2.3], we define $\mathcal{C}^*$ as the set of polynomials on $([0,1]^2)$. Since $(\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2))_{t_2>0}$ is a semi-martingale, by applying Itô's formula we obtain that $(\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2))_{t_2>0}$ is a $\mathcal{D}$-semi-martingale with corresponding operator

$$
\begin{aligned}
G_{\dagger}^{(2),[N^2]}&\colon\ (\mathcal{C}^*,[0,1]^2,(0,\infty),\Omega)\to\mathbb{R},\\
G_{\dagger}^{(2),[N^2]}(f,(x,y),t,\omega) &= \frac{K_2e_2}{1+K_0+K_1}\left[y-\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(N^2t,\omega)\right]\frac{\partial f}{\partial x}\\
&\quad+e_2\left[\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(N^2t,\omega)-y\right]\frac{\partial f}{\partial y}\\
&\quad+\frac{1}{2(1+K_0+K_1)^2}\frac{1}{N^2}\sum_{i\in[N^2]}g(x_i^{[N^2]}(N^2t,\omega))\frac{\partial^2 f}{\partial x^2}.
\end{aligned}
\tag{8.69}
$$

The conditions $H_1,\ H_2,\ H_3$ in [49, Proposition 3.2.3] are satisfied. Hence tightness follows from [49, Proposition 3.2.3]. □

§8.3.2 Stability of the 2-block estimators

Lemma 8.3.2 (Stability property of the 2-block estimator). *Let* $(\mathbf{\Theta}^{\mathrm{eff},(2),[N^2]}(t))_{t>0}$ *be defined as in* (8.23). *For any* $L(N)$ *such that* $\lim_{N\to\infty}L(N)=\infty$ *and* $\lim_{N\to\infty}L(N)/N=0$,

$$
\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\bar{\Theta}^{(2),[N^2]}(N^2t_2)-\bar{\Theta}^{(2),[N^2]}(N^2t_2-Nt)\right|=0\ \textit{in probability} \tag{8.70}
$$

and

$$
\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\Theta_{y_2}^{(2),[N^2]}(N^2t_2)-\Theta_{y_2}^{(2),[N^2]}(N^2t_2-Nt)\right|=0\ \textit{in probability.} \tag{8.71}
$$

Proof. Fix $\epsilon > 0$. From the SSDE in (8.67) we obtain that, for N large enough,

$$\begin{aligned}
&\mathbb{P}\left[\sup_{0\le t\le L(N)}\left|\bar\Theta^{(2),[N^2]}(N^2t_2)-\bar\Theta^{(2),[N^2]}(N^2t_2-Nt)\right|>\epsilon\right]\\
&=\mathbb{P}\Bigg[\sup_{0\le t\le L(N)}\frac{1}{1+K_0+K_1}\Bigg|\int_{N^2t_2-Nt}^{N^2t_2}\mathrm{d}w_i(r)\,\frac{1}{N^2}\sum_{i\in[N^2]}\sqrt{g(x_i^{[N^2]}(r))}\\
&\qquad+\int_{N^2t_2-Nt}^{N^2t_2}\mathrm{d}r\,\frac{K_2e_2}{N^2}\left[\Theta_{y_2}^{(2),[N^2]}(r)-\frac{1}{N^2}\sum_{i\in[N^2]}x_i^{[N^2]}(r)\right]\Bigg|>\epsilon\Bigg]\\
&\le\mathbb{P}\Bigg[\sup_{0\le t\le L(N)}\frac{1}{1+K_0+K_1}\left|\int_{N^2t_2-Nt}^{N^2t_2}\mathrm{d}w_i(r)\,\frac{1}{N^2}\sum_{i\in[N^2]}\sqrt{g(x_i^{[N^2]}(r))}\right|\\
&\qquad>\epsilon-\frac{K_2e_2}{1+K_0+K_1}\frac{L(N)N}{N^2}\Bigg]\\
&\le\mathbb{P}\left[\sup_{0\le t\le L(N)}\frac{1}{1+K_0+K_1}\left|\int_{N^2t_2-Nt}^{N^2t_2}\mathrm{d}w_i(r)\frac{1}{N^2}\sum_{i\in[N^2]}\sqrt{g(x_i^{[N^2]}(r))}\right|>\frac{\epsilon}{2}\right].
\end{aligned}\tag{8.72}$$

By a similar optional stopping time argument as in the proof of Lemma 6.2.15, the above computation shows that (8.70) holds. Equation (8.71) holds by a similar argument as given in the proof of Lemma 7.2.2. □

§8.3.3 Tightness of the 1-block estimators

Lemma 8.3.3 (Tightness of the 1-block estimator). *Let*

$$\begin{aligned}
&\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2+Nt_1)\\
&=(\bar\Theta^{(1),[N^2]}(N^2t_2+Nt_1),\Theta_{y_1}^{(1),[N^2]}(N^2t_2+Nt_1),\Theta_{y_2}^{(1),[N^2]}(N^2t_2+Nt_1))
\end{aligned}\tag{8.73}$$

be defined as in (8.20). *Then* $(\mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),[N]}(N^2t_2+Nt_1))_{t_1>0}])_{N\in\mathbb{N}}$ *is a tight sequence of probability measures on* $\mathcal{C}((0,\infty),[0,1]^3)$.

Proof. To prove the tightness of the 1-blocks, we again use [49, Proposition 3.2.3]. From (8.11) we find that the effective process $(\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2+Nt_1))_{t_1>0}$ evolves according to

$$\begin{aligned}
\mathrm{d}\bar{\Theta}^{(1),[N^2]}(Nt_1) &= \frac{1}{1+K_0} c_1 \left[\frac{1}{N^2} \sum_{j\in[N^2]} x_j^{[N^2]}(Nt_1) - \frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt_1) \right] \mathrm{d}t_1 \\
&\quad + \frac{1}{1+K_0} \sqrt{\frac{1}{N} \sum_{i\in[N]} g(x_i^{[N^2]}(Nt_1))}\, \mathrm{d}w_i(t_1) \\
&\quad + \frac{K_1 e_1}{1+K_0} \left[\Theta_{y_1}^{(1),[N^2]}(Nt_1) - \frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt_1) \right] \mathrm{d}t_1 \\
&\quad + \frac{K_2 e_2}{N(1+K_0)} \left[\Theta_{y_2}^{(1),[N^2]}(Nt_1) - \frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt_1) \right] \mathrm{d}t_1, \\
\mathrm{d}\Theta_{y_1}^{(1),[N^2]}(Nt_1) &= e_1 \left[\frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt_1) - \Theta_{y_1}^{(1),[N^2]}(Nt_1) \right] \mathrm{d}t_1, \\
\mathrm{d}\Theta_{y_2}^{(1),[N^2]}(Nt_1) &= \frac{e_2}{N} \left[\frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt_1) - \Theta_{y_2}^{(1),[N^2]}(Nt_1) \right] \mathrm{d}t_1.
\end{aligned} \tag{8.74}$$

To use [49, Proposition 3.2.3], we define $\mathcal{C}^*$ as the set of polynomials on $([0,1]^2)$. Since $(\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2 + Nt_1))_{t_1>0}$ is a semi-martingale, by applying Itô's formula we obtain that $(\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2+Nt_1))_{t_1>0}$ is a $\mathcal{D}$-semi-martingale with corresponding operator

$$\begin{aligned}
G_\dagger^{(1),[N^2]}&\colon (\mathcal{C}^*, [0,1]^3, (0,\infty), \Omega) \to \mathbb{R}, \\
G_\dagger^{(1),[N^2]}(f,(x,y_1,y_2),t,\omega) &= \frac{c_1}{1+K_0} \left[\frac{1}{N^2} \sum_{j\in[N^2]} x_j^{[N^2]}(Nt,\omega) - \frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt,\omega) \right] \frac{\partial f}{\partial x} \\
&\quad + \frac{K_1 e_1}{1+K_0} \left[y_1 - \frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt,\omega) \right] \frac{\partial f}{\partial x} \\
&\quad + \frac{K_2 e_2}{N(1+K_0)} \left[y_2(Nt,\omega) - \frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt) \right] \frac{\partial f}{\partial x} \\
&\quad + e_1 \left[\frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt,\omega) - y_1 \right] \frac{\partial f}{\partial y_1} \\
&\quad + \frac{e_2}{N} \left[\frac{1}{N} \sum_{i\in[N]} x_i^{[N^2]}(Nt,\omega) - y_2(Nt,\omega) \right] \frac{\partial f}{\partial y_2} \\
&\quad + \frac{1}{2(1+K_0)^2} \frac{1}{N} \sum_{i\in[N]} g(x_i^{[N^2]}(Nt,\omega)) \frac{\partial^2 f}{\partial x^2}.
\end{aligned} \tag{8.75}$$

The conditions H_1, H_2, H_3 in [49, Proposition 3.2.3] are satisfied as before. Hence we conclude that the sequence $(\mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2+Nt_1))_{t_1>0}])_{N\in\mathbb{N}}$ is tight. $\square$

§8.3.4 Stability of the 1-block estimators

Lemma 8.3.4 (Stability property of the 1-block estimator). *Let $\mathbf{\Theta}^{\mathrm{aux},(1),[N^2]}(t)$ be defined as in (7.14). For any $L(N)$ such that $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N=0$,*

$$\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\bar{\Theta}^{(1),[N^2]}(N^2t_2+Nt_1)-\bar{\Theta}^{(1),[N^2]}(N^2t_2+Nt_1-t)\right|=0 \text{ in probability,} \tag{8.76}$$

$$\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\Theta^{(1),[N^2]}_{y_1}(N^2t_2+Nt_1)-\Theta^{(1),[N^2]}_{y_1}(N^2t_2+Nt_1-t)\right|=0 \text{ in probability,} \tag{8.77}$$

$$\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\Theta^{(1),[N^2]}_{y_2}(N^2t_2+Nt_1)-\Theta^{(1),[N^2]}_{y_2}(N^2t_2+Nt_1-t)\right|=0 \text{ in probability.} \tag{8.78}$$

Proof. Define
$$u=N^2t_2+Nt_1. \tag{8.79}$$
From the SSDE in (8.7) we obtain that

$$\begin{aligned}
&\mathbb{P}\left[\sup_{0\le t\le L(N)}\left|\bar{\Theta}^{(1),[N^2]}(u)-\bar{\Theta}^{(1),[N^2]}(u-t)\right|>\epsilon\right]\\
&=\mathbb{P}\Bigg[\sup_{0\le t\le L(N)}\frac{1}{1+K_0}\Bigg|\int_{u-t}^{u}\mathrm{d}r\,\frac{c_1}{N}\left[\frac{1}{N^2}\sum_{j\in[N^2]}x_j^{[N^2]}(r)-\frac{1}{N}\sum_{i\in[N]}x_i^{[N^2]}(r)\right]\\
&\qquad+\int_{u-t}^{u}\mathrm{d}r\,\frac{K_1e_1}{N}\left[\Theta^{(1),[N^2]}_{y_1}(r)-\frac{1}{N}\sum_{i\in[N]}x_i^{[N^2]}(r)\right]\\
&\qquad+\frac{K_2e_2}{N^2}\left[\frac{1}{N}\sum_{i\in[N]}y_{i,2}^{[N^2]}(r)-\frac{1}{N}\sum_{i\in[N]}x_i^{[N^2]}(r)\right]\\
&\qquad+\int_{u-t}^{u}\mathrm{d}w_i(r)\,\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N^2]}(r))}\Bigg|>\epsilon\Bigg]\\
&\le\mathbb{P}\Bigg[\sup_{0\le t\le L(N)}\frac{1}{1+K_0}\Bigg|\int_{u-t}^{u}\mathrm{d}w_i(r)\,\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N^2]}(r))}\Bigg|\\
&\qquad>\epsilon-\frac{L(N)2(c_1+K_1e_1+\frac{K_2e_2}{N})}{N(1+K_0)}\Bigg]\\
&\le\mathbb{P}\left[\sup_{0\le t\le L(N)}\frac{1}{1+K_0}\left|\int_{u-t}^{u}\mathrm{d}w_i(r)\frac{1}{N}\sum_{i\in[N]}\sqrt{g(x_i^{[N^2]}(r))}\right|>\frac{\epsilon}{2}\right].
\end{aligned} \tag{8.80}$$

CHAPTER 8

Via the same optional stopping time argument as in the proof of Lemma 6.2.15, the above computation shows that (8.76) holds. Note that the extra drift term $\frac{K_2 e_2}{N}$ does not have any influence. Equations (8.77)–(8.78) hold by a similar argument as given in the proof of Lemma 7.2.2. □

§8.3.5 Limiting evolution for the single components

Proposition 8.3.5 (Equilibrium for the infinite system). *Fix $t_2, t_1 > 0$. Let $(N_k)_{k\in\mathbb{N}} \subset \mathbb{N}$ and let $L(N)$ be any sequence satisfying $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$ such that*

$$\begin{aligned}
&\lim_{k\to\infty} \mathcal{L}\left[\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right] = P_{t_1,t_2},\\
&\lim_{k\to\infty} \mathcal{L}\left[\left(Y_{1,0}^{[N_k^2]}(N_k^2 t_2 + N_k t_1), Y_{2,0}^{[N_k^2]}(N_k^2 t_2)\right) \middle| \boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right]\\
&\qquad = P^{z_1^{\mathrm{eff}}(t_1)},\\
&\lim_{k\to\infty} \mathcal{L}\Big[\sup_{0\le t\le L(N_k)} \left|\bar\Theta^{[N_k^2]}(N_k^2 t_2 + N_k t_1) - \bar\Theta^{[N_k^2]}(N_k^2 t_2 + N_k t_1 - t)\right|\\
&\qquad\qquad + \left|\Theta_{y_1}^{[N_k^2]}(N_k^2 t_2 + N_k t_1) - \Theta_{y_1}{}^{[N_k^2]}(N_k^2 t_2 + N_k t_1 - t)\right|\\
&\qquad\qquad + \left|\Theta_{y_2}^{[N_k^2]}(N_k^2 t_2 + N_k t_1) - \Theta_{y_2}{}^{[N_k^2]}(N_k^2 t_2 + N_k t_1 - t)\right|\Big] = \delta_0,\\
&\lim_{k\to\infty} \mathcal{L}\Big(Z^{[N_k^2]}(N_k^2 t_2 + N_k t_1),\Big) = \nu(t_1,t_2).
\end{aligned} \tag{8.81}$$

Then $\nu(t_1,t_2)$ is of the form

$$\nu(t_1,t_2) = \int_{[0,1]^2} P_{t_1,t_2}(\mathrm{d}\theta^{(1)}, \mathrm{d}\theta_y^{(1)}) \int_{[0,1]^{\mathbb{N}_0}} P^{(\theta^{(1)},\theta_y^{(1)})}(\mathrm{d}\boldsymbol{y})\, \nu_{\theta,\boldsymbol{y}}, \tag{8.82}$$

where

$$\nu_{\theta,\boldsymbol{y}_0} = \prod_{i\in\mathbb{N}_0} \Gamma_{(\theta,\boldsymbol{y}_{0,i})} \tag{8.83}$$

with $\Gamma_{(\theta,\boldsymbol{y}_{0,i})}$ the equilibrium measure for the i'th single colony defined in (8.24).

Note that by step 1 and step 3 we can find a subsequence $(N_k)_{k\in\mathbb{N}}$ such that the first and third line in (8.81) hold. The second line in (8.81) follows from assumptions (8.38) and (8.39). To prove Proposition 8.3.5 we proceed as in the proof of Proposition 7.2.3, but with the finite system in (7.4) replaced by the system in (8.6) and the infinite system in (7.11) replaced by the system in (8.18). Note that Lemma 7.2.4 holds also for the system in (8.18), after adding the non-interacting component $y_{2,0}$ to the equilibrium. The equivalent of Lemma 7.2.5 will again follow from the equivalent of Lemma 7.2.9. We will derive the analogue of Lemmas 7.2.6 and 7.2.7 (see Lemma's 8.3.6 and 8.3.7 below). Lemma 7.2.8 can be extended with an extra colour-2 seed-bank estimator by using the same proof. Since the infinite system for the single colonies in the two-layer three-colour mean-field system (see (8.89)) equals the one for the one-layer two-colour mean-field system, up to a non-interacting component, we obtain an equivalent of Lemma 7.2.9. Finally, also the equivalent of Lemma 7.2.10

holds under an additional assumption, see Lemma 8.3.8. Finally Corollary 8.3.9 states the equivalent of Corollary 7.2.11. With the help of the lemma's and the corollary, the proof of Proposition 8.3.5 follows from the same argument as used in the proof of Proposition 7.2.3.

Lemma 8.3.6 (Comparison of empirical averages).
Let $(\Theta_x^{(1),[N^2]}(t_0))_{t_0\geq 0}$ *and* $(\Theta_{y_0}^{(1),[N^2]}(t_0))_{t_0\geq 0}$ *be defined as in* (8.20). *Then*

$$\mathbb{E}\left[\left|\Theta_x^{(1),[N]}(t)-\Theta_{y_0}^{(1),[N]}(t)\right|\right] \leq \sqrt{\mathbb{E}\left[\left(\Theta_x^{(1),[N]}(0)-\Theta_{y_0}^{(1),[N]}(0)\right)^2\right]}\mathrm{e}^{-(K_0e_0+e_0)t} + \sqrt{\frac{1}{K_0e_0+e_0}\left[\frac{c_1}{N}+\frac{||g||}{N}+\frac{K_1e_1}{N}+\frac{K_2e_2}{N^2}\right]}. \tag{8.84}$$

Proof. The result follows by Itô-calculus on the SSDE in (8.6) and the same type of argument as used in the proof of Lemma 7.2.6. □

Like for the mean-field system with one colour, we need to compare the finite system in (8.6) with an infinite system. To derive the analogue of Lemma 7.2.7, let $L(N)$ satisfy $\lim_{N\to\infty} L(N)=\infty$ and $\lim_{N\to\infty} L(N)/N=0$. Define $[N]_i$ to be the 1-block that contains site $i\in[N^2]$. Since we start our system in an exchangeable measure and the dynamics are exchangeable, we will only consider the single colonies in $[N]_0$, the 1-block containing the site $0\in[N^2]$. In the rest of the prove, we will suppress the 0 from the notation i.e., $[N]_0=[N]$ and $\bar{\Theta}_0^{(1),[N^2]}=\bar{\Theta}^{(1),[N^2]}$. Define

$$u=N^2t_2+Nt_1 \tag{8.85}$$

and let μ_N be the measure on $([0,1]^3)^{\mathbb{N}_0}$ by continuing the configuration of

$$\begin{aligned}&\left(Z^{[N^2]}(u-L(N))\right)\\&=\left(X^{[N^2]}(u-L(N)),\left(Y_0^{[N^2]}(u-L(N)),Y_1^{[N^2]}(u-L(N)),Y_2^{[N^2]}(u-L(N))\right)\right)\end{aligned} \tag{8.86}$$

periodically to $([0,1]^4)^{\mathbb{N}_0}$, i.e., we continue the configuration of the single colonies in the first block to $([0,1]^4)^{\mathbb{N}_0}$. Let

$$\bar{\Theta}^{(1),[N^2]}=\frac{1}{N}\sum_{i\in[N]}\frac{x_i^{[N^2]}(u-L(N))+K_0y_{i,0}^{[N^2]}(u-L(N))}{1+K_0}. \tag{8.87}$$

Let

$$(Z^{\mu_N}(t))_{t\geq 0}=\left(X^{\mu_N}(t),(Y_0^{\mu_N}(t),Y_1^{\mu_N}(t),Y_2^{\mu_N}(t))\right)_{t\geq 0} \tag{8.88}$$

be the infinite system evolving according to

$$\begin{aligned}\mathrm{d}x_i^{\mu_N}(t)&=c_0\,[\bar{\Theta}^{(1),[N^2]}-x_i^{\mu_N}(t)]\,\mathrm{d}t+\sqrt{g\big(x_i^{\mu_N}(t)\big)}\,\mathrm{d}w_i(t)+K_0e_0\,[y_{i,0}^{\mu_N}(t)-x_i^{\mu_N}(t)]\,\mathrm{d}t,\\\mathrm{d}y_{i,0}^{\mu_N}(t)&=e_0\,[x_i^{\mu_N}(t)-y_{i,0}^{\mu_N}(t)]\,\mathrm{d}t,\\y_{i,1}^{\mu_N}(t)&=y_{i,1}^{\mu_N}(0),\\y_{i,2}^{\mu_N}(t)&=y_{i,2}^{\mu_N}(0),\qquad i\in\mathbb{N}_0,\end{aligned} \tag{8.89}$$

starting from initial distribution μ_N. Then the following Lemma 8.3.7 is the equivalent of Lemma 7.2.7 for the three-colour two-layer mean-field system. In particular, the infinite system considered in Lemma 8.3.7 is similar to the infinite system in Lemma 7.2.7. The only difference is that there is one more non-interacting component added in (8.89).

Lemma 8.3.7. *[Comparison of finite and infinite systems] Fix $t_1, t_2 > 0$, and let $u = N^2 t_2 + N t_1$. Let $L(N)$ satisfy $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N$. Suppose that*

$$\lim_{N\to\infty} \sup_{0\leq t\leq L(N)} \left|\bar{\Theta}^{(1),[N]}(u) - \bar{\Theta}^{(1),[N]}(u-t)\right| = 0 \quad \textit{in probability.} \tag{8.90}$$

Then, for all $t \geq 0$,

$$\lim_{k\to\infty} \left|\mathbb{E}\left[f\big(Z^{\mu_N}(t)\big) - f\big(Z^{[N^2]}(u - L(N) + t)\big)\right]\right| = 0 \qquad \forall\, f \in \mathcal{C}\big(([0,1]^3)^{\mathbb{N}_0}, \mathbb{R}\big). \tag{8.91}$$

Proof. We proceed as in the proof of Lemma 7.2.7 and couple the finite and infinite systems by their Brownian motion, exactly as was done there. The single components in the block around site 0 of the finite process $(Z^{[N^2]}(t))$ are evolving according to

$$\begin{aligned}
\mathrm{d}x_i^{[N^2]}(t) &= c_0\left[\Theta^{(1),[N^2]} - x_i^{[N^2]}(t)\right]\mathrm{d}t + c_0\left[\bar{\Theta}^{(1),[N^2]}(t) - \Theta^{(1),[N^2]}\right]\mathrm{d}t\\
&\quad + c_0\left[\Theta_x^{(1),[N^2]}(t) - \bar{\Theta}^{(1),[N^2]}(t)\right]\mathrm{d}t + \frac{c_1}{N}\left[\frac{1}{N^2}\sum_{i\in[N^2]} x_j^{[N^2]}(t) - x_i^{[N^2]}(t)\right]\mathrm{d}t\\
&\quad + \sqrt{g(x_i^{[N^2]}(t))}\,\mathrm{d}w_i(t) + K_0 e_0\,[y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t\\
&\quad + \frac{K_1 e_1}{N}\left[y_{i,1}^{[N^2]}(t) - x_i^{[N^2]}(t)\right]\mathrm{d}t + \frac{K_2 e_2}{N^2}\,[y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,0}^{[N^2]}(t) &= e_0\,[x_i^{[N^2]}(t) - y_{i,0}^{[N^2]}(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,1}^{[N^2]}(t) &= \frac{e_1}{N}\,[x_i^{[N^2]}(t) - y_{i,1}^{[N^2]}(t)]\,\mathrm{d}t,\\
\mathrm{d}y_{i,2}^{[N^2]}(t) &= \frac{e_2}{N^2}\,[x_i^{[N^2]}(t) - y_{i,2}^{[N^2]}(t)]\,\mathrm{d}t, \qquad i\in[N].
\end{aligned} \tag{8.92}$$

Using this SSDE we can exactly proceed as in the proof of Lemma 7.2.7 to obtain the result. Note that the colour-2 seed-bank can be treated just in the same way as the colour-1 seed-bank in the proof of Lemma 7.2.7, since its rate of interaction with the active population is even slower than the rate of interaction of the colour-1 seed-bank. □

Finally, we state the equivalent of Lemma 7.2.10 for the three-colour two-layer mean-field system.

Lemma 8.3.8 (Coupling of finite systems). *Let*

$$Z^{[N^2],1} = (X^{[N^2],1}, Y_0^{[N^2],1}, Y_1^{[N^2],1}, Y_2^{[N^2],1}) \tag{8.93}$$

Chapter 8

be the finite system evolving according to (8.6) *starting from an exchangeable initial measure. Let* $\mu^{[N],1}$ *be the measure obtain by periodic continuation of the configuration of* $Z^{[N^2],1}(0)$ *in the* 1*-block around* 0*. Similarly, let*

$$Z^{[N^2],2} = (X^{[N^2],2}, Y_0^{[N^2],2}, Y_1^{[N^2],2}, Y_2^{[N^2],2}) \tag{8.94}$$

be the finite system evolving according to (8.6) *starting from an exchangeable initial measure. Let* $\mu^{[N],2}$ *be the measure obtained by periodic continuation of the configuration of* $Z^{[N^2],1}(0)$ *in the* 1*-block around* 0*. Let* $\tilde{\mu}$ *be any weak limit point of the sequence of measures* $(\mu^{[N],1} \times \mu^{[N],2})_{N\in\mathbb{N}}$*. Define the variables* $\bar{\Theta}^{[N],1}$ *on* $(([0,1]^4, \mu^{[N],1})^{\mathbb{N}_0})$*,* $\bar{\Theta}^{[N],2}$ *on* $(([0,1]^4)^{\mathbb{N}_0}, \mu^{[N],2})$ *and* $\bar{\Theta}_1$ *and* $\bar{\Theta}_2$ *on* $(([0,1]^4)^{\mathbb{N}_0} \times ([0,1]^4)^{\mathbb{N}_0}, \mu)$ *by*

$$\begin{aligned} \bar{\Theta}^{[N],1} &= \frac{1}{N}\sum_{i\in[N]} \frac{x_i^{[N^2],1} + K_0 y_{i,0}^{[N^2],1}}{1+K_0}, \qquad & \bar{\Theta}^{[N],2} &= \frac{1}{N}\sum_{i\in[N]} \frac{x_i^{[N^2],2} + K_0 y_{i,0}^{[N^2],2}}{1+K_0}, \\ \bar{\Theta}^1 &= \lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]} \frac{x_i^1 + K_0 y_{i,0}^1}{1+K_0}, \qquad & \bar{\Theta}^2 &= \lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]} \frac{x_i^2 + K_0 y_{i,0}^2}{1+K_0}, \end{aligned} \tag{8.95}$$

and let $(\bar{\Theta}^{(1),[N],1}(t))_{t\geq 0}$ *and* $(\bar{\Theta}^{(1),[N],2}(t))_{t\geq 0}$ *be defined as in* (7.14) *for* $Z^{[N^2],1}$*, respectively,* $Z^{[N^2],2}$*. Suppose that*

$$\lim_{N\to\infty}\sup_{0\leq t\leq L(N)} \left(\left|\bar{\Theta}^{[N],k}(0) - \bar{\Theta}^{[N],k}(t)\right|\right) = 0 \quad \textit{in probability}, \qquad k\in\{1,2\}, \tag{8.96}$$

and suppose that $\tilde{\mu}(\{\bar{\Theta}_1 = \bar{\Theta}_2,\, Y_1^1 = Y_1^2, Y_2^1 = Y_2^2\}) = 1$*. Then, for any sequence* $(t(N))_{N\in\mathbb{N}}$ *with* $\lim_{N\to\infty} t(N) = \infty$*,*

$$\begin{aligned} \lim_{N\to\infty} \mathbb{E}\big[&|x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K_0|y_{i,0}^{[N],1}(t(N)) - y_{i,0}^{[N],2}(t(N))| \\ &+ K_1|y_{i,1}^{[N],1}(t(N)) - y_{i,1}^{[N],2}(t(N))| + K_2|y_{i,2}^{[N],1}(t(N)) - y_{i,2}^{[N],2}(t(N))|\big] = 0. \end{aligned} \tag{8.97}$$

Proof. Like in the proof of Lemma 7.2.10, we can show with Itô calculus that the function

$$\begin{aligned} t \to \mathbb{E}\big[&|x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K_0|y_{i,0}^{[N],1}(t(N)) - y_{i,0}^{[N],2}(t(N))| \\ &+ K_1|y_{i,1}^{[N],1}(t(N)) - y_{i,1}^{[N],2}(t(N))| + K_2|y_{i,2}^{[N],1}(t(N)) - y_{i,2}^{[N],2}(t(N))|\big] \end{aligned} \tag{8.98}$$

is monotonically decreasing. Hence we can proceed as in the proof of Lemma 6.2.13 to show that (8.97) is true. □

From the above couplings we can derive the following corollary, which is the analogue of Corollary 7.2.11 for the two-level three-colour mean-field system.

Corollary 8.3.9. *Fix* $t_1, t_2 > 0$ *and set* $u = N^2 t_2 + N t_1$*. Let* μ_N *be the measure obtained by periodic continuation of*

$$Z^{[N^2]}(u-L(N)) = \big(X^{[N^2]}(u-L(N)), Y_0^{[N^2]}(u-L(N)), Y_1^{[N^2]}(u-L(N)), Y_2^{[N^2]}(u-L(N))\big), \tag{8.99}$$

and let μ be a weak limit point of the sequence $(\mu_N)_{N\in\mathbb{N}}$. Let

$$\Theta = \lim_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} \frac{x_i^\mu + K y_i^\mu}{1+K} \qquad \text{in } L^2(\mu), \tag{8.100}$$

and let $(Z^{\nu_\Theta}(t))_{t>0} = (X^{\nu_\Theta}(t), Y_0^{\nu_\Theta}(t), Y_1^{\nu_\Theta}(t), Y_2^{\nu_\Theta}(t))_{t>0}$ be the infinite system with components evolving according to (8.18) with $\theta = \Theta$ and $y_{i,1,0}$ and $y_{i,2,0}$ determined by assumption (8.81) and starting from its equilibrium measure. Extend the finite system $Z^{[N^2]}$ as a system on $([0,1]^4)^{\mathbb{N}_0}$ by periodic continuation. Construct $(Z^{[N^2]}(t))_{t>0}$ and $(Z^{\nu_\Theta}(t))_{t>0}$ on one probability space. Then there exists a sequence $(\bar{L}(N))_{N\in\mathbb{N}}$ such that $\lim_{N\to\infty} \bar{L}(N) = \infty$, $\lim_{N\to\infty} \frac{\bar{L(N)}}{N} = 0$ and

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\left|x_i^{[N^2]}(Ns) - x_i^{\nu_\Theta}(\bar{L}(N))\right|\right] + K_0\,\mathbb{E}\left[\left|y_{i,0}^{[N^2]}(Ns) - y_{i,0}^{\nu_\Theta}(\bar{L}(N))\right|\right]\\
&+ K_1\,\mathbb{E}\left[\left|y_{i,1}^{[N^2]}(Ns) - y_{i,1}^{\nu_\Theta}(\bar{L}(N))\right|\right] + K_2\,\mathbb{E}\left[\left|y_{i,2}^{[N^2]}(Ns) - y_{i,2}^{\nu_\Theta}(\bar{L}(N))\right|\right] = 0,\ i\in[N].
\end{aligned} \tag{8.101}$$

Note that Lemmas 8.3.7, 8.3.8 and Corollary 8.3.9 do not only hold for sites i in the 1-block around 0, but hold for for all sites $i \in [N^2]$, after we replace $\Theta_0^{[N]_0}$ by $\Theta_i^{[N]_i}$.

§8.3.6 Limiting evolution of the 1-block estimator process

Proposition 8.3.10 (Limiting evolution of the 1-blocks). *Fix $t_2 > 0$. Let $(L(N))_{N\in\mathbb{N}}$ satisfy $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$. Let $(N_k)_{k\in\mathbb{N}}$ be a subsequence such that*

$$\begin{aligned}
&\lim_{k\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2 t_2)\right)\right] = P_{t_2}(\cdot),\\
&\lim_{k\to\infty} \mathcal{L}\left[y_{2,1}^{[N_k^2]}(N_k t_2)\middle|\boldsymbol{\Theta}^{(2),[N_k^2]}(N_k^2 t_2)\right] = P^{z_2(t_2)},\\
&\lim_{k\to\infty} \mathcal{L}\left[\left(Y_{1,0}^{[N_k^2]}(N_k^2 t_2 + N_k t_1), Y_{2,0}^{[N_k^2]}(N_k^2 t_2)\right)\middle|\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right] = P^{z_1^{\mathrm{eff}}(t_1)},\\
&\lim_{k\to\infty} \mathcal{L}\Bigg[\sup_{0\le t\le L(N_k)} \left|\bar{\Theta}^{(2),[N_k^2]}(N_k^2 t_2) - \bar{\Theta}^{(2),[N_k^2]}(N_k^2 t_2 - N_k t)\right|\\
&\qquad\qquad + \left|\Theta_{y_2}^{(2),[N_k^2]}(N_k^2 t_2) - \Theta_{y_2}{}^{(2),[N_k^2]}(N_k^2 t_2 - Nt)\right|\Bigg] = \delta_0.
\end{aligned} \tag{8.102}$$

Then, for the 1-block around 0,

$$\lim_{k\to\infty} \mathcal{L}\left[\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2)\right] = \int_{[0,1]^2} \int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v), \tag{8.103}$$

where $\Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}$ is the equilibrium measure of (8.29) *with θ replaced by u, and*

$$\lim_{k\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right)_{t_1>0}\right] = \mathcal{L}\left[(z_1^{\mathrm{aux}}(t_1))_{t_1>0})\right], \tag{8.104}$$

where $(z_1^{\mathrm{aux}}(t_1))_{t_1>0}$ is the process evolving according to (8.29) *with θ replaced by the random variable $\bar{\Theta}^{(2)}(t_2)$ and with initial measure $\int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v)$, and $y_{2,1}$ is a random variable.*

Note that by tightness of the 2-blocks and the assumptions in Proposition 8.1.1, we can always find a subsequence $(N_k)_{k\in\mathbb{N}}$ such that (8.102) holds and also (8.81) holds. To prepare for the proof of Proposition 8.3.10, we prove four lemmas: Lemma 8.3.11 shows that the limiting 1-block system has a unique equilibrium, Lemma 8.3.13 implies convergence of the active 2-block estimator and the combined 2-block estimator, Lemma 8.3.14 gives a regularity property for the 2-block estimator, and Lemma 8.3.15 shows the limiting evolution of the auxiliary 1-block estimator process. Lemma 8.3.17 proves equation (8.103). After we derive these lemmas we prove Proposition 8.3.10.

Lemma 8.3.11 (1-block equilibrium). *For any initial distribution $\mu \in ([0,1]^3)$, the process $(z_1^{\mathrm{aux}}(t_1))_{t_1>0}$ evolving according to* (8.29) *is well defined and converges to a unique equilibrium measure*

$$\lim_{t_1\to\infty} \mathcal{L}[z_1^{\mathrm{aux}}(t_1)] = \Gamma^{\mathrm{aux},(1)}_{\theta,y_{2,1}}. \tag{8.105}$$

Proof. By [72], the SSDE in (8.29) has a unique strong solution. By a similar argument as in the proof of Lemma 7.2.4, the SSDE in (8.29) converges to a unique equilibrium measure $\Gamma^{\mathrm{aux},(1)}_{\theta,y_{2,1}}$. □

Remark 8.3.12 (Equilibrium measure). Note that Lemma 8.3.11 still holds when we allow θ and $y_{2,1}$ to be the random variables $\bar{\Theta}(t_2)$ and $y_{2,1}$. Assuming (8.102), we can derive the distributions of $\bar{\Theta}(t_2)$ and $y_{2,1}$, and we can write the equilibrium as $\int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v)$. In what follows we abbreviate

$$\Gamma^{(1)}_{\bar{\Theta}(t_2),y_{2,1,i}} = \int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1,i}}\, P^{(u,v)}(\mathrm{d}y_{2,1,i})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v). \tag{8.106}$$

■

Lemma 8.3.13 (2-block averages). *Define*

$$\Delta^{(2),[N^2]}_{\Sigma}(Nt_1) = \frac{\Theta^{(2),[N^2]}_x(Nt_1) + K_0\Theta^{(2),[N^2]}_{y_0}(Nt_1)}{1+K_0} - \Theta^{(2),[N^2]}_{y_1}(Nt_1). \tag{8.107}$$

Then

$$\begin{aligned}
&\mathbb{E}\left[\left|\Delta_\Sigma^{(2),[N^2]}(Nt_1)\right|\right] \\
&\leq \sqrt{\mathbb{E}\left[\left(\Delta_\Sigma^{(2),[N^2]}(0)\right)^2\right]}\mathrm{e}^{-e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)t_1} \\
&+ \sqrt{\int_0^{t_1} \mathrm{d}s\, 2e_1 \left(\frac{1+K_0+K_1}{1+K_0}\right) \mathrm{e}^{-2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)(t_1-s)} \mathbb{E}\left[\left|\bar{\Theta}^{(1),[N^2]}(Ns) - \Theta_x^{(1),[N^2]}(Ns)\right|\right]} \\
&+ \sqrt{\frac{1}{e_1}\left[\frac{K_2e_2}{N(1+K_0+K_1)} + \frac{||g||}{2N(1+K_0+K_1)}\right]}.
\end{aligned} \tag{8.108}$$

Proof. For the two-level mean-field system we have the following SSDE for the 2-block averages:

$$\begin{aligned}
\mathrm{d}\Theta_x^{(2),[N^2]}(Nt_1) &= \sqrt{\frac{1}{N^3}\sum_{i\in[N^2]} g(x_i^{[N^2]}(Nt_1))\,\mathrm{d}\tilde{w}(t_1)} \\
&\quad + NK_0e_0\left[\Theta_{y_0}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1 \\
&\quad + K_1e_1\left[\Theta_{y_1}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1 \\
&\quad + \frac{K_2e_2}{N}\left[\Theta_{y_2}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1, \\
\mathrm{d}\Theta_{y_0}^{(2),[N^2]}(Nt_1) &= Ne_0\left[\Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_0}^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1, \\
\mathrm{d}\Theta_{y_1}^{(2),[N^2]}(Nt_1) &= e_1\left[\Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_1}^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1, \\
\mathrm{d}\Theta_{y_2}^{(2),[N^2]}(Nt_1) &= \frac{e_2}{N}\left[\Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_2}^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1.
\end{aligned} \tag{8.109}$$

Therefore

$$\begin{aligned}
&\mathrm{d}\left(\Delta_\Sigma^{(2),[N^2]}(Nt_1)\right)^2 = 2\Delta_\Sigma^{(2),[N^2]}(Nt_1)\,\mathrm{d}\Delta_\Sigma^{(2),[N^2]}(Nt_1) + \mathrm{d}\left\langle \Delta_\Sigma^{(2),[N^2]}(Nt_1)\right\rangle \\
&= 2\Delta_\Sigma^{(2),[N^2]}(Nt_1)\,\frac{1}{1+K_0}\sqrt{\frac{1}{N^3}\sum_{i\in[N^2]} g(x_i^{[N^2]}(Nt_1))\,\mathrm{d}\tilde{w}(t_1)} \\
&\quad + 2\Delta_\Sigma^{(2),[N^2]}(Nt_1)\,\frac{K_1e_1}{(1+K_0)}\left[\Theta_{y_1}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1, \\
&\quad + 2\Delta_\Sigma^{(2),[N^2]}(Nt_1)\,\frac{K_2e_2}{N(1+K_0)}\left[\Theta_{y_2}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1, \\
&\quad - 2\Delta_\Sigma^{(2),[N^2]}(Nt_1)\,e_1\left[\Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_1}^{(2),[N^2]}(Nt_1)\right]\mathrm{d}t_1 \\
&\quad + \frac{1}{(1+K_0)^2}\frac{1}{N^3}\sum_{i\in[N^2]} g(x_i^{[N^2]}(Nt_1))\,\mathrm{d}t_1.
\end{aligned} \tag{8.110}$$

Hence

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\left(\Delta_{\Sigma}^{(2),[N^2]}(Nt_1)\right)^2\right]\\
&= -2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)\mathbb{E}\left[\left(\Delta_{\Sigma}^{(2),[N^2]}(Nt_1)\right)^2\right]\\
&\quad + 2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)\\
&\qquad \times \mathbb{E}\left[\Delta_{\Sigma}^{(2),[N^2]}(Nt_1)\left(\frac{\Theta_x^{(2),[N^2]}(Nt_1)+K_0\Theta_{y_0}^{(2),[N^2]}(Nt_1)}{1+K_0}-\Theta_x^{(2),[N^2]}(Nt_1)\right)\right]\\
&\quad + \frac{K_2e_2}{N(1+K_0)}\,2\mathbb{E}\left[\Delta_{\Sigma}^{(2),[N^2]}(Nt_1)\,\left[\Theta_{y_2}^{(2),[N^2]}(Nt_1)-\Theta_x^{(2),[N^2]}(Nt_1)\right]\right]\\
&\quad + \frac{1}{(1+K_0)^2}\mathbb{E}\left[\frac{1}{N^3}\sum_{i\in[N^2]} g(x_i^{[N^2]}(Nt_1))\right],
\end{aligned} \tag{8.111}$$

and therefore

$$\begin{aligned}
&\mathbb{E}\left[\left(\Delta_{\Sigma}^{(2),[N^2]}(Nt_1)\right)^2\right]\\
&= \mathbb{E}\left[\left(\Delta_{\Sigma}^{(2),[N^2]}(0)\right)^2\right]\mathrm{e}^{-2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)t_1} + \int_0^{t_1}\mathrm{d}s\,\mathrm{e}^{-2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)(t_1-s)}h^{[N]}(s),
\end{aligned} \tag{8.112}$$

where

$$\begin{aligned}
h^{[N]}(s) &= 2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)\\
&\quad \times \mathbb{E}\left[\Delta_{\Sigma}^{(2),[N^2]}(Ns)\left(\frac{\Theta_x^{(2),[N^2]}(Ns)+K_0\Theta_{y_0}^{(2),[N^2]}(Ns)}{1+K_0}-\Theta_x^{(2),[N^2]}(Ns)\right)\right]\\
&\quad + \frac{2K_2e_2}{N(1+K_0)}\mathbb{E}\left[\Delta_{\Sigma}^{(2),[N^2]}(Ns)\,\left[\Theta_{y_2}^{(2),[N^2]}(Ns)-\Theta_x^{(2),[N^2]}(Ns)\right]\right]\\
&\quad + \frac{1}{(1+K_0)^2}\mathbb{E}\left[\frac{1}{N^3}\sum_{i\in[N^2]} g(x_i^{[N^2]}(Ns))\right].
\end{aligned} \tag{8.113}$$

Therefore

$$\begin{aligned}
&\mathbb{E}\left[\left|\Delta_{\Sigma}^{(2),[N^2]}(Nt_1)\right|\right]\\
&\leq \sqrt{\mathbb{E}\left[\left(\Delta_{\Sigma}^{(2),[N^2]}(0)\right)^2\right]}\mathrm{e}^{-e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)t_1}\\
&+ \sqrt{\int_0^{t_1}\mathrm{d}s\,2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)\mathrm{e}^{-2e_1\left(\frac{1+K_0+K_1}{1+K_0}\right)(t_1-s)}\mathbb{E}\left[\left|\bar{\Theta}^{(1),[N^2]}(Ns)-\Theta_x^{(1),[N^2]}(Ns)\right|\right]}\\
&+ \sqrt{\frac{1}{e_1}\left[\frac{K_2e_2}{N(1+K_0+K_1)}+\frac{||g||}{2N(1+K_0+K_1)}\right]}.
\end{aligned} \tag{8.114}$$

□

Let μ_{N_k} be the measure obtained by periodic continuation of the configuration

$$Z^{[N^2]}(N_k^2 t_2). \tag{8.115}$$

Since the state space $([0,1] \times [0,1]^3)^{\mathbb{N}_0}$ is compact, we can pass to a further subsequence, to obtain

$$\mu = \lim_{k\to\infty} \mu_{N_k}. \tag{8.116}$$

Lemma 8.3.14 (Regularity for 2-block estimator). *Let μ and μ_N be as defined above. Let $(x_i, y_{1,i}, y_{2,i})_{i\in\mathbb{N}_0}$ be distributed according to μ. Define the random variable*

$$\begin{aligned}
&\phi = (\phi_1, \phi_2),\\
&\phi_1 = \lim_{n\to\infty} \frac{1}{n^2} \sum_{i\in[n^2]} \frac{x_i + K_0 y_{i,0} + K_1 y_{i,1}}{1+K_0+K_1}, \qquad \phi_2 = \lim_{n\to\infty} \frac{1}{n^2} \sum_{i\in[n^2]} y_{i,2},
\end{aligned} \tag{8.117}$$

and the random variable $\phi^{[N]}$ on $(\mu_N, ([0,1]^3)^{\mathbb{N}_0})$ by putting

$$\begin{aligned}
&\phi^{[N^2]} = (\phi_1^{[N^2]}, \phi_2^{[N^2]}),\\
&\phi_1^{[N^2]} = \frac{1}{N^2} \sum_{i\in[N^2]} \frac{x_i^{[N^2]} + K_0 y_{i,0}^{[N^2]} + K_1 y_{i,1}^{[N^2]}}{1+K_0+K_1}, \qquad \phi_2 = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i\in[N^2]} y_{i,2}^{[N^2]}.
\end{aligned} \tag{8.118}$$

Then

$$\lim_{N\to\infty} \mathcal{L}[\phi^{[N^2]}] = \mathcal{L}[\phi]. \tag{8.119}$$

Proof. Use a similar argument as in the proof of Lemma 7.2.8. □

We will first determine the limiting evolution of $(\mathbf{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N t_1))_{t_1>0}$. To do so we consider all the N_k 1-blocks in $[N_k^2]$. After that we show that

$$\lim_{k\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}_i^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2)\right)_{i\in[N_k]}\right] = \prod_{i\in\mathbb{N}_0} \Gamma^{(1)}_{\bar{\Theta}(t_2), y_{2,1,i}}, \tag{8.120}$$

The limiting 1-block process for the auxiliary estimator process (recall (8.29)) is given by

$$\begin{aligned}
&(\mathbf{z}_1^{\mathrm{aux}}(t))_{t>0} = (\mathbf{x}_1^{\mathrm{aux}}(t), \mathbf{y}_{1,1}^{\mathrm{aux}}(t), \mathbf{y}_{2,1}^{\mathrm{aux}}(t))_{t>0},\\
&\mathbf{z}_1^{\mathrm{aux}}(t) = (z_{1,i}^{\mathrm{aux}}(t))_{i\in\mathbb{N}_0}, \qquad \mathbf{x}_1^{\mathrm{aux}}(t) = (x_{1,i}^{\mathrm{aux}}(t))_{i\in\mathbb{N}_0},\\
&\mathbf{y}_{1,1}^{\mathrm{aux}}(t) = (y_{1,1,i}^{\mathrm{aux}}(t))_{i\in\mathbb{N}_0} \qquad \mathbf{y}_{2,1}^{\mathrm{aux}}(t) = (y_{2,i}^{\mathrm{aux}}(t))_{i\in\mathbb{N}_0}
\end{aligned} \tag{8.121}$$

and its components evolve according to

$$\begin{aligned}
\mathrm{d}x_{1,i}^{\mathrm{aux}}(t) = \frac{1}{1+K_0}\Bigg[&c_1[\bar{\Theta}^{(2)}(t_2) - x_{1,i}^{\mathrm{aux}}(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(1)}g)(x_{1,i}^{\mathrm{aux}}(t))}\,\mathrm{d}w(t)\\
&+ K_1 e_1\,[y_{1,1,i}^{\mathrm{aux}}(t) - x_{1,i}^{\mathrm{aux}}(t)]\,\mathrm{d}t\Bigg],
\end{aligned} \tag{8.122}$$

$$\begin{aligned}
\mathrm{d}y_{1,1,i}^{\mathrm{aux}}(t) &= e_1\,[x_{1,i}^{\mathrm{aux}}(t) - y_{1,1,i}^{\mathrm{aux}}(t)]\,\mathrm{d}t,\\
y_{2,1,i}^{\mathrm{aux}}(t) &= y_{2,1,i}, \qquad i\in\mathbb{N}_0,
\end{aligned}$$

Chapter 8

where

$$\bar{\Theta}^{(2)}(t_2) = \lim_{N\to\infty} \sum_{i\in[N^2]} \frac{x_i^{[N^2]} + K_0 y_{i,0}^{[N^2]} + K_1 y_{i,1}^{[N^2]}}{1+K_0+K_1} \text{ in } L_2(\mu). \tag{8.123}$$

Let $\mu_{N_k}^{(1)}$ be the law obtained by periodic continuation of $(\boldsymbol{\Theta}_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2))_{i\in[N_k]}$, and let $\mu^{(1)} = \lim_{k\to\infty} \mu_{N_k}^{(1)}$ be any weak limit point of the sequence $(\mu_{N_k}^{(1)})_{k\in\mathbb{N}}$.

Lemma 8.3.15 (Limiting evolution of auxiliary 1-block estimator). *Let $\mathcal{L}[(\boldsymbol{z}_1^{\text{aux}}(0))] = \mu^{(1)}$. Then the following hold.*

(a) *For all $t_1 > 0$ and $i \in [N_k]$,*

$$\begin{aligned}\lim_{k\to\infty} \mathbb{E}\Bigg[&(1+K_0)\left(x_{1,i}^{\text{aux}}(t_1) - \bar{\Theta}_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right)^2\\ &+ K_1\left(y_{1,1,i}(t_1) - \Theta_{y_1,i}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1)\right)^2\\ &+ K_2\left(y_{2,1,i}^{\text{aux},(1),[N_k^2]}(t_1) - \Theta_{y_2,i}^{\text{aux},(1),[N_k]}(N_k^2 t_2 + N_k t_1)\right)^2\Bigg] = 0.\end{aligned} \tag{8.124}$$

(b) *For all $t_2 > 0$,*

$$\lim_{k\to\infty} \mathcal{L}\left[(\boldsymbol{\Theta}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N t_1))_{t_1>0}\right] = \mathcal{L}[(\boldsymbol{z}_1^{\text{aux}}(t_1))_{t_1>0}]. \tag{8.125}$$

Proof. Abbreviate

$$\begin{aligned}\Delta_i^{(1),[N_k^2]}(N_k t_1) &= x_{1,i}^{\text{aux}}(t_1) - \Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1),\\ \delta_{y_1,i}^{(1),[N_k^2]}(N_k t_1) &= y_{1,1,i}(t_1) - \Theta_{y_1,i}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1),\\ \delta_{y_2,i}^{(1),[N_k^2]}(N_k t_1) &= y_{2,1,i}^{\text{aux}}(t_1) - \Theta_{y_2,i}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1).\end{aligned} \tag{8.126}$$

Extending $(\boldsymbol{\Theta}^{\text{aux},(1),[N_k^2]}(N_k t_1))_{t_1>0}$ as a process on $\mathbb{N}_0$ by periodic continuation, we can construct $(\mathbf{z}_1^{\text{aux}}(t_1))_{t_1>0}$ and $(\boldsymbol{\Theta}^{\text{aux},(1)}(N_k t_1))_{t_1>0}$ on one probability space such that

$$\lim_{k\to\infty} \boldsymbol{\Theta}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2) = \mathbf{z}_1^{\text{aux},(1)}(0) \quad a.s. \tag{8.127}$$

We couple the processes $(\mathbf{z}_1^{\text{aux}}(t_1))_{t_1>0}$ and $(\boldsymbol{\Theta}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{t_1>0}$ by using the same Brownian motions for both processes. By Itô-calculus we obtain for the coupled process (recall (8.74))

$$
\begin{aligned}
&\mathbb{E}\left[(1+K_0)\left(\Delta_i^{(1),[N_k^2]}(N_k t_1)\right)^2 + K_1\left(\delta_{y_1,i}^{(1),[N_k^2]}(N_k t_1)\right)^2 + K_2\left(\delta_{y_2,i}^{(1),[N_k^2]}(N_k t_1)\right)^2\right]\\
=&\mathbb{E}\left[(1+K_0)\left(\Delta_i^{(1),[N_k^2]}(0)\right)^2 + K_1\left(\delta_{y_1,i}^{(1),[N_k^2]}(0)\right)^2 + K_2\left(\delta_{y_2,i}^{(1),[N_k^2]}(0)\right)^2\right]\\
&-2c_1\int_0^{t_1}\mathbb{E}\left[\left(\Delta_i^{(1),[N_k^2]}(N_k s)\right)^2\right]\mathrm{d}s\\
&-2K_1e_1\int_0^{t_1}\mathbb{E}\left[\left(\Delta_i^{(1),[N_k^2]}(N_k s)-\delta_{y_1,i}^{(1),[N_k^2]}(N_k s)\right)^2\right]\mathrm{d}s\\
&+2c_1\int_0^{t_1}\mathbb{E}\left[\Delta_i^{(1),[N_k^2]}(N_k s)\left(\Theta^{(2)}(t_2)-\frac{1}{N_k^2}\sum_{j\in[N_k^2]}x_j^{[N_k^2]}(N_k^2t_2+N_ks)\right)\right]\mathrm{d}s\\
&+(K_1e_1+c_1)\int_0^{t_1}\mathbb{E}\Bigg[\Delta_i^{(1),[N_k^2]}(N_k s)\\
&\qquad\times\left[\frac{1}{N_k}\sum_{j\in[N_k]_i}x_j^{[N_k^2]}(N_k^2t_2+N_ks)-\bar{\Theta}_i^{\mathrm{aux},(1),[N_k^2]}(N_k^2t_2+N_ks)\right]\Bigg]\mathrm{d}s\\
&+2K_1e_1\int_0^{t_1}\mathbb{E}\Bigg[\delta_{y_1,i}^{(1),[N_k^2]}(N_k s)\\
&\qquad\times\left[\bar{\Theta}_i^{\mathrm{aux},(1),[N_k^2]}(N_k^2t_2+N_ks)-\frac{1}{N_k}\sum_{j\in[N_k]_i}x_j^{[N_k^2]}(N_k^2t_2+N_ks)\right]\Bigg]\mathrm{d}s\\
&+2\frac{K_2e_2}{N_k}\int_0^{t_1}\mathbb{E}\Bigg[\left[\delta_{y_1,i}^{(1),[N_k^2]}(N_k s)-\Delta_i^{(1),[N_k^2]}(N_k^2t_2+N_ks)\right]\\
&\qquad\times\left[\frac{1}{N_k}\sum_{j\in[N_k]_i}x_j^{[N_k^2]}(N_k^2s)-\Theta_{y_2,i}^{\mathrm{aux},(1),[N_k^2]}(N_k^2t_2+Nt_1)\right]\Bigg]\mathrm{d}s\\
&+(1+K_0)^2\int_0^{t_1}\mathbb{E}\left[\left(\sqrt{(\mathcal{F}g)\left(x_1^{\mathrm{aux}}(s)\right)}-\sqrt{\frac{1}{N}\sum_{i\in[N]}g(x_i^{[N_k^2]}(N_k^2t_2+N_ks))}\right)^2\right]\mathrm{d}s.
\end{aligned}
\tag{8.128}
$$

Note that $|\Delta_i^{(1),[N_k^2]}|\le 1$ and $|\delta_{y_1,i}^{(1),[N_k^2]}|\le 1$. Note that the first term tends to 0 by (8.127). We show by dominated convergence that also all other positive terms in the right-hand side of (8.128) tend to 0 as $k\to\infty$.

For the third term, we can estimate

$$
\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}\left[\frac{c_1}{1+K_0}\left|\frac{1}{N_k^2}\sum_{i\in[N_k^2]} x_j^{[N_k^2]}(N_k^2 t_2 + N_k s) - \bar{\Theta}^{(2)}(t_2)\right|\right] \\
&\le \lim_{k\to\infty} \mathbb{E}\left[\frac{c_1}{1+K_0}\left|\Theta_x^{(2),[N_k^2]}(N_k^2 t_2 + N_k s)\right.\right. \\
&\qquad\left.\left. - \frac{\Theta_x^{(2),[N_k^2]}(N_k^2 t_2 + N_k s) + K_0\Theta_{y_0}^{(2),[N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0}\right|\right] \\
&\quad + \mathbb{E}\left[\frac{c_1}{1+K_0}\left|\frac{\Theta_x^{(2),[N_k^2]}(N_k^2 t_2 + N_k s) + K_0\Theta_{y_0}^{(2),[N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0}\right.\right. \\
&\quad \left.\left. - \frac{\Theta_x^{(2),[N_k^2]}(N_k^2 t_2 + N_k s) + K_0\Theta_{y_0}^{(2),[N_k^2]}(N_k^2 t_2 + N_k s) + K_1\Theta_{y_1}^{(2),[N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0+K_1}\right|\right] \\
&\quad + \mathbb{E}\left[\frac{c_1}{1+K_0}\left|\frac{\Theta_x^{(2),[N_k^2]}(N_k^2 t_2 + N_k s) + K_0\Theta_{y_0}^{(2),[N_k^2]}(N_k^2 t_2 + N_k s) + K_1\Theta_{y_1}^{(2),[N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0+K_1}\right.\right. \\
&\qquad\left.\left. - \bar{\Theta}^{(2)}(t_2)\right|\right].
\end{aligned}
\tag{8.129}
$$

The first term in (8.129) tends to zero by Lemma 8.3.6, the second term tends to zero by Lemma 8.3.13, while the third term tends to zero by Lemma 8.3.14 and Lemma 8.3.2, which is the third assumption in (8.102). Hence the third term in (8.128) tends to zero by dominated convergence as $k \to \infty$.

The fourth and fifth term in (8.128) tend to zero by Lemma 8.3.6 and dominated convergence. The sixth term in (8.128) tends to zero because the integral is bounded by t_1 and there is a factor $\frac{1}{N_k}$ in front. To see that the last term in the right-hand side in (8.128) tends to zero, recall that the subsequence N_k is chosen such that

$$
\lim_{k\to\infty} \mathcal{L}\left[(\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N t_1))_{t_1>0}\right] \tag{8.130}
$$

exists. Note that

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sqrt{(\mathcal{F}g)(x_1^{\mathrm{aux}}(s))} - \sqrt{\frac{1}{N}\sum_{i\in[N]} g(x_i^{[N_k^2]}(N_k^2 t_2 + N_k s))}\right)^2\right] \\
&\le \mathbb{E}\left[\left|(\mathcal{F}g)(x_1^{\mathrm{aux}}(s)) - \frac{1}{N}\sum_{i\in[N]} g(x_i^{[N_k^2]}(N_k^2 t_2 + N_k s))\right|\right],
\end{aligned}
\tag{8.131}
$$

and hence we can apply a similar reasoning as in (6.198) to see that (8.131) tends to zero as $k \to \infty$. Therefore we obtain

$$
\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}\left[(1+K_0)(\Delta_i^{(1),[N_k^2]}(N_k t_1))^2 + K_1(\delta_{y_1,i}^{(1),[N_k^2]}(N_k t_1))^2 + K_2(\delta_{y_2,i}^{(1),[N_k^2]}(N_k t_1))^2\right] \\
&\quad = 0.
\end{aligned}
\tag{8.132}
$$

CHAPTER 8

To prove (8.125), note that (8.132) implies convergence of the finite-dimensional distributions of $(\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{t_1>0}$ by a similar argument as given below (6.137). By Lemma 8.3.3 we see that the laws of the processes

$$\left(\mathcal{L}\left[(\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{t_1>0}\right]\right)_{k\in\mathbb{N}_0} \tag{8.133}$$

are tight. Therefore (8.125) indeed holds. □

Remark 8.3.16. Note that in the proof of Lemma 7.2.12 we could have proceeded as in the proof of Lemma 8.3.15, instead of using the criterion in [49, Theorem 3.3.1]. ■

Lemma 8.3.17 (Proof of (8.103)**).** *Under the assumptions in Proposition 8.3.10,*

$$\lim_{k\to\infty} \mathcal{L}\left[\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2)\right] = \int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v). \tag{8.134}$$

Proof. For ease of notation, we drop the subsequence notation in this proof. Let $(t_n)_{n\in\mathbb{N}_0}$ be any sequence satisfying $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} t_n/n = 0$. For each t_n, let $\mu^{(1)}_{N,t_n}$ be the law obtained by periodic continuation of the configuration of $(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N^2]}(N^2 t_2 - N t_n))_{i\in[N]}$. Recall that, since our state space is compact, the sequence $(\mu^{(1)}_{N,t_n})_{N\in\mathbb{N}}$ is tight. Let $\mu^{(1)}_{t_n}$ be any weak limit point of the sequence $(\mu^{(1)}_{N,t_n})_{N\in\mathbb{N}}$.

Let $\mathcal{L}[\mathbf{z}_1^{\mathrm{aux},n}(0)]$ be the law obtained by periodic continuation of $\boldsymbol{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2 t_2 - N t_n)$. By Lemma 8.3.3 we know that the sequence

$$\left(\mathcal{L}\left[\left(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N^2]}(N^2 t_2 - N t_n + N t_1)\right)_{t_1>0,i\in[N]}\right]\right)_{N\in\mathbb{N}} \tag{8.135}$$

is tight and hence for each t_n we can pass to a subsequence such that

$$\lim_{k\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_1)\right)_{t_1>0,i\in[N]}\right] \tag{8.136}$$

exists. By Lemmas 8.3.2–8.3.14, we obtain for all t_n that

$$\lim_{N\to\infty} \bar{\Theta}^{(2),[N_k^2]}(N_k^2 t_2 - N t_n) = \bar{\Theta}^{(2)}(t_2) \text{ in probability.} \tag{8.137}$$

Then, by (8.128) in the proof of Lemma 8.3.15, for fixed t_n and all $i \in [N]$,

$$\begin{aligned}\lim_{k\to\infty} \mathbb{E}\Bigg[&(1+K_0)\left(x_{1,i}^{\mathrm{aux},n}(t_n) - \bar{\Theta}_i^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_n)\right)^2 \\ &+ K_1\left(y_{1,1,i}^{\mathrm{aux},n}(t_n) - \Theta_{y_1,i}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_n)\right)^2 \\ &+ K_2\left(y_{2,1,i}^{\mathrm{aux},n}(t_n) - \Theta_{y_2,i}^{\mathrm{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_n)\right)^2\Bigg] = 0.\end{aligned} \tag{8.138}$$

By contradiction we can argue that

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\Bigg[&(1+K_0)\left(x^{\mathrm{aux},n}_{1,i}(t_n)-\bar{\Theta}^{\mathrm{aux},(1),[N^2]}_{i}(N^2t_2-Nt_n+Nt_n)\right)^2\\&+K_1\left(y^{\mathrm{aux},n}_{1,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N^2]}_{y_1,i}(N^2t_2-Nt_n+Nt_n)\right)^2\\&+K_2\left(y^{\mathrm{aux},n}_{2,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N^2]}_{y_2,i}(N^2t_2-Nt_n+Nt_n)\right)^2\Bigg]=0.\end{aligned}\tag{8.139}$$

To see why, suppose that (8.139) does not hold. Then for any $\delta>0$ we can construct a sequence $(N_l)_{l>0}$ such that, for $l\in\mathbb{N}$,

$$\begin{aligned}\mathbb{E}\Bigg[&(1+K_0)\left(x^{\mathrm{aux},n}_{1,i}(t_n)-\bar{\Theta}^{\mathrm{aux},(1),[N_l^2]}_{i}(N_l^2t_2-N_lt_n+N_lt_n)\right)^2\\&+K_1\left(y^{\mathrm{aux},n}_{1,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N_l^2]}_{y_1,i}(N_l^2t_2-N_lt_n+N_lt_n)\right)^2\\&+K_2\left(y^{\mathrm{aux},n}_{2,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N_l^2]}_{y_2,i}(N_l^2t_2-N_lt_n+N_lt_n)\right)^2\Bigg]>\delta.\end{aligned}\tag{8.140}$$

However, also the sequence

$$\left(\mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{aux},(1),[N_l^2]}_{i}(N_l^2t_2-N_lt_n+N_lt_1)\right)_{t_1>0,i\in[N]}\right]\right)_{l\in\mathbb{N}}\tag{8.141}$$

is tight. Hence we can pass to a further subsequence $(N_{\tilde{l}})_{\tilde{l}\in\mathbb{N}}$ for which (8.138) holds. But this contradicts (8.140). We conclude that (8.139) indeed holds. Moreover the argument holds for all t_n, so that (8.139) holds for all t_n.

Hence for every t_n there exists a N_n such that, for all $N\geq N_n$,

$$\begin{aligned}\mathbb{E}\Bigg[&(1+K_0)\left(x^{\mathrm{aux},n}_{1,i}(t_n)-\bar{\Theta}^{\mathrm{aux},(1),[N_k^2]}_{i}(N^2t_2)\right)^2\\&+K_1\left(y^{\mathrm{aux},n}_{1,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N_k^2]}_{y_1,i}(N^2t_2)\right)^2\\&+K_2\left(y^{\mathrm{aux},n}_{2,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N_k^2]}_{y_2,i}(N^2t_2)\right)^2\Bigg]<\frac{1}{n}.\end{aligned}\tag{8.142}$$

In particular, we may require that $N_n>N_{n-1}$. Setting $N=N_n$, we obtain

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\Bigg[&(1+K_0)\left(x^{\mathrm{aux},n}_{1,i}(t_n)-\bar{\Theta}^{\mathrm{aux},(1),[N_n^2]}_{i}(N_n^2t_2)\right)^2\\&+K_1\left(y^{\mathrm{aux},n}_{1,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N_n^2]}_{y_1,i}(N_n^2t_2)\right)^2\\&+K_2\left(y^{\mathrm{aux},n}_{2,1,i}(t_n)-\Theta^{\mathrm{aux},(1),[N_n^2]}_{y_2,i}(N_n^2t_2)\right)^2\Bigg]=0.\end{aligned}\tag{8.143}$$

If we can prove that

$$\lim_{n\to\infty} \mathcal{L}[z_i^{\mathrm{aux},n}(t_n)] = \Gamma^{(1)}_{\bar\Theta(t_2)}, \tag{8.144}$$

then we are done. To see why, note that, for all $f \in \mathcal{C}_b([0,1]\times[0,1]^2)$, f Lipschitz continuous

$$\begin{aligned}&\left|\mathbb{E}[f(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N^2]}(N^2t_2))] - \mathbb{E}^{\Gamma^{(1)}_{\bar\Theta(t_2)}}[f]\right|\\ &\leq \left|\mathbb{E}[f(\boldsymbol{\Theta}_i^{\mathrm{aux},(1),[N^2]}(N^2t_2)) - f({z_{1,i}}^{\mathrm{aux},n}(t_n))]\right| + \left|\mathbb{E}[f({z_{1,i}}^{\mathrm{aux},n}(t_n))] - \mathbb{E}^{\Gamma^{(1)}_{\bar\Theta(t_2)}}[f]\right|.\end{aligned} \tag{8.145}$$

Therefore if (8.144) holds, then for all $\epsilon > 0$ we can choose $\bar n$ such that, for all $n > \bar n$,

$$\left|\mathbb{E}[f({z_{1,i}}^{\mathrm{aux},n}(t_n))] - \mathbb{E}^{\Gamma^{(1)}_{\bar\Theta(t_2)}}[f]\right| < \frac{\epsilon}{2}. \tag{8.146}$$

By (8.143) we can find a $\hat n > \bar n$ such that for all $n > \hat n$

$$\left|\mathbb{E}[f(\boldsymbol{\Theta}_i^{\mathrm{aux}}(N_n^2t_2)) - f({z_{1,i}}^{\mathrm{aux},n}(t_{\hat n}))]\right| < \frac{\epsilon}{2}. \tag{8.147}$$

Using (8.144) and the fact that the Lipschitz functions are dense in $\mathcal{C}_b([0,1]\times[0,1]^2)$, we obtain (8.134).

Proof of (8.144)**.** We use that any two systems $(\mathbf{z}^{\mathrm{aux},1}(t_1))_{t_1>0}$ and $(\mathbf{z}^{\mathrm{aux},2}(t_1))_{t_1>0}$ evolving according to (8.29), and having the same $y_{2,1}$-components and the same $\bar\Theta(t_2)$, can be constructed on one probability space and can coupled by their Brownian motions. We obtain, for a component $i \in \mathbb{N}_0$,

$$\begin{aligned}&\mathbb{E}[|x_{1,i}^{\mathrm{aux},1}(t_n) - x_{1,i}^{\mathrm{aux},2}(t_n)| + K_1|y_{1,1,i}^{\mathrm{aux},1}(t_n) - y_{1,1,i}^{\mathrm{aux},2}(t_n)| + K_2|y_{2,1,i}^{\mathrm{aux},1}(t_n) - y_{2,1,i}^{\mathrm{aux},2}(t_n)|]\\ &= \mathbb{E}[|x_{1,i}^{\mathrm{aux},1}(0) - x_{1,i}^{\mathrm{aux},2}(0)| + K_1|y_{1,1,i}^{\mathrm{aux},1}(0) - y_{1,1,i}^{\mathrm{aux},2}(0)| + K_2|y_{2,1,i}^{\mathrm{aux},1}(0) - y_{2,1,i}^{\mathrm{aux},2}(0)|]\\ &- c\int_0^{t_n} \mathbb{E}[|x_{1,i}^{\mathrm{aux},1}(s) - x_{1,i}^{\mathrm{aux},2}(s)|]\mathrm{d}s\\ &- 2K_1e_1\int_0^{t_n} \mathbb{E}\big[[|x_{1,i}^{\mathrm{aux},1}(s) - x_{1,i}^{\mathrm{aux},2}(s)| + K_1|y_{1,1,i}^{\mathrm{aux},1}(s) - y_{1,1,i}^{\mathrm{aux},2}(s)|]\\ &\qquad\qquad \times 1_{\{\mathrm{sgn}\,(x_{1,i}^{\mathrm{aux},1}(s)-x_{1,i}^{\mathrm{aux},2}(s))\neq \mathrm{sgn}\,(y_{1,1,i}^{\mathrm{aux},1}(s)-y_{1,1,i}^{\mathrm{aux},2}(s))\}}\big]\mathrm{d}s.\end{aligned} \tag{8.148}$$

Therefore the difference between these two systems monotonically decreases.

Since the state space $[0,1]\times[0,1]^2$ is compact, the sequence of laws

$$(\mathcal{L}[z_i^{\mathrm{aux},n}(0)])_{n\in\mathbb{N}} \tag{8.149}$$

is tight. Therefore we can find converging subsequences such that

$$\lim_{k\to\infty} \mathcal{L}[{z_i}^{\mathrm{aux},n_k}(0)] = \mu \tag{8.150}$$

for some probability measure μ on $[0,1]\times[0,1]^2$.

Let $(z^{\mathrm{aux},0}(t_1))_{t_1>0}$ be the limiting system evolving according to (8.29) and starting from initial distribution μ. By Skorohod's theorem, we can construct the sequence of limiting systems $((z^{\mathrm{aux},n_k}(t_1))_{t_1>0})_{k\in\mathbb{N}}$ and $(z^{\mathrm{aux},0}(t_1))_{t_1>0}$ on one probability space such that

$$\lim_{k\to\infty} z^{\mathrm{aux},n_k}(0) = z^{\mathrm{aux},0}(0) \quad a.s. \tag{8.151}$$

Use the coupling of Brownian motions to obtain

$$\begin{aligned}
&\mathbb{E}[|x^{\mathrm{aux},n_k}_{1,i}(t_{n_k}) - x^{\mathrm{aux},0}_{1,i}(t_{n_k})| + K_1|y^{\mathrm{aux},n_k}_{1,1,i}(t_{n_k}) - y^{\mathrm{aux},0}_{1,1,i}(t_{n_k})| \\
&\qquad + K_2|y^{\mathrm{aux},n_k}_{2,1,i}(t_{n_k}) - y^{\mathrm{aux},0}_{2,1,i}(t_{n_k})|] \\
&= \mathbb{E}[|x^{\mathrm{aux},n_k}_{1,i}(0) - x^{\mathrm{aux},0}_{1,i}(0)| + K_1|y^{\mathrm{aux},n_k}_{1,1,i}(0) - y^{\mathrm{aux},0}_{1,1,i}(0)| + K_2|y^{\mathrm{aux},n_k}_{2,1,i}(0) - y^{\mathrm{aux},0}_{2,1,i}(0)|] \\
&- c\int_0^{t_{n_k}} \mathbb{E}[|x^{\mathrm{aux},n_k}_{1,i}(s) - x^{\mathrm{aux},0}_{1,i}(s)|]\mathrm{d}s \\
&- 2K_1e_1\int_0^{t_{n_k}} \mathbb{E}\big[[|x^{\mathrm{aux},n_k}_{1,i}(s) - x^{\mathrm{aux},0}_{1,i}(s)| + K_1|y^{\mathrm{aux},n_k}_{1,1,i}(s) - y^{\mathrm{aux},0}_{1,1,i}(s)|] \\
&\qquad\qquad \times 1_{\{\mathrm{sgn}\,(x^{\mathrm{aux},n_k}_{1,i}(s) - x^{\mathrm{aux},0}_{1,i}(s)) \neq \mathrm{sgn}\,(y^{\mathrm{aux},n_k}_{1,1,i}(s) - y^{\mathrm{aux},0}_{1,1,i}(s))\}}\big]\mathrm{d}s.
\end{aligned} \tag{8.152}$$

Taking the limit $k\to\infty$ on both sides of (8.152), we obtain

$$\begin{aligned}
\lim_{k\to\infty} &\mathbb{E}[|x^{\mathrm{aux},n_k}_{1,i}(t_{n_k}) - x^{\mathrm{aux},0}_{1,i}(t_{n_k})| + K_1|y^{\mathrm{aux},n_k}_{1,1,i}(t_{n_k}) - y^{\mathrm{aux},0}_{1,1,i}(t_{n_k})| \\
&+ K_2|y^{\mathrm{aux},n_k}_{2,1,i}(t_{n_k}) - y^{\mathrm{aux},0}_{2,1,i}(t_{n_k})|] = 0.
\end{aligned} \tag{8.153}$$

Note that $\lim_{n\to\infty} t_n = \infty$ implies that $\lim_{k\to\infty} t_{n_k} = \infty$, so $z_i^{\mathrm{aux},0}$ is the limiting system in (8.29) with θ replaced by the random variable $\bar\Theta(t_2)$ and $y_{2,1,i}$. Therefore we can condition on $\bar\Theta(t_2)$ and $y_{2,1,i}$, and use the assumption in (8.81), to obtain

$$\lim_{k\to\infty} \mathcal{L}\left[z_i^{\mathrm{aux},0}(t_{n_k})\right] = \int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v). \tag{8.154}$$

Hence we conclude that

$$\lim_{k\to\infty} \mathcal{L}[z_i^{\mathrm{aux},n_k}(t_{n_k})] = \int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v). \tag{8.155}$$

Equation (8.155) holds for all subsequences along which the initial distribution converges,

$$\lim_{k\to\infty} z_i^{\mathrm{aux},n_k}(0) = z_i^{\mathrm{aux},0}(0) \quad a.s. \tag{8.156}$$

We will show that this implies (8.144).

Suppose that

$$\lim_{n\to\infty} \mathcal{L}[z_i^{\mathrm{aux},n}(t_n)] \neq \int_{[0,1]^2}\int_{[0,1]} \Gamma^{\mathrm{aux},(1)}_{u,y_{2,1}}\, P^{(u,v)}(\mathrm{d}y_{2,1})\, P_{t_2}(\mathrm{d}u,\mathrm{d}v). \tag{8.157}$$

Then there exist $f\in\mathcal{C}_b([0,1]\times[0,1]^3)$ and $\delta>0$ such that for all $N\in\mathbb{N}$ there exists an $n\in\mathbb{N}$, $n>N$ such that

$$\left|\mathbb{E}[f(z_i^{\mathrm{aux},n}(t_n))] - \mathbb{E}^{\Gamma^{(1)}_{\bar\Theta(t_2)}}[f]\right| > \delta. \tag{8.158}$$

Hence we can construct a subsequence $(z_i^{\mathrm{aux},n_k}(t_1))_{t_1>0,\,k\in\mathbb{N}}$ such that (8.158) holds for each $k\in\mathbb{N}$. However, also for this sequence $(\mathcal{L}[z_i^{\mathrm{aux},n_k}(0)])_{k\in\mathbb{N}}$ is tight. Passing to a possibly further subsequence of converging initial distributions, we argue like before to obtain that along this subsequence

$$\lim_{k\to\infty}\left|\mathbb{E}[f(z_i^{\mathrm{aux},n_k}(t_{n_k}))]-\mathbb{E}^{\Gamma^{(1)}_{\Theta(t_2)}}[f]\right|=0. \tag{8.159}$$

This contradicts (8.158) and so (8.144) is indeed true. □

Proof of Proposition 8.3.10

Proof. Lemma 8.3.17 implies (8.103). Therefore Lemma 8.3.15 implies (8.104). □

§8.3.7 Convergence of 2-block process

In this section we derive the limiting evolution of the effective 2-block process.

Lemma 8.3.18 (Convergence of the 2-block averages). *Assume that* $(N_k)_{k\in\mathbb{N}}\subset\mathbb{N}$ *is a subsequence satisfying*

$$\begin{aligned}&\lim_{k\to\infty}\mathcal{L}\left[y_{2,1}^{[N_k^2]}(N_kt_2)\middle|\boldsymbol{\Theta}^{(2),[N_k^2]}(N_k^2t_2)\right]=P^{z_2(t_2)},\\&\lim_{k\to\infty}\mathcal{L}\left[\left(Y_{1,0}^{[N_k^2]}(N_k^2t_2+N_kt_1),Y_{2,0}^{[N_k^2]}(N_k^2t_2)\right)\middle|\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2t_2+N_kt_1)\right]\\&\qquad=P^{z_1^{\mathrm{eff}}(t_1)}.\end{aligned} \tag{8.160}$$

Then, for the effective 2*-block estimator process defined in* (8.23)*,*

$$\lim_{k\to\infty}\mathcal{L}\left[\left(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2t_2)\right)_{t_2>0}\right]=\mathcal{L}\left[\left(z_2^{\mathrm{eff}}(t_2)\right)_{t_2>0}\right], \tag{8.161}$$

where the limit is determined by the unique solution of the SSDE (8.36) *with initial state*

$$z_2^{\mathrm{eff}}(0)=\left(x_2^{\mathrm{eff}}(0),y_2^{\mathrm{eff}}(0)\right)=(\vartheta_1,\theta_{y_2}). \tag{8.162}$$

Proof. Again we use [49, Theorem 3.3.1]. By a similar argument as used in the proof of Lemma 7.2.12 we can show that

$$\lim_{t_2\downarrow 0}\mathcal{L}\left[\boldsymbol{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2t_2)\right]=\delta_{(\vartheta_1,\theta_{y_2})}. \tag{8.163}$$

Note that by steps 1-4 of the scheme in Section 8.2 we can choose the subsequence $(N_k)_{k\in\mathbb{N}}$ such that both (8.81) and (8.102) hold. Since we already established the tightness of the 2-block in Lemma 8.3.1, we are left to show that, for all $t_2>0$,

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\Big[\Big|&G_{\dagger}^{(2),[N_k^2]}(f,\boldsymbol{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2t_2),t_2,\omega)\\&-G^{(2)}f\left(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N_k^2]}(N_k^2t_2)\right)\Big|\Big]=0,\end{aligned} \tag{8.164}$$

where $G^{(2),[N_k^2]}_{\dagger}$ is the $\mathcal{D}$-semi-martingale operator defined in (8.69), $G^{(2)}$ is the generator of the process $(z_2^{\mathrm{eff}}(t_2))_{t_2>0}$ defined in (8.36), and both generators work on a probability space driven by one set of Brownian motions. Note that, for all $t_2>0$,

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg|G^{(2),[N_k^2]}_{\dagger}\left(f,\left(\bar\Theta^{(2),[N_k^2]}(N_k^2t_2),\Theta^{(2),[N_k^2]}_{y_2}(N_k^2t_2)\right),t_2,\omega\right)\\
&\qquad\qquad - G^{(2)}f\left(\bar\Theta^{(2),[N_k^2]}(N_k^2t_2),\Theta^{(2),[N_k^2]}_{y_2}(N_k^2t_2)\right)\Bigg|\Bigg]\\
&\le \frac{K_2e_2}{1+K_0+K_1}\mathbb{E}\left[\left|\bar\Theta^{(2),[N_k^2]}(N_k^2t_2)-\Theta^{(2),[N_k^2]}_{x}(N_k^2t_2,\omega)\right|\left|\frac{\partial f}{\partial x}\right|\right]\\
&\quad + e_2\mathbb{E}\left[\left|\Theta^{(2),[N_k^2]}_{x}(N_k^2t_2,\omega)-\bar\Theta^{(2),[N_k^2]}(N_k^2t_2)\right|\left|\frac{\partial f}{\partial y}\right|\right]\\
&\quad + \frac{1}{(1+K_0+K_1)^2}\\
&\qquad \times\mathbb{E}\left[\left|\frac{1}{N_k^2}\sum_{i\in[N_k^2]} g(x_i^{[N_k^2]}(N_k^2t_2,\omega))-(\mathcal{F}^{(2)}g)(\bar\Theta^{(2),[N_k^2]}(N_k^2t_2))\right|\left|\frac{\partial^2 f}{\partial x^2}\right|\right].
\end{aligned}\tag{8.165}$$

The first and second term on the right-hand side tend to 0 as $k\to\infty$ by a similar argument as used in (8.129) and below. For the third let $[N]_i$ denote the 1-block that contains site i and let $(z^{\nu_{\bar\Theta_i^{(1)}}}(t))_{t>0}$ be the limiting single colony system, with drift towards the random variable $\bar\Theta_i$ and starting from the equilibrium measure $\nu_{\bar\Theta_i^{(1)}}$. We construct the single colony system $Z^{[N_k^2]}(N_k^2t_2-L(N)+t)_{t\ge0}$ and the limiting system $(z^{\nu_{\bar\Theta_i^{(1)}}}(t))_{t>0}$ on one probability space, such that by Skohorod's theorem, we can assume that the convergence is almost surely. Note that $\bar\Theta_i$ is the limiting one block. Then we can write

$$\begin{aligned}
&\mathbb{E}\left[\left|\frac{1}{N_k^2}\sum_{i\in[N_k^2]} g(x_i^{[N_k^2]}(N_k^2t_2,\omega))-(\mathcal{F}^{(2)}g)(\bar\Theta^{(2),[N_k^2]}(N_k^2t_2))\right|\left|\frac{\partial^2 f}{\partial x^2}\right|\right]\\
&\quad\le \frac{1}{N_k}\sum_{i\in[N]}\mathbb{E}\left[\left|\frac{1}{N_k}\sum_{j\in[N]_i} g(x_j^{[N_k^2]}(N_k^2t_2,\omega))-\frac{1}{N_k}\sum_{j\in[N]_i} g(x_j^{\nu_{\bar\Theta_i^{(1)}}}(L(N))\right|\left|\frac{\partial^2 f}{\partial x^2}\right|\right]\\
&\quad + \frac{1}{N_k}\sum_{i\in[N]}\mathbb{E}\left[\left|\frac{1}{N_k}\sum_{j\in[N]_i} g(x_j^{\nu_{\bar\Theta_i^{(1)}}}(L(N))-(\mathcal{F}^{(1)}g)(\bar\Theta_i^{(1)})\right|\left|\frac{\partial^2 f}{\partial x^2}\right|\right]\\
&\quad + \mathbb{E}\left[\left|\frac{1}{N_k}\sum_{i\in[N]}(\mathcal{F}^{(1)}g)(\bar\Theta_i^{(1)})-(\mathcal{F}^{(2)}g)(\bar\Theta_i^{(2)})\right|\left|\frac{\partial^2 f}{\partial x^2}\right|\right]\\
&\quad + \mathbb{E}\left[\left|(\mathcal{F}^{(2)}g)(\bar\Theta_i^{(2)})-(\mathcal{F}^{(2)}g)(\bar\Theta_i^{(2)}(N_k^2t_2))\right|\left|\frac{\partial^2 f}{\partial x^2}\right|\right].
\end{aligned}\tag{8.166}$$

The first term on the right-hand side tends to zero by Lipschitz continuity for g and Corollary 8.3.9. The second term tends to zero by the law of large numbers, since the limiting single colonies are i.i.d. given the value of the random variable $\bar\Theta_i^{(1)}$.

CHAPTER 8

The third term tends to zero since by Proposition 8.3.10 also the limiting 1-blocks become independent given the value of the 2-block. Hence we can again apply the law of large numbers. Finally, for the last term, note that since we construct the single components and the limiting process on one probability space, we can argue like in the proof of Lemma 7.2.8 that

$$\lim_{k\to\infty} \mathbb{E}\left[\left|\bar{\Theta}_i^{(2)} - \bar{\Theta}_i^{(2)}(N_k^2 t_2)\right|\right] = 0. \tag{8.167}$$

Hence the last term tends to zero by the Lipschitz property of $\mathcal{F}^{(2)}g$. □

Remark 8.3.19. Instead of [49, Theorem 3.3.1] we could have used a similar strategy as in the proof of Lemma 8.3.15 to obtain Lemma 8.3.18. ■

§8.3.8 State of the slow seed-banks

On time scale t_0, i.e., space-time scale 0, the colour-1 seed-bank is a "slow seed-bank," since it does not move on this time scale. Because we study the two-layer three-colour mean-field system from time $N^2 t_2$ onwards, the 1-block averages of the colour 1-dormant population are already in equilibrium. As a consequence we can exactly describe the single 1-dormant colonies, which turn out to be in a state that equals the current 1-block average of the dormant population of colour 1. To obtain the formal result we will first prove the following lemma.

Lemma 8.3.20 (Slow seed-banks). *Fix $t_2, t_1 > 0$, for $i \in [N^2]$ and all $t_0 \geq 0$,*

$$\lim_{N\to\infty} \left[y_{i,1}^{[N^2]}(N^2 t_2 + N t_1 + t_0) - \Theta_{y_1,i}^{(1),[N^2]}(N^2 t_2 + N t_1 + t_0)\right] = 0 \qquad a.s., \tag{8.168}$$

where $\Theta_{y_1,i}^{(1),[N^2]}$ is the 1-block average to which $y_{i,1}^{[N^2]}$ contributes.

To prove Lemma 8.3.20, we need the kernel $b^{[N^2]}(\cdot,\cdot)$ defined in 4.31, which becomes in the current setting

$$b^{[N^2]}((i,R_i),(j,R_j)) = \begin{cases} 1_{\{d_{[N^2]}(i,j)\leq 1\}} \frac{c_0}{N} + \frac{c_1}{N^3}, & \text{if } R_i = R_j = A,\\ K_m \frac{e_m}{N^m}, & \text{if } i=j,\ R_i = A,\ R_j = D_m,\ m \in \{0,1,2\},\\ \frac{e_m}{N^m}, & \text{if } i=j,\ R_i = D_m,\ R_j = A,\ m \in \{0,1,2\},\\ 0, & \text{otherwise.} \end{cases} \tag{8.169}$$

The corresponding semigroup of the kernel $b^{[N^2]}(\cdot,\cdot)$ is denoted by $b_t^{[N^2]}(\cdot,\cdot)$.

To prove Lemma 8.3.20 we will use the following lemma, which was proved in [43][Lemma 6.1] and for our setting reads as follows.

Lemma 8.3.21 (First and second moment).
Let $\mathbb{E}_{z^{[N^2]}}$ the expectation if the process start from some state $z^{[N^2]} \in ([0,1]\times[0,1]^3)^{[N^2]}$. For $z^{[N^2]} \in ([0,1]\times[0,1]^2)^{[N^2]}$, $t \geq 0$ and $(i,R_i),(j,R_j) \in [N^2]\times\{A,D_0,D_1,D_2\}$,

$$\mathbb{E}_{z^{[N^2]}}[z_{(i,R_i)}^{[N^2]}(t)] = \sum_{\substack{(k,R_k)\in\\ \Omega_N\times\{A,D_0,D_1,D_2\}}} b_t^{[N^2]}\big((i,R_i),(k,R_k)\big)\, z_{(k,R_k)}^{[N^2]} \tag{8.170}$$

and

$$\begin{aligned}
&\mathbb{E}_{z^{[N^2]}}[z^{[N^2]}_{(i,R_i)}(t) z^{[N^2]}_{(j,R_j)}(t)] \\
&= \sum_{\substack{(k,R_k),(l,R_l)\in \\ \Omega_N\times\{A,D_0,D_1,D_2\}}} b^{[N^2]}_t\big((i,R_i),(k,R_k)\big)\, b^{[N^2]}_t\big((j,R_j),(l,R_l)\big)\, z^{[N^2]}_{(k,R_k)} z^{[N^2]}_{(l,R_l)} \\
&\quad + 2\int_0^t \mathrm{d}s \sum_{k\in\Omega_N} b^{[N^2]}_{(t-s)}((i,R_i),(k,A))\, b^{[N^2]}_{(t-s)}((j,R_j),(k,A))\, \mathbb{E}^{[N^2]}_z[g(x^{[N^2]}_k(s))].
\end{aligned} \tag{8.171}$$

Proof of Lemma 8.3.20. The argument is given in such a way that it can easily be generalised to more complicated systems, which we treat later. Let $\bar{t}(N) = N^2 t_2 + N t_1 + t_0$. We will show that if $i,j \in [N]_i$, i.e., i and j belong to the same 1-block, then

$$\lim_{N\to\infty} \mathbb{E}\left[\left(y^{[N^2]}_{i,1}(\bar{t}(N)) - y^{[N^2]}_{j,1}(\bar{t}(N))\right)^2\right] = 0. \tag{8.172}$$

This implies (8.168). By Lemma 8.3.21, we can write

$$\begin{aligned}
&\mathbb{E}\left[\left(y^{[N^2]}_{i,1}(\bar{t}(N)) - y^{[N^2]}_{j,1}(\bar{t}(N))\right)^2\right] \\
&= \sum_{\substack{(k,R_k),(l,R_l)\in \\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left(b^{[N^2]}_{\bar{t}(N)}\left((i,D_1),(k,R_k)\right) - b^{[N^2]}_{\bar{t}(N)}\left((j,D_1),(k,R_k)\right)\right) \\
&\quad \times \left(b^{[N^2]}_{\bar{t}(N)}\left((i,D_1),(l,R_l)\right) - b^{[N^2]}_{\bar{t}(N)}\left((j,D_1),(l,R_l)\right)\right) \mathbb{E}[z^{[N^2]}_{(k,R_k)} z^{[N^2]}_{(l,R_l)}] \\
&\quad + 2\int_0^{\bar{t}(N)} \mathrm{d}s \sum_{k\in[N^2]} \left(b^{[N^2]}_{(\bar{t}(N)-s)}((i,D_1),(k,A)) - b^{[N^2]}_{(\bar{t}(N)-s)}((j,D_1),(k,A))\right)^2 \\
&\quad \times \mathbb{E}[g(x^{[N^2]}_k(s))].
\end{aligned} \tag{8.173}$$

Using a coupling argument, we show that both terms in (8.173) tend to 0 as $N\to\infty$. To prove that the first term tends to 0, we will show that

$$\lim_{N\to\infty} \sum_{\substack{(k,R_k)\in \\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left| b^{[N^2]}_{\bar{t}(N)}\left((i,D_1),(k,R_k)\right) - b^{[N^2]}_{\bar{t}(N)}\left((j,D_1),(k,R_k)\right)\right| = 0. \tag{8.174}$$

To do so, let $(RW^{[N^2]}(t))_{t\geq 0}$ and $(RW'^{[N^2]}(t))_{t\geq 0}$ be two independent random walks, starting from $RW^{[N^2]}(0) = (i,D_1)$ and $RW'^{[N^2]}(0) = (j,D_1)$, where i and j are in the same 1-block. Let $RW^{[N^2]}$ and $RW'^{[N^2]}$ both evolve according to the kernel $b^{[N^2]}(\cdot,\cdot)$, so $b^{[N^2]}_t(\cdot,\cdot)$ is their corresponding semigroup. Since $RW^{[N^2]}$ and $RW'^{[N^2]}$ both start from the colour 1-seed-bank, we can perfectly couple their switches between A, D_0, D_1 and D_2. Since this implies that both $RW^{[N^2]}$ and $RW'^{[N^2]}$ are always simultaneously active, we can also couple the times when they jump due to migration and the distance over which they migrate. However, we do not couple their migrations, i.e. $RW^{[N^2]}$ and $RW'^{[N^2]}$ jump at the same time and over the same distance, but they can jump to different sites. This implies that the coupled process $(RW^{[N^2]}(t), RW'^{[N^2]}(t))_{t\geq 0}$ has transition rates

$$((i,R_i),(j,R_i)) \to \begin{cases} ((k,A),(l,A)) & \text{if } R_i = R_j = A \text{ and } d_{[N^2]}(i,k) = d_{[N^2]}(j,l) \\ & \text{at rate } 1_{\{d_{[N^2]}(i,j)\leq 1\}} \frac{c_0}{N} + \frac{c_1}{N^3}, \\ ((i,D_m),(j,D_m)) & \text{if } R_i = R_j = A \text{ at rate } \frac{K_m e_m}{N^m},\ m \in \{0,1,2\}, \\ ((i,A),(j,A)) & \text{if } R_i = R_j = D_m \text{ at rate } \frac{e_m}{N^m},\ m \in \{0,1,2\}. \end{cases} \tag{8.175}$$

Define the event

$$H_t^{[N^2]} = \{RW^{[N^2]} \text{ has migrated at least once up to time } t\}. \tag{8.176}$$

Note that if $H_t^{[N^2]}$ has happened, then also $RW'^{[N^2]}$ has migrated. Hence

$$\begin{aligned} & b_{\bar{t}(N)}^{[N^2]}((i,D_1),(k,R_k)) - b_{\bar{t}(N)}^{[N^2]}((j,D_1),(k,R_k)) \\ &\quad = \mathbb{P}_{(i,D_1)}\left(RW^{[N^2]}(\bar{t}(N)) = (k,R_k)\right) - \mathbb{P}_{(j,D_1)}\left(RW'^{[N^2]}(\bar{t}(N)) = (k,R_k)\right) \\ &\quad = \tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW^{[N^2]}(\bar{t}(N)) = (k,R_k),\ H_{\bar{t}(N)}^{[N^2]}\right) \\ &\qquad + \tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW^{[N^2]}(\bar{t}(N)) = (k,R_k),\ (H_{\bar{t}(N)}^{[N^2]})^c\right) \\ &\qquad - \tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW'^{[N^2]}(\bar{t}(N)) = (k,R_k),\ H_{\bar{t}(N)}^{[N^2]}\right) \\ &\qquad - \tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW'^{[N^2]}(\bar{t}(N)) = (k,R_k),\ (H_{\bar{t}(N)}^{[N^2]})^c\right) \\ &\quad = \tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW^{[N^2]}(\bar{t}(N)) = (k,R_k),\ (H_{\bar{t}(N)}^{[N^2]})^c\right) \\ &\qquad - \tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW'^{[N^2]}(\bar{t}(N)) = (k,R_k),\ (H_{\bar{t}(N)}^{[N^2]})^c\right), \end{aligned} \tag{8.177}$$

where the last equality follows because, once the random walks have just jumped once, $RW^{[N^2]}$ and $RW'^{[N^2]}$ are uniformly distributed over $[N] \times A$ if their jump horizon was 1 and they are uniformly distributed of $[N^2] \times A$ if they jumped over distance 2. Hence if $H_t^{[N^2]}$ has occurred, then $RW^{[N^2]}$ and $RW'^{[N^2]}$ have the same distribution. Therefore

$$\sum_{\substack{(k,R_k)\in \\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left| b_{\bar{t}(N)}^{[N^2]}((i,D_1),(k,R_k)) - b_{\bar{t}(N)}^{[N^2]}((j,D_1),(k,R_k)) \right| \leq 2\tilde{\mathbb{P}}((H_{\bar{t}(N)}^{[N^2]})^c) \tag{8.178}$$

and we are left to show that

$$\lim_{N\to\infty} \tilde{\mathbb{P}}((H_{\bar{t}(N)}^{[N^2]})^c) = 0. \tag{8.179}$$

The event $(H_{\bar{t}(N)}^{[N^2]})^c$ occurs either when the random walks do not wake up before time $\bar{t}(N)$ or when the random walks wake up before time $\bar{t}(N)$ but do not migrate. By the coupling we only have to consider one of the random walks. Therefore the probability that $RW^{[N^2]}$ and $RW'^{[N^2]}$ do not wake up before time $\bar{t}(N)$ is given by

$$\tilde{\mathbb{P}}_{(i,D_1),(j,D_1)}\left(RW^{[N^2]} \text{ does not wake up before } \bar{t}(N)\right) = \mathrm{e}^{-\frac{e_1}{N}\bar{t}(N)} = \mathrm{e}^{-\frac{e_1(N^2t_2+Nt_1+t_0)}{N}} \tag{8.180}$$

and hence

$$\lim_{N\to\infty} \mathbb{P}_{(i,D_1),(j,D_1)}\left(RW^{[N^2]} \text{ does not wake up before } \bar{t}(N)\right) = 0. \tag{8.181}$$

The probability that the random walks do wake up, but do not migrate is a little more complicated, since each time they wake up with positive probability they go to sleep before they migrate. Define

$$\begin{aligned} C^{[N^2]}(t) &= \{\#\,\text{times } RW^{[N^2]} \text{ gets active before time } t\},\\ T_A^{[N^2]}(t) &= \{\text{total time } RW^{[N^2]} \text{ is active up to time } t\},\\ T_D^{[N^2]}(t) &= \{\text{total time } RW^{[N^2]} \text{ is dormant up to time } t\}. \end{aligned} \tag{8.182}$$

Thus, $C^{[N^2]}(t)$ counts the number of active/dormant cycles. Define $T_{A,n}^{[N^2]}$, $T_{D,n}^{[N^2]}$ as the active respectively, dormant time during the nth cycle. Define

$$\chi = K_0e_0 + \frac{K_1e_1}{N} + \frac{K_2e_2}{N^2}, \tag{8.183}$$

so χ is the total rate at which RW and RW' become dormant when they are active. Define

$$c = c_0 + \frac{c_1}{N}, \tag{8.184}$$

so c is the total rate at which RW and RW' migrate when they are active. Then

$$T_A^{[N^2]}(t) = \sum_{n=1}^{C^{[N^2]}(t)} T_{A,n}^{[N^2]}, \qquad T_D^{[N^2]}(t) = \sum_{n=1}^{C^{[N^2]}(t)} T_{D,n}^{[N^2]}, \tag{8.185}$$

where $T_{A,n}^{[N^2]} \overset{d}{=} \exp(\chi)$ and $T_{D,n}^{[N^2]} \overset{d}{=} \frac{1}{\chi}K_0e_0 \exp(e_0) + \frac{1}{\chi}\frac{K_1e_1}{N}\exp(\frac{e_1}{N}) + \frac{1}{\chi}\frac{K_2e_2}{N^2}\exp(\frac{e_2}{N^2})$. Once awake, $RW^{[N^2]}$ migrates at rate c and hence the probability to migrate before time $\bar{t}(N)$ is given by $1-\mathrm{e}^{-cT_A^{[N^2]}(\bar{t}(N))}$. Therefore we are left to show that

$$\lim_{N\to\infty} cT_A^{[N^2]}(\bar{t}(N)) = \lim_{N\to\infty} c \sum_{n=1}^{C^{[N^2]}(\bar{t}(N))} T_{A,n}^{[N^2]} = \infty, \qquad a.s. \tag{8.186}$$

Since $T_{A,n}^{[N^2]} \overset{d}{=} \exp(\chi)$, it is enough to show that

$$\lim_{N\to\infty} C^{[N^2]}(\bar{t}(N)) = \infty \qquad a.s. \tag{8.187}$$

To do so, we assume the contrary, i.e., there exists an $R \in \mathbb{N}$ such that for all $\bar{N} \in \mathbb{N}$ there exists an $N > \bar{N}$ such that

$$\mathbb{P}_{(i,D_1)}(C^{[N^2]}(\bar{t}(N)) \le R) > 0. \tag{8.188}$$

CHAPTER 8

Let $L(N)$ be such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$. Note that, by (8.181), we can condition on the first wake-up time and estimate

$$\begin{aligned}
\mathbb{P}_{(i,D_1)}(C^{[N^2]}(\bar{t}(N)) \leq R) &= \int_0^{\bar{t}(N)} \mathrm{d}s\, \mathbb{P}_{(i,A)}(C^{[N^2]}(\bar{t}(N)-s) \leq R)\, \frac{e_1}{N} \mathrm{e}^{-\frac{e_1}{N}s} \\
&= \int_0^{\bar{t}(N)-L(N)} \mathrm{d}s\, \mathbb{P}_{(i,A)}(C^{[N^2]}(\bar{t}(N)-s) \leq R)\, \frac{e_1}{N} \mathrm{e}^{-\frac{e_1}{N}s} \\
&\quad + \int_{\bar{t}(N)-L(N)}^{\bar{t}(N)} \mathrm{d}s\, \mathbb{P}_{(i,A)}(C^{[N^2]}(\bar{t}(N)-s) \leq R)\, \frac{e_1}{N} \mathrm{e}^{-\frac{e_1}{N}s} \\
&\leq \mathbb{P}_{(i,A)}(C^{[N^2]}(L(N)) \leq R) + \mathrm{e}^{-\frac{e_1}{N}\bar{t}(N)} \left[\mathrm{e}^{\frac{e_1}{N}L(N)} - 1\right].
\end{aligned} \tag{8.189}$$

Note that the second term in the last inequality tends to 0 as $N \to \infty$. For the first term, note that we are now looking at time $L(N)$, i.e., time scale N^0. Since

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{P}_{i,A}\left(RW^{[N^2]} \text{ jumps to } D_1 \text{ or } D_2 \text{ before time } L(N)\right) \\
&\quad = \lim_{N\to\infty} 1 - \mathrm{e}^{-(\frac{K_1 e_1}{N} + \frac{K_2 e_2}{N^2})L(N)} = 0.
\end{aligned} \tag{8.190}$$

we have $\lim_{N\to\infty} \mathbb{P}_{i,A}\left(\{RW^{[N^2]}(s) \in \{A, D_0\} \text{ for } s \in [0, L(N)]\}\right) = 1$. Hence, conditioned on the event $\{RW^{[N^2]} \in \{A, D_0\}\}$, $T^{[N^2]}_{A,n} \stackrel{d}{=} \exp(K_0 e_0)$ and $T^{[N^2]}_{D,n} \stackrel{d}{=} \exp(e_0)$. We therefore obtain

$$\begin{aligned}
\mathbb{P}_{(i,A)}(C^{[N^2]}(L(N)) \leq R) &= \mathbb{P}_{(i,A)}\left(\sum_{n=1}^{R} (T^{[N^2]}_{A,n} + T^{[N^2]}_{D,n}) \geq L(N)\right) \\
&\leq \frac{R}{L(N)} \mathbb{E}_{(i,A)}\left[T^{[N^2]}_{A,n} + T^{[N^2]}_{D,n}\right] \\
&= \frac{R}{L(N)} \left[\frac{1}{K_0 e_0} + \frac{1}{e_0}\right].
\end{aligned} \tag{8.191}$$

Taking the limit $N \to \infty$ in (8.191) and combining this with (8.189), we conclude that (8.187) indeed holds. Hence also (8.179) and (8.174) hold.

We are left to show that

$$\begin{aligned}
&\lim_{N\to\infty} 2 \int_0^{\bar{t}(N)} \mathrm{d}s \sum_{k\in[N^2]} \left(b^{[N^2]}_{(\bar{t}(N)-s)}((i,D_1),(k,A)) - b^{[N^2]}_{(\bar{t}(N)-s)}((j,D_1),(k,A))\right)^2 \\
&\qquad \times \mathbb{E}[g(x_k(s))] = 0.
\end{aligned} \tag{8.192}$$

Also here the idea is to make a similar coupling. As soon as the random walks migrate, they are equally distributed. On time scale N, after waking up from the colour 1 seed-bank they will almost immediately migrate, since migration happens on time scale 1, i.e., by time $L(N)$ they have migrated with probability tending to 1. This will again be the key to show that (8.192) tends to 0 as $N \to \infty$.

Note that, by (8.173),

$$\left| 2\int_0^{\bar{t}(N)} \mathrm{d}s \sum_{k\in[N^2]} \Big(b^{[N^2]}_{(\bar{t}(N)-s)}((i,D_1),(k,A)) - b^{[N^2]}_{(\bar{t}(N)-s)}((j,D_1),(k,A))\Big)^2 \times \mathbb{E}[g(x_k^{[N^2]}(s))] \right| \le 2. \tag{8.193}$$

We will again use the coupling in (8.176). Define

$$\tau^{[N^2]} = \inf\{t \ge 0 : RW^{[N^2]}(t) = (k,A) \text{ for some } k \in [N^2]\}. \tag{8.194}$$

Then, for all $l \in [N^2]$,

$$\mathbb{P}_{(l,D_1)}(\tau^{[N^2]} \le t) = 1 - \mathrm{e}^{-\frac{e_1}{N}t}. \tag{8.195}$$

Setting $s = \bar{t}(N) - s$, we can rewrite the integral in (8.192) as

$$\begin{aligned}
&2\int_0^{\bar{t}(N)} \mathrm{d}s \sum_{k\in[N^2]} \Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,D_1),(j,D_1)}\Big(RW^{[N^2]}(s) = (k,A)\Big) \\
&\quad - \tilde{\mathbb{P}}^{[N^2]}_{(i,D_1),(j,D_1)}\Big(RW'^{[N^2]}(s) = (k,A),\Big)\Big)^2 \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N)-s))] \\
&= 2\int_0^{\bar{t}(N)} \mathrm{d}s \sum_{k\in[N^2]} \Bigg[\int_0^s \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \\
&\quad \Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW^{[N^2]}(s-r) = (k,A)\Big) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW'^{[N^2]}(s-r) = (k,A)\Big)\Big)\Bigg] \\
&\quad \times \Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,D_1),(j,D_1)}\Big(RW^{[N^2]}(s) = (k,A)\Big) - \tilde{\mathbb{P}}^{[N^2]}_{(i,D_1),(j,D_1)}\Big(RW'^{[N^2]}(s) = (k,A)\Big)\Big) \\
&\quad \times \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N)-s))].
\end{aligned} \tag{8.196}$$

In what follows we will abbreviate

$$\begin{aligned}
&P_{\Delta^{[N^2]}_{(i,A),(j,A)}}(t,((l,R_l),(j,R_j))) \\
&= \tilde{\mathbb{P}}_{(i,A),(j,A)}\Big(RW^{[N^2]}(t) = (l,R_l)\Big) - \tilde{\mathbb{P}}_{(i,A),(j,A)}\Big(RW'^{[N^2]}(t) = (j,R_j)\Big),
\end{aligned} \tag{8.197}$$

and similarly, for $m \in \{0,1,2\}$,

$$\begin{aligned}
&P_{\Delta^{[N^2]}_{(i,D_m),(j,D_m)}}(t,((l,R_l),(j,R_j))) \\
&= \tilde{\mathbb{P}}_{(i,D_m),(j,D_m)}\Big(RW^{[N^2]}(t) = (l,R_l)\Big) - \tilde{\mathbb{P}}_{(i,D_m),(j,D_m)}\Big(RW'^{[N^2]}(t) = (j,R_j)\Big).
\end{aligned} \tag{8.198}$$

By (8.193), we can use Fubini to swap the order of integration and subsequently substitute $v = s - r$, to obtain

$$
\begin{aligned}
&2\int_0^{\bar{t}(N)} \mathrm{d}r\,\tilde{\mathbb{P}}(\tau^{[N^2]}=r)\int_0^{\bar{t}(N)-r}\mathrm{d}v \sum_{k\in[N^2]} P_{\Delta^{[N^2]}_{(i,A),(j,A)}}(v,((k,A),(k,A)))\\
&\qquad\times P_{\Delta^{[N^2]}_{(i,D_1),(j,D_1)}}(v+r,((k,A),(k,A)))\mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N)-r-v))]\\
&= 2\int_0^{\bar{t}(N)} \mathrm{d}r\,\tilde{\mathbb{P}}(\tau^{[N^2]}=r)\\
&\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left(\mathbb{P}_{(i,D_1)}\left(RW^{[N^2]}(r)=(l,R_l)\right)-\mathbb{P}_{(j,D_1)}\left(RW'^{[N^2]}(r)=(l,R_l)\right)\right)\\
&\quad\times\int_0^{\bar{t}(N)-r}\mathrm{d}v\sum_{k\in[N^2]} P_{\Delta^{[N^2]}_{(i,A),(j,A)}}(v,((k,A),(k,A)))\\
&\quad\times \mathbb{P}_{(l,R_l)}\left(RW^{[N^2]}(v)=(k,A)\right)\mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N)-r-v))],
\end{aligned}
\tag{8.199}
$$

where in the last equality we use that the random walks move according to the same kernel $b(\cdot,\cdot)$.

We can continue by writing

$$
\begin{aligned}
&2\int_0^{\bar{t}(N)} \mathrm{d}r\,\tilde{\mathbb{P}}(\tau^{[N^2]}=r) \\
&\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \Big(\mathbb{P}^{[N^2]}_{(i,D_1)}\Big(RW^{[N^2]}(r)=(l,R_l)\Big)-\mathbb{P}^{[N^2]}_{(j,D_1)}\Big(RW'^{[N^2]}(r)=(l,R_l)\Big)\Big)\\
&\times\int_0^{\bar{t}(N)-r}\mathrm{d}v\sum_{k\in[N^2]}\Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW^{[N^2]}(v)=(k,A)\Big)\\
&\qquad\qquad -\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW'^{[N^2]}(v)=(k,A)\Big)\Big)\\
&\times\mathbb{P}^{[N^2]}_{(l,R_l)}\Big(RW^{[N^2]}(v)=(k,A)\Big)\,\mathbb{E}[g(x^{[N^2]}_k(\bar{t}(N)-r-v))]\\
&=2\int_0^{\bar{t}(N)}\mathrm{d}r\,\tilde{\mathbb{P}}(\tau^{[N^2]}=r)\int_0^r\mathrm{d}u\,\tilde{\mathbb{P}}(\tau^{[N^2]}=u)\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}}\\
&\Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW^{[N^2]}(r-u)=(l,R_l)\Big)-\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW'^{[N^2]}(r-u)=(l,R_l)\Big)\Big)\\
&\times\int_0^{\bar{t}(N)-r}\mathrm{d}v\sum_{k\in[N^2]}\Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW^{[N^2]}(v)=(k,A)\Big)\\
&\qquad\qquad -\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW'^{[N^2]}(v)=(k,A)\Big)\Big)\\
&\times\mathbb{P}^{[N^2]}_{(l,R_l)}\Big(RW^{[N^2]}(v)=(k,A)\Big)\,\mathbb{E}[g(x^{[N^2]}_k(\bar{t}(N)-r-v))]\\
&+2\int_0^{\bar{t}(N)}\mathrm{d}r\,\tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r)\\
&\times\int_0^{\bar{t}(N)-r}\mathrm{d}v\sum_{k\in[N^2]}\Big(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW^{[N^2]}(v)=(k,A)\Big)\\
&\qquad\qquad -\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\Big(RW'^{[N^2]}(v)=(k,A)\Big)\Big)\\
&\times\Big(\mathbb{P}^{[N^2]}_{(i,D_1)}\Big(RW^{[N^2]}(v)=(k,A)\Big)-\mathbb{P}^{[N^2]}_{(j,D_1)}\Big(RW'^{[N^2]}(v)=(k,A)\Big)\Big)\\
&\times\mathbb{E}[g(x^{[N^2]}_k(\bar{t}(N)-r-v))].
\end{aligned}
\tag{8.200}
$$

We will show that both terms in the last equality of (8.200) tends to 0 as $N\to\infty$.

For the first term note that, by (8.171), (8.173) and (8.195), we have

$$
\begin{aligned}
&2\int_0^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\int_0^r \mathrm{d}u\, \tilde{\mathbb{P}}(\tau^{[N^2]}=u)\\
&\quad\times\Bigg[\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(r-u)=(l,R_l)\right)\\
&\qquad\qquad - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(r-u)=(l,R_l)\right)\Bigg]\\
&\quad\times\int_0^{\bar t(N)-r}\mathrm{d}v \sum_{k\in[N^2]}\Bigg(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(v)=(k,A)\right)\\
&\qquad\qquad - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(v)=(k,A)\right)\Bigg)\\
&\quad\times \mathbb{P}^{[N^2]}_{(l,R_l)}\left(RW^{[N^2]}(v)=(k,A)\right)\mathbb{E}[g(x_k(\bar t(N)-r-v))]\\
&\le 4\int_0^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\int_0^r \mathrm{d}u\, \tilde{\mathbb{P}}(\tau^{[N^2]}=u)\\
&\quad\times\Bigg|\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(r-u)=(l,R_l)\right)\\
&\qquad\qquad - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(r-u)=(l,R_l)\right)\Bigg| \qquad (8.201)\\
&\le 4\int_0^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\int_0^{r-L(N)} \mathrm{d}u\, \tilde{\mathbb{P}}(\tau^{[N^2]}=u)\\
&\quad\times\Bigg|\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(r-u)=(l,R_l)\right)\\
&\qquad\qquad - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(r-u)=(l,R_l)\right)\Bigg|\\
&+8\int_{L(N)}^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\in[r-L(N),r]) + 8\mathbb{P}[\tau^{[N^2]}\in[0,L(N)]]\\
&\le 4\int_0^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\int_0^{r-L(N)} \mathrm{d}u\, \tilde{\mathbb{P}}(\tau^{[N^2]}=u)\\
&\quad\times\Bigg|\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(r-u)=(l,R_l)\right)\\
&\qquad\qquad - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(r-u)=(l,R_l)\right)\Bigg|\\
&+16[1-\mathrm{e}^{-\frac{e_1}{N}L(N)}].
\end{aligned}
$$

Hence the last term in the last inequality tends to 0 as $N \to \infty$.

To show that the first term in the last inequality tends to 0 we use the coupling again. Recall the definition of $H_t^{[N^2]}$ in (8.176). Note that we can rewrite the sum as

$$\begin{aligned}
&\sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(r-u)=(l,R_l)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(r-u)=(l,R_l)\right)\\
&\leq \sum_{\substack{(l,R_l)\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \sum_{\substack{(l',R_{l'})\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left(\mathbb{P}^{[N^2]}_{(i,A)}\left(RW^{[N^2]}(L(N))=(l',R'_l)\right)\right.\\
&\qquad \times \mathbb{P}^{[N^2]}_{(l',R'_l)}\left(RW^{[N^2]}(r-u-L(N))=(l,R_l)\right)\\
&\leq \sum_{\substack{(l',R_{l'})\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left(\mathbb{P}^{[N^2]}_{(i,A)}\left(RW^{[N^2]}(L(N))=(l',R'_l)\right) - \mathbb{P}^{[N^2]}_{(j,A)}\left(RW'^{[N^2]}(L(N))=(l',R'_l)\right)\right)\\
&= \sum_{\substack{(l',R_{l'})\in\\ [N^2]\times\{A,D_0,D_1,D_2\}}} \left(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(L(N))=(l',R'_l),(H_t^{[N^2]})^c\right)\right.\\
&\qquad\qquad \left. - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(L(N))=(l',R'_l),(H_t^{[N^2]})^c\right)\right)\\
&\leq 2\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left((H_t^{[N^2]})^c\right).
\end{aligned} \tag{8.202}$$

To show that

$$\lim_{n\to\infty} \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left((H_t^{[N^2]})^c\right) = 0, \tag{8.203}$$

we can use a similar strategy as between (8.189) and (8.191), but note that we now start from two active sites instead of two 1-dormant sites. Therefore (8.187) directly follows from (8.190) and (8.191).

To show the second term in (8.200) tends to 0, we write it as

$$\begin{aligned}
&2\int_0^{\bar{t}(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r)\\
&\times \int_0^{\bar{t}(N)-r} \mathrm{d}v \sum_{k\in[N^2]} \left(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(v)=(k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(v)=(k,A)\right)\right)\\
&\times \left[\mathbb{P}^{[N^2]}_{(i,D_1)}\left(RW^{[N^2]}(v)=(k,A)\right) - \mathbb{P}^{[N^2]}_{(j,D_1)}\left(RW'^{[N^2]}(v)=(k,A)\right)\right] \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N)-r-v))]\\
&= 2\int_0^{\bar{t}(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r)\\
&\times \int_0^{\bar{t}(N)-r} \mathrm{d}v \sum_{k\in[N^2]} \left(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(v)=(k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(v)=(k,A)\right)\right)\\
&\times \int_0^{v} \mathrm{d}u\, \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]}=u)\\
&\quad \left(\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(v-u)=(k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(v-u)=(k,A)\right)\right)\\
&\times \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N)-r-v))].
\end{aligned} \tag{8.204}$$

Changing the order of integration and setting $w = v - u$, we obtain

$$
\begin{aligned}
&2\int_0^{\bar{t}(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]} = r)\tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
&\times \int_0^{\bar{t}(N)-r} \mathrm{d}u\, \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \int_u^{\bar{t}(N)-r} \mathrm{d}v \\
&\sum_{k\in[N^2]} \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW^{[N^2]}(v) = (k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW'^{[N^2]}(v) = (k,A)\right)\right] \\
&\times \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW^{[N^2]}(v-u) = (k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW'^{[N^2]}(v-u) = (k,A)\right)\right] \\
&\times \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N) - r - v))] \\
&= 2\int_0^{\bar{t}(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]} = r)\tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
&\times \int_0^{\bar{t}(N)-r} \mathrm{d}u\, \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \int_0^{\bar{t}(N)-r-u} \mathrm{d}w \\
&\sum_{k\in[N^2]} \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW^{[N^2]}(w+u) = (k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW'^{[N^2]}(w+u) = (k,A)\right)\right] \\
&\times \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW^{[N^2]}(w) = (k,A)\right) - \tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)} \left(RW'^{[N^2]}(w) = (k,A)\right)\right] \\
&\times \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N) - r - u - w))].
\end{aligned}
\tag{8.205}
$$

CHAPTER 8

This can be rewritten as

$$
\begin{aligned}
&2\int_0^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r)\\
&\int_0^{\bar t(N)-r} \mathrm{d}u\, \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]}=u)\\
&\quad\times \sum_{(l,R_l)\in[N^2]} \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(u)=(l,R_l)\right)-\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(u)=(l,R_l)\right)\right]\\
&\quad\times \int_0^{\bar t(N)-r-u} \mathrm{d}w \sum_{k\in[N^2]} \left[\mathbb{P}^{[N^2]}_{(l,R_l)}\left(RW^{[N^2]}(w)=(k,A)\right)\right]\\
&\quad\times \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(w)=(k,A)\right)-\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(w)=(k,A)\right)\right]\\
&\quad\times \mathbb{E}[g(x_k^{[N^2]}(\bar t(N)-r-u-w))]\\
&\leq 8\int_{L(N)}^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r)\\
&\int_{L(N)}^{\bar t(N)-r} \mathrm{d}u\, \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]}=u)\\
&\quad\times \sum_{(l,R_l)\in[N^2]} \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(u)=(l,R_l)\right)-\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(u)=(l,R_l)\right)\right]\\
&+8\int_{L(N)}^{\bar t(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r)\\
&\int_0^{L(N)} \mathrm{d}u\, \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]}=u)\\
&\quad\times \sum_{(l,R_l)\in[N^2]} \left[\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW^{[N^2]}(u)=(l,R_l)\right)-\tilde{\mathbb{P}}^{[N^2]}_{(i,A),(j,A)}\left(RW'^{[N^2]}(u)=(l,R_l)\right)\right]\\
&+16\int_0^{L(N)} \mathrm{d}r\, \tilde{\mathbb{P}}(\tau^{[N^2]}=r)\tilde{\mathbb{P}}(\tau^{[N^2]}\geq r).
\end{aligned}
\tag{8.206}
$$

This tends to 0 by (8.202) and the reasoning below (8.203). □

§8.3.9 Limiting evolution of the estimator processes

In this section we show that the results along the subsequences used in steps 5-8 of the scheme for the two-level three-colour mean-field system actually hold for all subsequences. Therefore the limiting evolution holds for $N\to\infty$. Recall that Lemma 8.3.20 tells us that all single 1-dormant colonies equal the value of the 1-dormant 1-block average. Therefore the second assumption in (8.81) in Proposition (8.3.5) can be replaced by

$$
\lim_{k\to\infty}\mathcal{L}\left[\left(Y_{2,0}^{[N_k^2]}(N_k^2t_2)\right)\Big|\boldsymbol{\Theta}^{\mathrm{aux},(1),[N_k^2]}(N_k^2t_2+N_kt_1)\right]=P^{z_1^{\mathrm{eff}}(t_1)},\tag{8.207}
$$

since, by Lemma 8.3.20, the limiting law

$$
\lim_{k\to\infty}\mathcal{L}\left[\left(Y_{1,0}^{[N_k^2]}(N_k^2t_2)\right)\right]\tag{8.208}
$$

is completely determined by the first line in (8.81). Hence, the assumptions in Proposition 8.3.10 and Lemma 8.3.18 can be weakened in the same way. Using that in Proposition 8.1.1 we assume (8.38) and (8.39), we find that the 2-block convergence stated in Lemma 8.3.18 holds along all subsequences we choose in Step 5. We conclude that Proposition 8.1.1(a) is indeed true. Combining Proposition 8.1.1(a) with steps 1-4 of the scheme and Lemma 8.3.20, we find that the assumptions in Proposition 8.3.10 are true for all subsequences, and we obtain the limiting evolution of the 1-block estimator process. Projecting this limiting evolution onto the active 1-block average and the 1-dormant 1-block average, we obtain Proposition 8.1.1(b). Finally, combining Proposition 8.1.1(a), steps1-4, and the fact that Proposition 8.3.10 is true along all subsequences, we obtain Proposition 8.1.1(c) and (f).

§8.3.10 Convergence in the Meyer-Zheng topology

In this section we show how the results on the effective and estimator processes can be use to show convergence of the full 1- and 2-block processes.

Lemma 8.3.22 ([Convergence of 1-process in the Meyer-Zheng topology]). *Assume that for the 1-block estimator process defined in* (8.21)

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\mathrm{aux},(1),[N^2]}(N^2t_2+Nt_1)\right)_{t_1>0}\right] = \mathcal{L}\left[(z_1^{\mathrm{aux}}(t_1))_{t_1>0}\right], \tag{8.209}$$

where, conditional on $x_2^{\mathrm{eff}}(t_2)=u$, the limit process is the unique solution of the SSDE in (8.30) *with θ replaced by u and with initial measure $\Gamma_u^{\mathrm{eff},(1)}$. Then*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{(1),[N^2]}(N^2t_2+Nt_1)\right)_{t_1>0}\right] = \mathcal{L}\left[(z_1^{\Gamma^{(1)}(t_2)}(t_1))_{t_1>0}\right] \quad \text{in the Meyer-Zheng topology,} \tag{8.210}$$

where $\Gamma^{(1)}(t_2)$ is defined as in (8.46) *and $(z_1^{\Gamma^{(1)}(t_2)}(t_1))_{t_1>0}$ is the process moving according to* (8.28) *with initial measure $\Gamma^{(1)}(t_2)$.*

Proof. By assumption 8.209 and Lemma 8.3.6, we can proceed as in the proof of Proposition 7.2.13 to find (8.210). □

Lemma 8.3.23 ([Convergence of 2-process in the Meyer-Zheng topology]).

Assume that for the effective 2-block process defined in (8.23)

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{\mathrm{eff},(2),[N^2]}(N^2t_2)\right)_{t_2>0}\right] = \mathcal{L}\left[(z_2^{\mathrm{aux}}(t_2))_{t_2>0}\right], \tag{8.211}$$

where $(z_2^{\mathrm{eff}}(t_2))_{t_2>0}$ is the process evolving according to (8.36) *and starting from $(\vartheta_1,\theta_{y_2})$. Then for the 2-block estimator process defined in* (8.23)

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\boldsymbol{\Theta}^{(2),[N^2]}(N^2t_2)\right)_{t_2>0}\right] = \mathcal{L}\left[(z_2(t_2))_{t_2>0}\right] \quad \text{in the Meyer-Zheng topology,} \tag{8.212}$$

where $(z_2(t_2))_{t_2>0}$ *is the process evolving according to* (8.35) *and starting in state* $(\vartheta_1, \vartheta_1, \vartheta_1, \theta_{y_2})$.

Proof. Combining Lemmas 8.3.6 and 8.3.13, we find for $t_2 > 0$

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \Theta_x^{(2),[N^2]}(N^2 t_2)\right|\right] = 0,\\
&\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \Theta_{y_0}^{(2),[N^2]}(N^2 t_2)\right|\right] = 0,\\
&\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \Theta_{y_1}^{(2),[N^2]}(N^2 t_2)\right|\right] = 0.
\end{aligned} \tag{8.213}$$

Therefore we can again proceed as in the proof of Proposition 7.2.13 to find (8.212).
□

§8.3.11 Proof of the two-level three-colour mean-field finite-systems scheme

In Section 8.3.9 we already proved Proposition 8.1.1(a),(b),(c) and (f). The proof of Proposition 8.1.1(d) follows from Proposition 8.1.1(a) by applying Lemma 8.3.23. The proof of Proposition 8.1.1(e) follows from Proposition 8.1.1(b) by applying Lemma 8.3.22. This completes the proof of Proposition 8.1.1.

CHAPTER 9

Proofs of the hierarchical multi-scale limit theorems

In this chapter we prove the hierarchical multi-scale limit theorems stated in Theorem 4.4.2 and Theorem 4.4.4. In Section 9.1 we first introduce the finite-level mean-field finite-systems scheme. In Section 9.2 we given an outline of how to prove the finite-level mean-field finite-systems scheme. In Section 9.3 we show how Theorems 4.4.2–4.4.4 can be obtained by a simple generalisation of the finite-level mean-field finite-systems scheme. The proof of the finite-level mean-field finite-systems scheme follows a similar line of argument as in Section 8.2 once we incorporate more levels. Since the proofs for the finite-level mean-field systems scheme are similar as the proofs in Section 8.3, we will not write out the full proof, but only give an outline and a sketch.

§9.1 Finite-level mean-field finite-systems scheme and interaction chain

In this section we extend the two-level three-colour system to a k-level $(k+1)$-colour system with an "outside world" for any $k \in \mathbb{N}$. This outside world allows also the highest level, the k-block, to start from equilibrium. It is also needed to generalize the results in this subsection to the infinite hierarchical group.

► **Definitions.** To set up the system, fix $k \in \mathbb{N}$ and consider the geographic space Ω_N^{k+1} obtained by truncating the hierarchical goup Ω_N (recall (4.2)) after hierarchical level $k+1$, i.e., $\Omega_N^{k+1} = B_{k+1}(0)$ the $(k+1)$-block centred at the origin (recall (4.2), (4.4) and Fig. 4.2). Note that the $k+1$-block consists of N k-blocks i.e., $B_{k+1}(0) = \bigcup_{i=0}^{N} B_k(i)$ and $B_{k+1}(0) = [N^{k+1}]$. The seed-bank in this model consists of the $k+2$ layers corresponding to colours $\{0, \cdots, k\} \cup \{k+1\}$. On this space we again consider a restricted version of the SSDE in (4.20) to the geographic space Ω_N^{k+1}. The migration kernel $a^{\Omega_N}(\cdot,\cdot)$ is restricted to Ω_N^{k+1} by setting all migration outside $B_{k+1}(0)$ equal to 0, i.e.,

$$a^{[\Omega_N^{k+1}]}(\xi,\eta) = \sum_{l=1}^{k+1} 1_{\left\{d_{\Omega_N^{k+1}}(\xi,\eta) \leq l\right\}} \frac{c_l}{N^{l-1}} \frac{1}{N^l}, \tag{9.1}$$

where $d_{\Omega_N^{k+1}}$ is the hierarchical distance d_{Ω_N} restricted to the space Ω_N^{k+1}. The colour-l dormant population exchanges individuals with the active population at rates $\frac{e_l}{N^l}$, $\frac{K_l e_l}{N^l}$ for all $0 \leq l \leq k$. We set the interaction of the active population with the colour $(k+1)$-dormant population equal to 0. This seed-bank is only needed later, namely for the "outside world".

The state space of the finite-level mean-field system is

$$S = (\mathfrak{s}^{k+1})^{\Omega_N^{k+1}}, \quad \mathfrak{s}^{k+1} = [0,1] \times [0,1]^{k+2}, \tag{9.2}$$

and the system is denoted by

$$\begin{aligned} &Z^{\Omega_N^{k+1}} = (Z^{\Omega_N^{k+1}}(t))_{t \geq 0}, \qquad Z^{\Omega_N^{k+1}}(t) = (z_\xi^{\Omega_N^{k+1}}(t))_{\xi \in \Omega_N^{k+1}}, \\ &z_\xi^{\Omega_N^{k+1}}(t) = (x_\xi^{\Omega_N^{k+1}}(t), (y_{\xi,m}^{\Omega_N^{k+1}}(t))_{m=0}^{k+1}). \end{aligned} \tag{9.3}$$

The components evolve according to the SSDE

$$\begin{aligned} \mathrm{d}x_\xi^{\Omega_N^{k+1}}(t) &= \sum_{l=1}^{k} \frac{c_{l-1}}{N^{l-1}} \frac{1}{N^l} \sum_{\eta \in B_l(\xi)} \left[x_\eta^{\Omega_N^{k+1}}(t) - x_\xi^{\Omega_N^{k+1}}(t) \right] \mathrm{d}t \\ &\quad + \sqrt{g(x_\xi^{\Omega_N^{k+1}}(t))}\, \mathrm{d}w_\xi(t) + \sum_{m=0}^{k} \frac{K_m e_m}{N^m} \left[y_{\xi,m}^{\Omega_N^{k+1}}(t) - x_\xi^{\Omega_N^{k+1}}(t) \right] \mathrm{d}t, \\ \mathrm{d}y_{\xi,m}^{\Omega_N^{k+1}}(t) &= \frac{e_m}{N^m} \left[x_\xi^{\Omega_N^{k+1}}(t) - y_{\xi,m}^{\Omega_N^{k+1}}(t) \right] \mathrm{d}t, \quad 0 \leq m \leq k, \\ \mathrm{d}y_{\xi,m}^{\Omega_N^{k+1}}(t) &= 0, \qquad \xi \in \Omega_N^{k+1}, \end{aligned} \tag{9.4}$$

with $B_l(\xi)$ the ball of radius l around $\xi \in \Omega_N^{k+1}$.

Note that this system is the hierarchical SSDE in (4.20) with all interactions at distance $> k$ switched off (i.e., $c_l = 0$ for $l > k+1$), and also the exchange with dormant populations of colour $m > k$ is switched off. As before, by [67] the martingale problem associated with (9.4) is well-posed, and for every initial state in S the SSDE has a unique strong solution. We will analyse (9.4) on time scales $1, N, N^2, \cdots, N^k$. If for $0 \leq l \leq k$ time runs on time scale N^l, then we write $N^l t_l$ with $t_l > 0$.

To study the k-level mean-field system, we analyse the equivalent of the block averages defined in (4.2.3). For the k-level mean-field system these are given by

$$\begin{aligned} x_l^{\Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{\eta \in B_l(0)} x_\eta^{\Omega_N^{k+1}}(N^l t), \\ y_{m,l}^{\Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{\eta \in B_l(0)} y_{\eta,m}^{\Omega_N^{k+1}}(N^l t), \qquad 0 \leq m \leq k+1, \qquad 0 \leq l \leq k. \end{aligned} \tag{9.5}$$

For $0 \le l \le k$ these block averages evolve according to the SSDE

$$\begin{aligned} \mathrm{d}x_l^{\Omega_N^{k+1}}(t) &= \sum_{n=1}^{k-(l-1)} \frac{c_{l+n-1}}{N^{n-1}} \left[x_{l+n}^{\Omega_N^{k+1}}(N^{-n}t) - x_l^{\Omega_N^{k+1}}(t)\right] \mathrm{d}t \\ &\quad + \sqrt{\frac{1}{N^l} \sum_{i \in B_l(0)} g(x_i(N^l t))}\, \mathrm{d}w_l(t) \\ &\quad + \sum_{m=0}^{k} N^l \frac{K_m e_m}{N^m} \left[y_{m,l}^{\Omega_N^{k+1}}(t) - x_l^{\Omega_N^{k+1}}(t)\right] \mathrm{d}t, \end{aligned} \tag{9.6}$$

$$\mathrm{d}y_{m,l}^{\Omega_N^{k+1}}(t) = N^l \frac{e_m}{N^m} \left[x_l^{\Omega_N^{k+1}}(t) - y_{m,l}^{\Omega_N^{k+1}}(t)\right] \mathrm{d}t, \qquad 0 \le m \le k, \tag{9.7}$$

$$\mathrm{d}y_{k+1,l}^{\Omega_N^{k+1}}(t) = 0, \tag{9.8}$$

In the limit as $N \to \infty$, the active l-block average feels a drift towards the active $(l+1)$-block average, which is not moving on time scale N^l, at rate c_l. The diffusion term for the l-block average becomes the average diffusion over the l-block. The drift of the active l-block average towards the l-block average of m-dormant populations $y_{m,l}^{\Omega_N^{k+1}}$ with $m > l$ vanishes in the limit as $N \to \infty$. Therefore, the $m > l$ m-dormant populations are *slow seed-banks* on space-time scale l. The l-block average of the colour-l dormant population $y_{l,l}^{\Omega_N^{k+1}}$ has a non-trivial drift towards the active l-block average, written $x_l^{\Omega_N^{k+1}}$. Therefore the l-dormant population is the *effective seed-bank* on space-time scale l. For the colour m-dormant populations $y_{m,l}^{\Omega_N^{k+1}}$ with $m < l$, we see that infinite rates appear. Therefore the m-dormant populations with $m < l$ are *fast seed-banks* on space-time scale l. We again need the Meyer-Zheng topology to show that $\lim_{N\to\infty} y_{m,l}^{\Omega_N^{k+1}} = \lim_{N\to\infty} x_l^{\Omega_N^{k+1}}$. On space-time scale l, the colour-l dormant population is the effective seed-bank. To get rid of the infinite rates we again look at combinations. From the above discussion and the SSDE in (9.6)–(9.6), we see that if we consider the quantity

$$\frac{x_l^{\Omega_N^{k+1}}(t) + \sum_{m=0}^{l-1} K_m y_{l,m}^{\Omega_N^{k+1}}(t)}{1 + \sum_{m=0}^{l-1} K_m}, \tag{9.9}$$

then all infinite rates cancel out. Therefore

$$\left(\frac{x_l^{\Omega_N^{k+1}}(t) + \sum_{m=0}^{l-1} K_m y_{l,m}^{\Omega_N^{k+1}}(t)}{1 + \sum_{m=0}^{l-1} K_m}, y_{l,l}^{\Omega_N^{k+1}}(t) \right)_{t>0} \tag{9.10}$$

is called the *effective process* on space-time scale l. Like in the simpler mean-field finite-systems scheme, the effective process allows us to analyse our system in path space.

An important difference between the finite-level mean-field system in (9.6)–(9.7) and the two-level three-colour mean-field system in Section (8.1) is that in the finite-level mean-field system also the highest level k has a drift towards the outside world.

This outside world is the active $k+1$-block average, which does not evolve on time scale N^k. This drift allows the finite-level mean-field system to equilibrate to a non-trivial equilibrium. In the two-level mean-field system, the highest level, i.e., the active 2-block average, does not feel a drift due to migration. Consequently, the 2-block averages will eventually cluster.

▶ **Scaling limit.** To state and prove the finite-level multi-scale limit, we need the following three limiting processes. Recall (4.64) and (4.62). For $0 \leq l \leq k$, let

$$(z_{l,(\theta,(y_{m,l})_{m=0}^{k+1})}(t))_{t\geq 0}) = (x_l(t), (y_{m,l}(t))_{m=0}^{k+1})_{t\geq 0}) \tag{9.11}$$

be the process evolving according to

$$\begin{aligned}
\mathrm{d}x_l(t) &= E_l\left[c_l[\theta - x_l(t)]\,\mathrm{d}t + \sqrt{\mathcal{F}^{(l)}g(x_1(t))}\,\mathrm{d}w(t) + K_l e_l\,[y_{l,l}(t) - x_l(t)]\,\mathrm{d}t\right],\\
y_{m,l}(t) &= x_l(t), \qquad \text{for } 0 \leq m < l\\
\mathrm{d}y_{l,l}(t) &= e_l\,[x_l(t) - y_{l,l}(t)]\,\mathrm{d}t,\\
y_{m,l}(t) &= y_{m,l}, \qquad \text{for } l < m \leq k+1,
\end{aligned} \tag{9.12}$$

where $\theta \in [0,1]$ and $y_{m,l} \in [0,1]$ for $l < m \leq k+1$.

For $0 \leq l \leq k$, let

$$(z^{\mathrm{aux}}_{l,(\theta,(y_{m,l})_{m=l+1}^{k+1})}(t))_{t\geq 0} = (x^{\mathrm{aux}}_l(t), (y^{\mathrm{aux}}_{m,l}(t))_{m=l+1}^{k+1})_{t\geq 0} \tag{9.13}$$

be the process evolving according to

$$\begin{aligned}
\mathrm{d}x^{\mathrm{aux}}_l(t) = E_l\Bigg[&c_l[\theta - x^{\mathrm{aux}}_l(t)]\,\mathrm{d}t + \sqrt{\mathcal{F}^{(l)}g(x^{\mathrm{aux}}_1(t))}\,\mathrm{d}w(t)\\
&+ K_l e_l\,[y^{\mathrm{aux}}_{l,l}(t) - x^{\mathrm{aux}}_l(t)]\,\mathrm{d}t\Bigg],
\end{aligned} \tag{9.14}$$

$$\begin{aligned}
\mathrm{d}y^{\mathrm{aux}}_{l,l}(t) &= e_l\,[x^{\mathrm{aux}}_l(t) - y^{\mathrm{aux}}_{l,l}(t)]\,\mathrm{d}t,\\
y^{\mathrm{aux}}_{m,l}(t) &= y_{m,l}, \qquad \text{for } l < m \leq k+1,
\end{aligned}$$

where $\theta \in [0,1]$ and $y^{\mathrm{aux}}_{m,l} \in [0,1]$, for $l < m \leq k+1$.

For $0 \leq l \leq k$, let

$$(z^{\mathrm{eff}}_{l,\theta}(t))_{t\geq 0} = \left(x^{\mathrm{eff}}_l(t), y^{\mathrm{eff}}_{l,l}(t)\right)_{t\geq 0} \tag{9.15}$$

be the *effective process* evolving according to

$$\begin{aligned}
\mathrm{d}x^{\mathrm{eff}}_l(t) &= E_l\left[c_l\,[\theta - x^{\mathrm{eff}}_l(t)]\,\mathrm{d}t + \sqrt{(\mathcal{F}^{(l)}g)(x^{\mathrm{eff}}_l(t))}\,\mathrm{d}w(t) + K_l e_l\,[y^{\mathrm{eff}}_{l,l}(t) - x^{\mathrm{eff}}_l(t)]\,\mathrm{d}t\right],\\
\mathrm{d}y^{\mathrm{eff}}_{l,l}(t) &= e_l\,[x^{\mathrm{eff}}_l(t) - y^{\mathrm{eff}}_{l,l}(t)]\,\mathrm{d}t.
\end{aligned} \tag{9.16}$$

Comparing (9.12) with (9.16), we see that the effective process looks at the non-trivial components of the full process. The auxiliary process in (9.14) looks at the active population, the effective seed-bank and the slow seed-banks.

To state and prove the finite-level multi-scale limit, we need the following list of ingredients:

(a) For $t > 0$ and for $0 \leq l \leq k$, define the *l-block estimators*

$$\begin{aligned}
\bar{\Theta}^{(l),\Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{i \in B_l} \frac{x_i^{\Omega_N^{k+1}}(t) + \sum_{m=0}^{l-1} K_m y_{i,0}^{\Omega_N^{k+1}}(t)}{1+K_0}, \\
\Theta_x^{(l),\Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{i \in B_l} x_i^{\Omega_N^{k+1}}(t), \\
\Theta_{y_m}^{(l),\Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{i \in B_l} y_{i,m}^{\Omega_N^{k+1}}(t), \qquad 0 \leq m \leq k+1,
\end{aligned} \tag{9.17}$$

and put

$$\begin{aligned}
\mathbf{\Theta}^{(l),\Omega_N^{k+1}}(t), &= \left(\Theta_x^{(l),\Omega_N^{k+1}}(t), \left(\Theta_{y_m}^{(l),\Omega_N^{k+1}}(t)\right)_{m=0}^{k+1}\right), \\
\mathbf{\Theta}^{\mathrm{aux},(l),\Omega_N^{k+1}}(t) &= \left(\bar{\Theta}^{(l),\Omega_N^{k+1}}(t), \left(\Theta_{y_l}^{(l),\Omega_N^{k+1}}(t)\right)_{m=l}^{k+1}\right), \\
\mathbf{\Theta}^{\mathrm{eff},(l),\Omega_N^{k+1}}(t) &= \left(\bar{\Theta}^{(l),\Omega_N^{k+1}}(t), \Theta_{y_l}^{(l),\Omega_N^{k+1}}(t)\right).
\end{aligned} \tag{9.18}$$

We call $(\mathbf{\Theta}^{(l),\Omega_N^{k+1}}(t))_{t>0}$ the *l-block estimator process*, $(\mathbf{\Theta}^{\mathrm{aux},(l),\Omega_N^{k+1}}(t))_{t>0}$ the *auxiliary l-block estimator process* and $(\mathbf{\Theta}^{\mathrm{eff},(l),\Omega_N^{k+1}}(t))_{t>0}$ the *effective l-block estimator process*.

(b) For $0 \leq l \leq k$, define the *time scales* N^l such that

$$\mathcal{L}[\bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l - L(N)N^{l-1}) - \bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l)] = \delta_0 \tag{9.19}$$

for all $L(N)$ such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$, but not for $L(N) = N$. In words, N is the time scale on which $\bar{\Theta}^{(l),\Omega_N^{k+1}}(\cdot)$ starts evolving, i.e., $\left(\bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l)\right)_{t_l>0}$, is no longer a fixed process.

(c) The *invariant measure* for the evolution of the l-block average in (9.12), denoted by

$$\Gamma_{\theta,y_l}^{(l)}, \qquad y_l = (y_{m,l})_{m=0}^{k+1}. \tag{9.20}$$

The *invariant measures* of the auxiliary l-block process in (9.15) and the effective l-block process in (9.16), denoted by, respectively,

$$\Gamma_{\theta,y_l}^{(l),\mathrm{aux}}, \qquad y_l = (y_{m,l})_{m=l+1}^{k+1} \tag{9.21}$$

and

$$\Gamma_\theta^{(l),\mathrm{eff}}. \tag{9.22}$$

(d) For $0 \leq l \leq k$, let $\mathcal{F}^{E_l,c_l,K_l,e_l}$ denote the *renormalisation transformation* acting on $\mathcal{G}$ defined by

$$(\mathcal{F}^{E_l,c_l,K_l,e_l} g)(\theta) = \int_{[0,1]^2} g(x)\, \Gamma_\theta^{(l)}(\mathrm{d}x, (\mathrm{d}y)_{m=0}^{k+1}), \qquad \theta \in [0,1], \tag{9.23}$$

and define the *iterates* $\mathcal{F}^{(n)}$, $0 \leq n \leq k$, of the renormalisation transformation as the compositions

$$\mathcal{F}^{(l)} = \mathcal{F}^{E_{n-1},c_{n-1},K_{n-1},e_{n-1}} \circ \cdots \circ \mathcal{F}^{E_0,c_0,K_0,e_0}, \qquad 0 \leq l \leq k. \tag{9.24}$$

(Recall (4.76).)

(e) To give a detailed description of the multi-scale behaviour of the SSDE in (4.20), define the *interaction chain*

$$(M^k_{-l})_{-l=-(k+1),-k,\dots,0} \tag{9.25}$$

as the *time-inhomogeneous* Markov chain on $[0,1] \times [0,1]^{k+1}$ with initial state

$$M^k_{-(k+1)} = (\vartheta_k, \overbrace{\vartheta_k, \cdots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y,k+1}) \tag{9.26}$$

that evolves according to the transition kernel $Q^{[l]}$ from time $-(l+1)$ to time $-l$ given by

$$Q^{[l]}(u, \mathrm{d}v) = \Gamma^{(l)}_u(\mathrm{d}v), \qquad 0 \leq l \leq k. \tag{9.27}$$

(Recall (4.77).)

We are now ready to state the scaling limit for the evolution of the averages in (7.7).

Proposition 9.1.1. *[Finite-level mean-field: finite-systems scheme] Suppose that the initial state of the system in* (9.4) *is given by* $\mu(0) = \mu^{\otimes[\Omega^{k+1}_N]}$ *for some* $\mu \in \mathcal{P}([0,1] \times [0,1]^{k+2})$*. Let* $L(N)$ *be such that* $\lim_{N\to\infty} L(N) = \infty$ *and* $\lim_{N\to\infty} L(N)/N = 0$*, and for* $t_k, \dots, t_0 \in (0,\infty)$ *set* $\bar{t} = L(N)N^k + \sum_{n=0}^k t_n N^n$*.*

(a) For every $t_k, \dots, t_0 \in (0,\infty)$*,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(l),\Omega^{k+1}_N}(\bar{t})\right)_{l=k+1,k,\dots,0}\right] = \mathcal{L}\left[(M^k_{-l})_{-l=-(k+1),-k,\dots,0}\right], \tag{9.28}$$

where $(M^k_{-l})_{-l=-(k+1),-k,\dots,0}$ *is the interaction chain in* (9.25) *starting from*

$$M^k_{-(k+1)} = (\vartheta_k, \overbrace{\vartheta_k, \cdots, \vartheta_k}^{k+1 \textit{ times}}, \theta_{y,k+1}). \tag{9.29}$$

(b) For all $0 \leq l \leq k$*,*

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{(l),\Omega^{k+1}_N}(\bar{t} + t_l N^l)\right)_{t_l>0}\right] = \mathcal{L}\left[\left(z_{l,M^k_{-(l+1)}}(t_l)\right)_{t_l>0}\right], \tag{9.30}$$

in the Meyer-Zheng topology,

where

$$(z_{l,M^k_{-(l+1)}}(t_l))_{t_l>0} \tag{9.31}$$

is the processes defined (9.12) *with* θ, $(y_m, l)_{m=l+1}^{k+1}$ *given by the corresponding components in* $M^k_{-(l+1)}$ *and with initial measure*

$$\mathcal{L}\left[z_{l,M^k_{-(l+1)}}(0)\right] = \Gamma^{(l)}_{M^k_{-(l+1)}}$$
$$\Gamma^{(l)}_{M^k_{-(l+1)}} = \int_{\mathfrak{s}^{k+1}} \cdots \int_{\mathfrak{s}^{k+1}} \int_{\mathfrak{s}^{k+1}} \Gamma^{(k)}_{M^k_{-(k+1)}}(\mathrm{d}u_k)\Gamma^{(k-1)}_{u_k}(\mathrm{d}u_{k-1}) \cdots \Gamma^{(l+1)}_{u_2}(\mathrm{d}u_{l+2})\Gamma^{(l)}_{u_{l+1}}. \tag{9.32}$$

In Part (a), the limit does not depend on the choice of the times $t_k, \dots, t_0$, since we let time start from a time larger than $L(N)N^k$, so that in the limit as $N \to \infty$ all the l-block averages with $l \leq k$ are already in quasi-equilibrium. In Part (b), for $l < k$ the center of the drift for the active population is *random* and is determined by the first component of the interaction chain. Also the states of the m-dormant populations with $l < m \leq k+1$ are determined by the interaction chain.

Remark 9.1.2. In contrast to Propositions 7.1.2–8.1.1, there are no assumptions on the seed-bank behaviour in Proposition 9.1.1. This is because all the block-averages that we consider are in equilibrium at time $\bar{t}$. Consequently on space-time scales $l < m$ the m-dormant l-block average will equal the state of the m-dormant m-block average at time $\bar{t}$. Therefore we say that the state of the *slow seed-banks* is determined by the space-time scaleon which this seed-bank is effective. Hence the state of the slow seed-banks is determined by the interaction chain. ■

The proof of Proposition 9.1.1 will be given in Section 9.2.

§9.2 Proof of the mean-field finite-systems scheme: finite-level

We give a sketch of the proof Proposition 9.1.1. The proof uses a similar scheme as the proof of Proposition 8.1.1. We state the scheme and indicate at each step how it can be proved.

1 Tightness of the auxiliary l-block estimator processes, for $0 \leq l \leq k$,

$$\left(\left(\mathbf{\Theta}^{\mathrm{aux},(l),\Omega_N^{k+1}}(N^l t_l)\right)_{t_l > 0}\right)_{N \in \mathbb{N}}. \tag{9.33}$$

Proof. For each $0 \leq l \leq k$ we use the tightness criterion in [49, Proposition 3.2.3.]. □

2 Stability property of the 2-block estimators, i.e., for $L(N)$ such that $\lim_{N\to\infty} L(N) = \infty$ and $\lim_{N\to\infty} L(N)/N = 0$,

$$\lim_{N\to\infty} \sup_{0 \leq t \leq L(N)} \left|\bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l) - \bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l - N^{l-1} t)\right| = 0 \text{ in probability} \tag{9.34}$$

and, for all $l \leq m \leq k+1$,

$$\lim_{N\to\infty} \sup_{0\leq t\leq L(N)} \left| \Theta_{y_m}^{(l),\Omega_N^{k+1}}(N^l t_l) - \Theta_{y_m}^{(l),\Omega_N^{k+1}}(N^l t_l - N^{l-1} t_l) \right| = 0 \text{ in probability.} \tag{9.35}$$

Proof. Use a similar computation as in the proof of Lemma 8.3.4. □

3 We analyse the behaviour of the slow seed-banks by proving the following lemma.

Lemma 9.2.1. *[Slow seed-banks in the multi-level system] Let $\Theta_{y_m,i}^{(l)}$ denote the m-dormant l-block average containing colony $i \in \Omega_N^{k+1}$. Then for all $i \in \Omega_N^{k+1}$, $m < k+1$, $l < m$ and $t_l > 0$,*

$$\lim_{N\to\infty} \left[y_{i,m}^{\Omega_N^{k+1}}(\bar{t} + N^l t_l) - \Theta_{y_m,i}^{(m),\Omega_N^{k+1}}(\bar{t} + N^l t_l) \right] = 0 \quad a.s. \tag{9.36}$$

and hence

$$\lim_{N\to\infty} \left[\Theta_{y_m,i}^{(l),\Omega_N^{k+1}}(\bar{t} + N^l t_l) - \Theta_{y_m,i}^{(m),\Omega_N^{k+1}}(\bar{t} + N^l t_l) \right] = 0 \quad a.s. \tag{9.37}$$

Proof. We can proceed as in the proof of Lemma 8.3.20, after adapting the kernel $b^{[N^2]}(\cdot,\cdot)$ to the kernel $b^{\Omega_N^{k+1}}(\cdot,\cdot)$. Then we can use that, from each of the $m < k+1$ m-dormant populations, individuals wake up before time $\bar{t}$ with probability 1. For individuals starting from an m-dormant state, we define the coupling event

$$H_t^{m,\Omega_N^{k+1}} = \{RW^{\Omega_N^{k+1}} \text{ has migrated over distance } m \text{ at least once up to time } t\}. \tag{9.38}$$

The migration over distance m is needed because we need m-dormant individuals to be uniformly distributed over the m-block in order to almost surely equal the state of the m-block. □

4 We prove the convergence of the single components. Recall that there are N^{k+1-l} l-blocks in Ω_N^{k+1}. Since tightness of components implies tightness of the process, step 1 implies that for $0 \leq l \leq k$ the full l-block processes

$$\left(\left(\mathbf{\Theta}_i^{\text{aux},(l),\Omega_N^{k+1}}(\bar{t} + N^l t_l) \right)_{t_l>0,\, i\in[N^{k+1-l}]} \right)_{N\in\mathbb{N}} \tag{9.39}$$

are tight. From the tightness in steps 1 we can construct a subsequence $(N_n)_{n\in\mathbb{N}}$ along which, for all $0 \leq l \leq k$,

$$\lim_{n\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}_i^{\text{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t} + N_n^l t_l) \right)_{t_l>0,\, i\in[N_n^{k+1-l}]} \right] \tag{9.40}$$

exists. Note that $\bar{t}$ depends on the subsequence. For example, along the subsequence $(N_{\tilde{n}})_{\tilde{n}\in\mathbb{N}}$,

$$\bar{t} = L(N)N^k + \sum_{n=0}^{k} t_n N_{\tilde{n}}^n. \tag{9.41}$$

We define the measure

$$\nu_{\mathbf{\Theta}}^{(0)} = \prod_{i\in\mathbb{N}_0} \Gamma_{\Theta_i}^{(0)}(\bar{t}), \tag{9.42}$$

where

$$\Theta_i \in \mathfrak{s}^{k+1}. \tag{9.43}$$

In this step we show that along the same subsequence the single components converge to the infinite system. We show that if

$$\lim_{n\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i\in[N_n^k]}] = P^{(1)}, \tag{9.44}$$

then

$$\lim_{n\to\infty} \mathcal{L}\Big[\Big(Z^{\Omega_{N_n}^{k+1}}(\bar{t}+t_0)\Big)_{t_0\geq 0}\Big] = \mathcal{L}\left[(Z^{\nu^{(0)}(\bar{t})}(t_0))_{t_0\geq 0}\right], \tag{9.45}$$

where

$$\nu^{(0)}(\bar{t}) = \int \nu_u^{(0)} P^{(1)}(\mathrm{d}u). \tag{9.46}$$

Here, $(Z^{\nu^{(0)}(\bar{t})}(t_0))_{t_0\geq 0}$ is the process starting from $\nu^{(0)}(\bar{t})$ with components evolving according to (8.18), where θ is now a random variable that inherits its law from

$$\lim_{n\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i\in[N_n^k]}] \tag{9.47}$$

and, similarly, the laws of $y_{m,0}$, $1 \leq l \leq k+1$ in the limiting process $(Z^{\nu^{(0)}(t_2)}(t_0))_{t_0\geq 0}$ are determined by

$$\lim_{n\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i\in[N_n^k]}]. \tag{9.48}$$

Note that we choose the subsequence $(N_n)_{n\in\mathbb{N}}$ in such a way that we know that the law $P^{(1)}$ in (9.44) exists.

Proof. Proceed as in the proof of Proposition 8.3.5. Note that the assumptions on the seed-banks in Proposition 8.3.5 follow from the choice of the subsequence and Lemma 9.2.1. □

5 Using the limiting evolution of the single colonies obtained in step 4, we can identify the limiting l-block process along the same subsequence. For $1 \leq l \leq k$, we show that if

$$\lim_{n\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(l+1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i\in[N_n^k]}] = P^{(l+1)}, \tag{9.49}$$

then

$$\lim_{n\to\infty} \mathcal{L}\left[\Big(\mathbf{\Theta}^{\mathrm{aux},(l),\Omega_{N_n}^{k+1}}(\bar{t}+N_n^l t_l)\Big)_{t_l>0,\, i\in[N_n^{k+1-l}]}\right] = \mathcal{L}\left[(z_{l,\mathbf{\Theta}^{(l+1)}}^{\mathrm{aux}}(t))_{t\geq 0}\right], \tag{9.50}$$

where $\mathbf{\Theta}^{(l+1)} = (\mathbf{\Theta}_x^{(l+1)}, (\mathbf{\Theta}_{y_m,l}^{(l+1)})_{m=l+1}^{k+1}) \in [0,1] \times [0,1]^{k+2-(l+1)}$,

$$\begin{aligned}\mathcal{L}\left[z^{\mathrm{aux}}_{l,\mathbf{\Theta}^{(l)}}(0)\right] &= \Gamma^{(l),\mathrm{aux}}_{\mathbf{\Theta}^{(l+1)}},\\ \Gamma^{(l),\mathrm{aux}}_{\mathbf{\Theta}^{(l+1)}} &= \int_{[0,1]\times[0,1]^{k+2-(l+1)}} \Gamma^{(l),\mathrm{aux}}_{u} P^{(l+1)}(\mathrm{d}u)\end{aligned} \tag{9.51}$$

and $(z^{\mathrm{aux}}_{l,\mathbf{\Theta}^{(l+1)}}(t))_{t\geq 0}$ is the process evolving according to (9.14) with $\theta, (y_{m,l})_{m=l}^{k+1}$ replaced by the random variables $\mathbf{\Theta}_x^{(l+1)}, (\mathbf{\Theta}_{y_m,l}^{(l+1)})_{m=l}^{k+1}$. Note that by the choice of the subsequence $(N_n)_{n\in\mathbb{N}}$ we know that for $1 \leq l \leq k$ the limiting laws in (9.49) exist.

Proof. The proof goes by induction. Using the convergence of the single components, we can proceed as in the proof of Proposition 8.3.10 to prove the convergence of the 1-blocks averages

$$\lim_{n\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}_i^{\mathrm{aux},(1),\Omega^{k+1}_{N_n}}(\bar{t} + N_n t_1)\right)_{t_1>0,\, i\in[N_n^{k+1-l}]}\right]. \tag{9.52}$$

Then, assuming that we have the convergence for all $0 \leq l \leq L$, we get

$$\lim_{n\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}_i^{\mathrm{aux},(l),\Omega^{k+1}_{N_n}}(\bar{t} + N_n^l t_l)\right)_{t_l>0,\, i\in[N_n^{k+1-l}]}\right], \tag{9.53}$$

and we prove the convergence of

$$\lim_{n\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}_i^{\mathrm{aux},(L+1),\Omega^{k+1}_{N_n}}(\bar{t} + N_n^{(L+1)} t_{(L+1)})\right)_{t_{(L+1)}>0,\, i\in[N_n^{k+1-(L+1)}]}\right]. \tag{9.54}$$

This is done using a similar proof strategy as in the proof of Proposition 8.3.10. In particular, we need to derive the l-level equivalent of Lemma 8.3.13. Since this lemma is also key to proving convergence in the Meyer-Zheng topology, we state it explicitly below. □

Chapter 9

Lemma 9.2.2 (l-block averages). *Define*

$$\begin{aligned}&\Delta_{\Sigma}^{(l),\Omega_N^{k+1}}(N^{l-1}t_{l-1})\\ &= \frac{\Theta_x^{(l),\Omega_N^{k+1}}(N^{l-1}t_{l-1}) + \sum_{m=0}^{l-2} K_m \Theta_{y_m}^{(l),\Omega_N^{k+1}}(N^{l-1}t_{l-1})}{1+\sum_{m=0}^{l-2}K_m} - \Theta_{y_{l-1}}^{(l),\Omega_N^{k+1}}(N^{l-1}t_{l-1})\end{aligned} \tag{9.55}$$

and

$$R_l = \frac{1+\sum_{m=0}^{l-1}K_m}{1+\sum_{m=0}^{l-2}K_m}. \tag{9.56}$$

For $t \geq 0$ *set* $\bar{\Theta}^{(0)}(t) = \Theta_x^{(0)}(t) = x_0(t)$. *Then, for* $1 \leq l \leq k$,

$$
\begin{aligned}
&\mathbb{E}\left[\left|\Delta_{\Sigma}^{(l),\Omega_N^{k+1}}(N^{l-1}t_{l-1})\right|\right] \\
&\leq \sqrt{\mathbb{E}\left[\left(\Delta_{\Sigma}^{(l),\Omega_N^{k+1}}(0)\right)^2\right]} \mathrm{e}^{-e_l R_l t_{l-1}} \\
&\quad + \sqrt{\int_0^{t_1} \mathrm{d}s\, 2e_l R_l \mathrm{e}^{-2e_l R_l(t_1-s)} \mathbb{E}\left[\left|\bar{\Theta}^{(l-1),\Omega_N^{k+1}}(N^{l-1}s) - \Theta_x^{(l-1),\Omega_N^{k+1}}(N^{l-1}s)\right|\right]} \\
&\quad + \sqrt{\frac{1}{N}\frac{1}{2e_l(1+\sum_{m=0}^{l-1}K_m)}\left[\sum_{n=l+1}^{k}\frac{c_{n-1}}{N^{n-(l+1)}} + \sum_{m=l}^{k}\frac{K_m e_m}{N^{m-l}} + E_{l-1}||g||\right]}.
\end{aligned}
\tag{9.57}
$$

Proof. Proceed like in the proof of Lemma 8.3.13, using the SSDE in (9.4) instead of the SSDE in (8.6). □

We obtain the following useful corollary from Lemma 9.2.2.

Corollary 9.2.3. *For all* $1 \leq l \leq k$, $s > 0$ *and* $\tilde{l} \geq l$,

$$
\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(l-1),\Omega_N^{k+1}}(N^{\tilde{l}-1}s) - \Theta_x^{(l-1),\Omega_N^{k+1}}(N^{\tilde{l}-1}s)\right|\right] = 0. \tag{9.58}
$$

Proof. We proceed by induction. The result for $l = 1$ is trivial. Suppose that the result holds for $l = L$. Then for $l = L + 1$ we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\left|\bar{\Theta}^{(L),\Omega_N^{k+1}}(N^L s) - \Theta_x^{(L),\Omega_N^{k+1}}(N^L s)\right|\right] \\
&\leq \mathbb{E}\left[\left|\bar{\Theta}^{(L),\Omega_N^{k+1}}(N^L s) - \frac{1}{N}\sum_{i=0}^{N-1}\bar{\Theta}_i^{(L-1),\Omega_N^{k+1}}(N^L s)\right|\right] \\
&\quad + \frac{1}{N}\sum_{i=0}^{N-1}\mathbb{E}\left[\left|\bar{\Theta}_i^{(L-1),\Omega_N^{k+1}}(N^L s) - \Theta_{x,i}^{(L-1),\Omega_N^{k+1}}(N^L s)\right|\right].
\end{aligned}
\tag{9.59}
$$

Note that the second term in the right-hand side of (9.59) tends to 0 as $N \to \infty$ by the induction hypothesis. For the first term in the right-hand side of (9.59),

note that

$$\begin{aligned}
&\mathbb{E}\left[\left|\bar{\Theta}^{(L),\Omega_N^{k+1}}(N^L s) - \frac{1}{N}\sum_{i=0}^{N-1}\bar{\Theta}_i^{(L-1),\Omega_N^{k+1}}(N^L s)\right|\right] \\
&= \mathbb{E}\left[\left|\frac{\Theta_x^{(L),\Omega_N^{k+1}}(N^L s) + \sum_{m=0}^{L-1} K_m \Theta_{y_m}^{(L),\Omega_N^{k+1}}(N^L s)}{1+\sum_{m=0}^{L-1} K_m}\right.\right. \\
&\qquad \left.\left. - \frac{\Theta_x^{(L),\Omega_N^{k+1}}(N^L s) + \sum_{m=0}^{L-2} K_m \Theta_{y_m}^{(L),\Omega_N^{k+1}}(N^L s)}{1+\sum_{m=0}^{L-2} K_m}\right|\right] \\
&= \frac{K_{L-1}}{1+\sum_{m=0}^{L-1} K_m}\mathbb{E}\left[\left|\Theta_{y_{L-1}}^{(L),\Omega_N^{k+1}}(N^L s) - \bar{\Theta}^{(L),\Omega_N^{k+1}}(N^L s)\right|\right].
\end{aligned} \tag{9.60}$$

Invoking Lemma 9.2.2 and using the induction hypothesis, we see that for $s > 0$ and $\tilde{l} \geq L$ indeed

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(L),\Omega_N^{k+1}}(N^{\tilde{l}} s) - \Theta_x^{(L),\Omega_N^{k+1}}(N^{\tilde{l}} s)\right|\right] = 0. \tag{9.61}$$

□

6 Show that the convergence in step 4 and step 5 actually holds along each subsequence. Therefore we obtain the limiting evolution of the single colonies, the auxiliary 1-block process and the effective 2-block process. This follows from the fact that the auxiliary k-estimator process converges to the same limit along every subsequence. Consequently, the same holds for the auxiliary $k-1$-estimator process. In this way we can traverse back through the levels to obtain that all l-estimator process converges as $N \to \infty$.

Define, for $0 \leq l \leq k$,

$$\mathfrak{s}_l^{k+1} = [0,1] \times [0,1]^{k+2-l}. \tag{9.62}$$

We obtain, for $0 \leq l \leq k-1$,

$$\begin{aligned}
&\lim_{N\to\infty} \mathcal{L}[(\mathbf{\Theta}^{\mathrm{aux},(l+1),\Omega_N^{k+1}}(\bar{t}))]] = \Gamma_{\mathbf{\Theta}^{(l+2)}}^{(l+1),\mathrm{aux}}, \\
&\Gamma_{\mathbf{\Theta}^{(l+2)}}^{(l+1),\mathrm{aux}} = \int_{\mathfrak{s}_{l+2}^{k+1}} \cdots \int_{\mathfrak{s}_k^{k+1}} \Gamma_{\vartheta_k}^{(k),\mathrm{aux}}(\mathrm{d}u_k)\cdots\Gamma_{u_{l+3}}^{(l+2),\mathrm{aux}}(\mathrm{d}u_{l+2})\Gamma_{u_{l+2}}^{(l+1),\mathrm{aux}}.
\end{aligned} \tag{9.63}$$

Therefore by step 5

$$\lim_{N\to\infty} \mathcal{L}\left[\left(\mathbf{\Theta}^{\mathrm{aux},(l),\Omega_N^{k+1}}(\bar{t}+N_n^l t_l)\right)_{t_l>0,\, i\in[N^{k+1-l}]}\right] = \mathcal{L}\left[(z_{l,\mathbf{\Theta}^{(l+1)}}^{\mathrm{aux}}(t))_{t\geq 0}\right], \tag{9.64}$$

where $\mathbf{\Theta}^{(l+1)} = (\mathbf{\Theta}_x^{(l+1)}, (\mathbf{\Theta}_{y_m,l}^{(l+1)})_{m=l}^{k+1}) \in \mathfrak{s}_l^{(k+1)}$ are random variables with law

$$\mathcal{L}\left[\mathbf{\Theta}^{(l+1)}\right] = \Gamma_{\mathbf{\Theta}^{(l+2)}}^{(l+1),\mathrm{aux}}. \tag{9.65}$$

The initial state of the limiting process in (9.64) is given by

$$\begin{aligned}
&\mathcal{L}\left[z^{\mathrm{aux}}_{l,\boldsymbol{\Theta}^{(l)}}(0)\right] = \Gamma^{(l),\mathrm{aux}}_{\boldsymbol{\Theta}^{(l+1)}},\\
&\qquad \Gamma^{(l),\mathrm{aux}}_{\boldsymbol{\Theta}^{(l+1)}} = \int_{\mathfrak{s}^{k+1}_{l+2}} \cdots \int_{\mathfrak{s}^{k+1}_{k}} \Gamma^{(k),\mathrm{aux}}_{\vartheta+k}(\mathrm{d}u_k)\cdots\Gamma^{(l+2),\mathrm{aux}}_{u_{l+2}}(\mathrm{d}u_{l+1})\Gamma^{(l+1),\mathrm{aux}}_{u_{l+1}}
\end{aligned} \tag{9.66}$$

and $(z^{\mathrm{aux}}_{l,\boldsymbol{\Theta}^{(l+1)}}(t))_{t\geq 0}$ is the process evolving according to (9.14) with θ, $(y_{m,l})_{m=l+1}^{k+1}$ replaced by the random variables $\boldsymbol{\Theta}^{(l+1)}_x, (\boldsymbol{\Theta}^{(l+1)}_{y_m})_{m=l+1}^{k+1}$. Recall that, by Lemma 9.2.1, we have, for $l+1\leq m\leq k+1$,

$$\boldsymbol{\Theta}^{(l+1)}_{y_m} = \boldsymbol{\Theta}^{(m)}_{y_m}. \tag{9.67}$$

7 Use the Meyer-Zheng topology to obtain Proposition 9.1.1(b).

Proof. Note that Lemma 9.2.2 and Corollary 9.2.3 together imply

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(l),\Omega^{k+1}_N}(N^l t_l) - \Theta^{(l),\Omega^{k+1}_N}_x(N^l t_l)\right|\right] = 0,\\
&\lim_{N\to\infty} \mathbb{E}\left[\left|\bar{\Theta}^{(l),\Omega^{k+1}_N}(N^l t_l) - \Theta^{(l),\Omega^{k+1}_N}_{y_m}(N^l t_l)\right|\right] = 0, \text{ for } 0\leq m\leq l-1.
\end{aligned} \tag{9.68}$$

Combining the result obtained in step 6 with the proof strategy followed in Section 8.3.10, we get the claim. □

8 Finally, we prove Proposition 9.1.1(a).

Proof. Step 6 and step 7 yield the laws of the components $\mathcal{L}[M^k_l]$ of the interaction chain $(M^k_{-l})^0_{-l=-(k+1)}$. Note that the state space $([0,1]\times[0,1]^{k+2})^{k+2}$ is compact, and therefore the sequence of random variables

$$\left(\left(\boldsymbol{\Theta}^{(l),\Omega^{k+1}_N}(\bar{t})\right)_{l=k+1,k,\dots,0}\right)_{N\in\mathbb{N}} \tag{9.69}$$

is tight. For any

$$\begin{aligned}
&f\colon ([0,1]\times[0,1]^{k+2})^{n+2}\to\mathbb{R},\\
&f(x) = \prod_{i=1}^n f_i(x_i),\\
&f_i\in\mathcal{C}_b([0,1]\to\mathbb{R}),
\end{aligned} \tag{9.70}$$

we can use conditioning on the previous block average to obtain

$$\lim_{N\to\infty} \mathbb{E}\left[f\left(\left(\boldsymbol{\Theta}^{(l),\Omega^{k+1}_N}(\bar{t})\right)_{l=k+1,k,\dots,0}\right)\right] = \mathbb{E}\left[f\left((M^k_{-l})^0_{-l=-(k+1)}\right)\right]. \tag{9.71}$$

Using that the set of functions of the form (9.70) is separating, we obtain the claim. □

§9.3 Proof: of the hierarchical multi-scale limit theorems.

In this section we prove Theorems 4.4.2 and 4.4.4. We start by proving Theorem 4.4.4. Theorem 4.4.2 will follow from Theorem 4.4.4 by projection onto the effective components.

Proof of Theorem 4.4.2

Proof. Recall the estimators in (4.70). Like for the finite-level hierarchical mean-field system, we can define the auxiliary estimator process by

$$\boldsymbol{\Theta}^{(l),\mathrm{aux},\Omega_N}(t) = \big(\bar{\Theta}^{(l),\Omega_N}(t), \big(\Theta_{y_m}^{(l),\Omega_N}(t)\big)_{m=l}^{\infty}\big). \tag{9.72}$$

For $l,k \in \mathbb{N}$ the processes $(\boldsymbol{\Theta}^{(l),\mathrm{aux},\Omega_N}(\bar{t}+N^k t))_{t>0}$ evolve according to (recall, (4.114))

$$\begin{aligned}
\mathrm{d}\bar{\Theta}^{(l),\Omega_N}(N^k t) &= E_l \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1-k}} \Big[\Theta_x^{(n),\Omega_N}(N^k t) - \Theta_x^{(l),\Omega_N}(N^k t)\Big]\,\mathrm{d}t \\
&\quad + E_l \sqrt{\frac{N^k}{N^{2l}} \sum_{\xi\in B_l} g\big(x_\xi(N^k t)\big)}\,\mathrm{d}w(t) \\
&\quad + E_l \sum_{m=l}^{\infty} \frac{K_m e_m}{N^{m-k}} \Big[\Theta_{y_m}^{(l),\Omega_N}(N^k t) - \Theta_x^{(l),\Omega_N}(N^k t)\Big]\,\mathrm{d}t,
\end{aligned} \tag{9.73}$$

$$\mathrm{d}\Theta_{y_m}^{(l),\Omega_N}(N^k t) = \frac{e_m}{N^{m-k}} \Big[\Theta_x^{(l),\Omega_N}(N^k t) - \Theta_{y_m}^{(l),\Omega_N}(N^k t)\Big]\,\mathrm{d}t, \qquad l \le m \le \infty.$$

Therefore, for $l > k$ and all $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\Bigg[\sup_{0\le t\le L(N)} \Big|\bar{\Theta}^{(l),\Omega_N}(\bar{t}) - \bar{\Theta}^{(l),\Omega_N}(\bar{t}+N^k t)\Big| > \epsilon\Bigg] \\
&= \mathbb{P}\Bigg[\sup_{0\le t\le L(N)} E_l \Bigg| \int_{\bar{t}}^{\bar{t}+N^k t} \mathrm{d}r \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1}} \Big[\Theta_x^{(n),\Omega_N}(r) - \Theta_x^{(l),\Omega_N}(r)\Big] \\
&\qquad + \int_{\bar{t}}^{\bar{t}+N^k t} \mathrm{d}r \sum_{m=l}^{\infty} \frac{K_m e_m}{N^m} \Big[\Theta_{y_m}^{(l),\Omega_N}(r) - \Theta_x^{(l),\Omega_N}(r)\Big] \\
&\qquad + \int_{\bar{t}}^{\bar{t}+N^k t} \mathrm{d}w_i(r) \sqrt{\frac{1}{N^{2l}} \sum_{\xi\in B_l} g\big(x_\xi(r)\big)}\,\Bigg| > \epsilon\Bigg] \\
&\le \mathbb{P}\Bigg[\sup_{0\le t\le L(N)} E_l \Bigg| \int_{\bar{t}}^{\bar{t}+N^k t} \mathrm{d}w_i(r) \sqrt{\frac{1}{N^{2l}} \sum_{\xi\in B_l} g\big(x_\xi(r)\big)}\,\Bigg| \\
&\qquad > \epsilon - t\Bigg[\sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1-k}} + \sum_{m=l}^{\infty} \frac{K_m e_m}{N^{m-k}}\Bigg]\Bigg].
\end{aligned} \tag{9.74}$$

Note that, since $l > k$,

$$\lim_{N\to\infty} t\left[\sum_{n=l+1}^{\infty}\frac{c_{n-1}}{N^{n-1-k}} + \sum_{m=l}^{\infty}\frac{K_m e_m}{N^{m-k}}\right] = 0. \tag{9.75}$$

Hence, like in the proof of Lemma 8.3.4, we can use an optional stopping argument to obtain

$$\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\bar{\Theta}^{(l),\Omega_N}(\bar{t}) - \bar{\Theta}^{(l),\Omega_N}(\bar{t}+N^k t)\right| = 0 \qquad \text{in probability.} \tag{9.76}$$

Using a similar computation as in (9.74), we can show

$$\lim_{N\to\infty}\sup_{0\le t\le L(N)}\left|\bar{\Theta}_{y_m}^{(l),\Omega_N}(\bar{t}) - \Theta_{y_m}^{(l),\Omega_N}(\bar{t}+N^k t)\right| = 0 \qquad \text{in probability.} \tag{9.77}$$

Hence we obtain that, on time scale N^k as $N \to \infty$, the process $(\boldsymbol{\Theta}^{(l),\mathrm{aux},\Omega_N}(\bar{t} + N^k t))_{t>0}$ does not evolve and therefore is still in its initial state$(\boldsymbol{\Theta}^{(l),\mathrm{aux},\Omega_N}(\bar{t}))$.

Using that the l-auxiliary estimator processes do not move for $l > k$, they function like the "outside world" for the finite-level mean-field system in Section 9.1. Therefore we can proceed as in the proof of Proposition 9.1.1 to prove the second and third line in (4.88) in Theorem 4.4.2. The l-block estimator process $(\boldsymbol{\Theta}^{(l),\mathrm{aux},\Omega_N}(\bar{t} + N^l t))_{t>0}$ evolves according to (9.73) with $l = k$. Note that the extra interactions due to migration over larger blocks $l > k$ and exchange with deeper seed-banks $m > k$ in (4.126) are of order $\mathcal{O}(1/N)$. Therefore these terms vanish as $N \to \infty$, and we can just proceed as in the scheme of Section 9.2, to obtain the second and third line in (4.88) in Theorem 4.4.2. Using these results, we obtain that, for $l > k$,

$$\boldsymbol{\Theta}^{(l),\mathrm{aux},\Omega_N}(\bar{t}) = \delta_{M^k_{-(k+1)}}. \tag{9.78}$$

□

CHAPTER 9

CHAPTER 10

Orbit of the renormalisation transformation

In this chapter we analyse the orbit of the renormalisation transformation and show that it has the Fisher-Wright diffusion as a global attractor. In Section 10.1 we write down moment relations for the equilibrium defined in (4.67) for single colonies (Proposition 10.1.1) and for block averages (Proposition 10.1.2). In Section 10.2 we derive the iterates of these moment relations for single colonies (Proposition 10.2.1) and for blocks (Proposition 10.2.2). In Section 10.3 we prove clustering (Propositions 10.3.1–10.3.2). In Section 10.4 we prove Theorems 4.5.1 and 4.5.3, and work out the dichotomy of a finite seed-bank ($\rho < \infty$) versus infinite seed-bank ($\rho = \infty$).

§10.1 Moment relations

We use Itô-calculus to compute the mixed moments. Recall $\theta_x, (\theta_{y_m})$ as defined in (4.21), ϑ_k as defined in (4.62) and $\bar{\vartheta}^{(k)}$ (4.135). Also recall E_k as defined in (4.64). Abbreviate

$$A_0^n = \frac{1}{2}\sum_{k=0}^{n} \frac{E_k}{c_k}\frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}, \qquad n \in \mathbb{N}, \tag{10.1}$$

and

$$B_0 = \frac{1}{2}\frac{E_0^2}{(E_0 c_0 + e_0) + E_0 K_0 e_0}. \tag{10.2}$$

In the following proposition, the first five equations are first and second moment relations, while the last equation is the definition of the renormalisation transformation. Later we will see that this set of equations can be iterated.

Proposition 10.1.1 (Moment relations: single colonies).
Let ϑ_0 be as defined in (4.62), *and let $\Gamma^{(0)}_{(\vartheta_0,y_l)} = \Gamma^{g,c_0,E_0,K_0,e_0}_{(\vartheta_0,y_l)}$ be the equilibrium of* (4.67) *measure defined in* (4.73) *with $k = 0$, with $g \in \mathcal{G}$, $c_0 \in (0,\infty)$, $E_0 \in [0,1]$ and*

$K_0, e_0 \in (0, \infty)$. *Then the following moment relations hold:*

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0 \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0) = \vartheta_0, \tag{10.3}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0} \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0) = \vartheta_0, \tag{10.4}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0 y_{0,0} \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0) = \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}^2 \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0), \tag{10.5}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2 \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0) = \vartheta_0^2 + A_0^0(\mathcal{F}g)(\vartheta_0), \tag{10.6}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}^2 \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0) = \vartheta_0^2 + (A_0^0 - B_0)(\mathcal{F}g)(\vartheta_0), \tag{10.7}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} g(x_0) \, \Gamma^{g,E_0,c_0,K_0,e_0}_{(\vartheta_0,y_l)}(\mathrm{d}z_0) = (\mathcal{F}g)(\vartheta_0). \tag{10.8}$$

Proof. For ease of notation we write x, y_0 instead of $x_0, y_{0,0}$ for the single colonies. We use Itô's formula to compute the first and second moments, and invoke the equilibrium condition to get the above formulas, except for the last formula, which is the definition of $\mathcal{F}$ in (4.75).

1. We begin with the first moments of x and y_0. For $k = 0$, (4.67) becomes

$$\begin{aligned} \mathrm{d}x(t) &= E_0 \Big[c_0 \big[\vartheta_0 - x(t)\big]\, \mathrm{d}t + \sqrt{g(x(t))}\, \mathrm{d}w_0(t) \\ &\qquad + K_0 e_0 \left[y_0(t) - x(t)\right] \mathrm{d}t \right], \end{aligned} \tag{10.9}$$

$$\mathrm{d}y_0(t) = e_0 \left[x(t) - y_0(t)\right] \mathrm{d}t, \tag{10.10}$$

$$\mathrm{d}y_m(t) = 0.$$

In equilibrium the distribution of $x(t)$ is constant in time, and so $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[x(t)] = 0$, where $\mathbb{E}$ denotes expectation w.r.t. $\Gamma^{g,c_0,E_0,K_0,e_0}_{\theta}$. Integrating (10.9) and taking the expectation, we get

$$\begin{aligned} \mathbb{E}[x(t) - x(0)] &= E_0 \left[\mathbb{E}\left[\int_0^t \mathrm{d}s\, c_0 \big[\vartheta_0 - x(s)\big] + K_0 e_0 \int_0^t \mathrm{d}s \left[y_0(s) - x(s)\right] \right]\right], \\ &= E_0 \left[\int_0^t \mathrm{d}s\, \mathbb{E}\Big[c_0 \big[\vartheta_0 - x(s)\big] + K_0 e_0 \big[y_0(s) - x(s)\big]\Big]\right], \end{aligned} \tag{10.11}$$

where in the second equation we use Fubini. Turning back to differential notation, we see from (10.11) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\, \mathbb{E}[x(0)] = 0 = E_0 \left\{ \mathbb{E}\Big[c_0 \big[\vartheta_0 - x(t)\big] + K_0 e_0 \big[y_0(t) - x(t)\big]\Big]\right\}, \tag{10.12}$$

and it follows that

$$\mathbb{E}\Big[c_0 \big[\vartheta_0 - x\big] + K_0 e_0 \big[y_0 - x\big]\Big] = 0. \tag{10.13}$$

In the same way it follows from (10.10) that

$$\mathbb{E}\Big[e_0 \big[x - y_0\big]\Big] = 0. \tag{10.14}$$

Therefore we obtain from (10.13)–(10.14) that

$$\mathbb{E}[x] = \mathbb{E}[y_0] = \vartheta_0. \tag{10.15}$$

2. We next compute the second moments. By Itô's formula,

$$\begin{aligned} \mathrm{d}(x(t))^2 &= 2x(t)\,\mathrm{d}x(t) + (\mathrm{d}x(t))^2 \qquad (10.16)\\ &= 2c_0 x(t) E_0[\vartheta_0 - x(t)]\,\mathrm{d}t + 2x(t)E_0\sqrt{g(x(t))}\,\mathrm{d}w_0(t)\\ &\quad + E_0\left[2K_0e_0\,x(t)y_0(t) - 2K_0e_0\,x_t^2\right]\mathrm{d}t + E_0^2 g(x(t))\,\mathrm{d}t. \end{aligned}$$

Taking expectations and using that we are in equilibrium, we get

$$0 = 2c_0\vartheta_0^2 - 2c_0\mathbb{E}[x^2] + 2K_0e_0\mathbb{E}[xy_0] - 2K_0e_0\mathbb{E}[x^2] + E_0\mathbb{E}[g(x)]. \tag{10.17}$$

Using $\mathbb{E}[g(x)] = (\mathcal{F}g)(\vartheta_0)$, we find

$$\mathbb{E}[x^2] = \frac{c_0}{(c_0 + K_0e_0)}\vartheta_0^2 + \frac{K_0e_0}{(c_0+K_0e_0)}\mathbb{E}[xy_0] + \frac{E_0}{2(c_0+K_0e_0)}(\mathcal{F}g)(\vartheta_0). \tag{10.18}$$

In the same way we find

$$\mathbb{E}[y_0^2] = \mathbb{E}[xy_0], \tag{10.19}$$

and for the mixed second moment

$$\mathbb{E}[xy_0] = \frac{E_0c_0}{(E_0c_0+e_0)}\vartheta_0^2 + \frac{e_0}{(E_0c_0+e_0)}\mathbb{E}[x^2]. \tag{10.20}$$

Substituting (10.20) into (10.18), we find $\mathbb{E}[x^2]$ and hence also $\mathbb{E}[y_0^2]$ and $\mathbb{E}[x_0y_0]$. This finishes the proof of Proposition 10.1.1. □

Similar moment relations can be derived for the equilibrium measures of the block averages. Define

$$A_m^n = \frac{1}{2}\sum_{k=m}^{n} \frac{E_k}{c_k}\,\frac{(E_kc_k+e_k)}{(E_kc_k+e_k)+E_kK_ke_k}, \qquad m \in \mathbb{N}_0,\ n \in \mathbb{N}, \tag{10.21}$$

and

$$B_m = \frac{1}{2}\,\frac{E_m^2}{(E_mc_m+e_m)+E_mK_me_m}, \qquad m \in \mathbb{N}_0. \tag{10.22}$$

Recall the definition of $\mathcal{F}^{(n)}$ in (4.76).

Proposition 10.1.2 (Moment relations: blocks).
Let ϑ_m be as defined in (4.62), and let $\Gamma^{(m)}_{(\vartheta_m,y_m)} = \Gamma^{\mathcal{F}^{(m)}g,c_m,E_m,K_m,e_m}_{(\vartheta_m,y_m)}$ be the equilibrium measure of (4.67) with $k = m$, with $g \in \mathcal{G}$, $c_0 \in (0,\infty)$, $E_0 \in [0,1]$ and

$K_0, e_0 \in (0, \infty)$. *Then the following moment relations hold:*

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_m \, \Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m) = \vartheta_m, \tag{10.23}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{m,m} \, \Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m) = \vartheta_m, \tag{10.24}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_m y_{m,m} \, \Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m) \tag{10.25}$$
$$= \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y^2_{m,m} \Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m),$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x^2_m \, \Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m) \tag{10.26}$$
$$= \vartheta_m^2 + A^m_m(\mathcal{F}^{(m+1)}g)(\vartheta_m), \tag{10.27}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y^2_{m,m} \, \Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m) \tag{10.28}$$
$$= \vartheta_m^2 + (A^m_m - B_m)(\mathcal{F}^{(m+1)}g)(\vartheta_m),$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (\mathcal{F}^{(m)}g)(x_m) \, \mathrm{d}\Gamma^{\mathcal{F}^{(m)}g, E_m, c_m, K_m, e_m}_{(\vartheta_m, y_m)}(\mathrm{d}z_m) \tag{10.29}$$
$$= (\mathcal{F}^{(m+1)}g)(\vartheta_m). \tag{10.30}$$

Proof. The proof follows the same line of argument as the proof of Proposition 10.1.1. □

§10.2 Iterate moment relations

Recall the kernels defined in (4.79), the iterates of the kernels defined in (4.134) and $\bar{\vartheta}^{(n)}$. Recall that $Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0)$ is the probability density to see the population of a single colony in state z_0 given that the $(n+1)$-block averages equal $\bar{\vartheta}^{(n)}$.

Proposition 10.2.1 (Iterated moment relations: single components). *For* $n \in \mathbb{N}_0$*,*

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0 \, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \vartheta_n, \tag{10.31}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0} \, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \vartheta_n, \tag{10.32}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2 \, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \vartheta_n^2 + A^n_0 \, (\mathcal{F}^{(n+1)}g)(\vartheta_n), \tag{10.33}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}^2 \, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \vartheta_n^2 + (A^n_0 - B_0)(\mathcal{F}^{(n+1)}g)(\vartheta_n), \tag{10.34}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0 y_{0,0} \, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}^2 \, Q^{(n)}(\vartheta_n, \mathrm{d}z_0), \tag{10.35}$$

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} g(x) \, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = (\mathcal{F}^{(n+1)}g)(\vartheta_n). \tag{10.36}$$

Proof. We prove the claim for x_0^2 only. The other relations follow in a similar way. The proof proceeds by induction. The result for $n = 0$ follows directly from Proposition 10.1.1. Assume the result holds true for $n = n$, for $n = n + 1$ we write

$$\begin{aligned}
&\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\, Q^{(n+1)}(\bar{\vartheta}^{(n+1)}, \mathrm{d}z_0) \qquad (10.37)\\
&= \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\, (Q^{[n+1]} \circ Q^{(n)})(\bar{\vartheta}^{(n+1)}, \mathrm{d}z_0)\\
&= \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\, Q^{[n+1]}(\bar{\vartheta}^{(n+1)}, \mathrm{d}z_{n+1})\, Q^{(n)}(z_{n+1}, \mathrm{d}z_0)\\
&= \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} \left[\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\, Q^{(n)}(z_{n+1}, \mathrm{d}z_0)\right] Q^{[n+1]}(\bar{\vartheta}^{(n+1)}, \mathrm{d}z_{n+1})\\
&= \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} \left[x_{n+1}^2 + A_0^n\, (\mathcal{F}^{(n+1)}g)(x_{n+1})\right] \Gamma_{\bar{\vartheta}^{(n+1)}}(\mathrm{d}z_{n+1})\\
&= \vartheta_{n+1}^2 + A_0^n\, (\mathcal{F}^{(n+2)}g)(\vartheta_{n+1}) + A_{n+1}^{n+1}\, (\mathcal{F}^{(n+2)}g)(\vartheta_{n+1})\\
&= \theta_{n+1}^2 + A_0^{n+1}\, (\mathcal{F}^{(n+2)}g)(\vartheta_{n+1}).
\end{aligned}$$

The first and second equality use the definition in (4.134), the third equality uses Fubini, the fourth equality is the induction step, the fifth equality uses Proposition 10.1.2, in particular, (10.26) and (10.29). □

Similar iterate moment relations hold for blocks. Define, for $m, n \in \mathbb{N}_0$ with $n \geq m$,

$$Q_m^{(n)} = Q^{[n]} \circ \cdots \circ Q^{[m]}. \qquad (10.38)$$

Proposition 10.2.2 (Iterated moment relations: blocks of components). *For $n, m \in \mathbb{N}_0$ with $n \geq m$,*

$$\begin{aligned}
\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_m\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m) &= \vartheta_n, & (10.39)\\
\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{m,m}\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m) &= \vartheta_n, & (10.40)\\
\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_m^2\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m) &= \vartheta_n^2 + A_m^n\, (\mathcal{F}^{(n+1)}g)(\vartheta_n), & (10.41)\\
\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{m,m}^2\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m) &= \vartheta_n^2 + (A_m^n - B_m)(\mathcal{F}^{(n+1)}g)(\vartheta_n), & (10.42)\\
\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_m y_{m,m}\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m) &= \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{m,m}^2\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m), & (10.43)\\
\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (\mathcal{F}^{(m)}g)(x_m)\, Q_m^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_m) &= (\mathcal{F}^{(n+1)}g)(\vartheta_n). & (10.44)
\end{aligned}$$

Proof. Follow a similar induction argument as in the proof of Proposition 10.2.1. □

§10.3 Clustering

To prove Theorem 4.5.1, we proceed as in [5]. The following clustering property holds for the kernels associated with single colonies.

Proposition 10.3.1 (Clustering: single colonies). *Assume*

$$\lim_{n\to\infty} \vartheta_n = \theta. \tag{10.45}$$

Then

$$\begin{aligned}\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, \{(x_0, y_{(0,0)}) = (0,0)\}\big) &= 1-\theta,\\ \lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, \{(x_0, y_{(0,0)}) = (1,1)\}\big) &= \theta,\end{aligned} \tag{10.46}$$

if and only if

$$\lim_{n\to\infty} A_0^n = \infty. \tag{10.47}$$

Consequently,

$$\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, \{(x_0, y_{0,0})\} \notin \{(0,0),(1,1)\}\big) = 0. \tag{10.48}$$

Proof. The proof exploits the iterated moment relations. First assume (10.47)

1. By Proposition 10.2.1

$$\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0(1-x_0)\, Q^{(n)}(\bar\vartheta^{(n)}, \mathrm{d}z_0) = \vartheta_n(1-\vartheta_n) - A_0^n(\mathcal{F}^{(n+1)}g)(\vartheta_n). \tag{10.49}$$

Because $x_0(1-x_0) \geq 0$ for $x \in [0,1]$, we have

$$\vartheta_n(1-\vartheta_n) \geq A_0^n(\mathcal{F}^{(n+1)}g)(\vartheta_n) \qquad \forall\, n \in \mathbb{N}. \tag{10.50}$$

Since $\lim_{n\to\infty} A_0^n = \infty$, it follows that $\lim_{n\to\infty}(\mathcal{F}^{(n+1)}g)(\vartheta_n) = 0$. On the other hand,

$$(\mathcal{F}^{(n+1)}g)(\vartheta_n) = \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} g(x_0)\, Q^{(n)}(\bar\vartheta^{(n)}, \mathrm{d}z_0) \tag{10.51}$$

and, because $g(x) > 0$ for $x \in (0,1)$, $Q^{(n)}(\vartheta_n, \mathrm{d}z_0)$ puts all its mass on $x_0 = 0$ and $x_0 = 1$ in the limit as $n \to \infty$. Let

$$Q^{(n)}\big(\bar\vartheta^{(n)}, \{x_0 = 0 \text{ or } x_0 = 1\}\big) = \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} \mathbf{1}_{\{x_0=0 \text{ or } x_0=1\}}(z_0)\, Q^{(n)}(\bar\vartheta^{(n)}, \mathrm{d}z_0), \tag{10.52}$$

then

$$\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, \{x_0 = 0 \text{ or } x_1 = 1\}\big) = 1. \tag{10.53}$$

The first moment of x_0 converges to (recall (4.63))

$$\lim_{n\to\infty} \int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0\, Q^{(n)}(\bar\vartheta^{(n)}, \mathrm{d}z_0) = \lim_{n\to\infty} \vartheta_n = \theta. \tag{10.54}$$

Hence

$$\begin{aligned}\lim_{n\to\infty} Q^{(n)}(\bar\vartheta^{(n)}, \{x_0 = 0\}) &= 1-\theta,\\ \lim_{n\to\infty} Q^{(n)}(\bar\vartheta^{(n)}, \{x_0 = 1\}) &= \theta,\end{aligned} \tag{10.55}$$

and

$$\begin{aligned}
&\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0)\\
&\quad = \lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\Big(\mathbf{1}_{\{1\}}(x_0)+\mathbf{1}_{\{0\}}(x_0)+\mathbf{1}_{\{(0,1)\}}(x_0)\Big)\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \theta.
\end{aligned} \tag{10.56}$$

On the other hand,

$$\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0^2\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \theta^2 + \lim_{n\to\infty} A_0^n(\mathcal{F}^{(n+1)}g)(\theta), \tag{10.57}$$

and so, combining (10.56)–(10.57), we obtain

$$\lim_{n\to\infty} A_0^n(\mathcal{F}^{(n+1)}g)(\theta) = \theta(1-\theta). \tag{10.58}$$

2. We know also that

$$\begin{aligned}
&\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0 y_{0,0}\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) \\
&\quad = \lim_{n\to\infty} \vartheta_n^2 + \left(A_0^n - \frac{E_0^2}{E_0c_0+e_0+E_0K_0e_0}\right)(\mathcal{F}^{(n+1)}g)(\vartheta_n)\\
&\quad = \theta^2+\theta(1-\theta) = \theta
\end{aligned} \tag{10.59}$$

and

$$\begin{aligned}
&\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} xy_0\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0)\\
&\quad = \lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} x_0y_{0,0}\Big(\mathbf{1}_{\{1\}}(x_0)+\mathbf{1}_{\{0\}}(x_0)+\mathbf{1}_{\{(0,1)\}}(x_0)\Big)\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0)\\
&\quad = \lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}\,\mathbf{1}_{\{1\}}(x_0)\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0).
\end{aligned} \tag{10.60}$$

Therefore

$$\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}\,\mathbf{1}_{\{1\}}(x_0)\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \lim_{n\to\infty}\vartheta_n = \theta, \tag{10.61}$$

and hence

$$\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (1-y_{(0,0)})\,\mathbf{1}_{\{1\}}(x_0)\, Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = \theta-\theta = 0. \tag{10.62}$$

Since $1-y_{(0,0)} \geq 0$, we conclude that if $x_0 = 1$, then $Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0)$ puts all its mass on $y_{0,0} = 1$ in the limit as $n\to\infty$. Hence

$$\lim_{n\to\infty} Q^{(n)}\big(\bar{\vartheta}^{(n)}, \{(x_0, y_{0,0}) = (1,1)\}\big) = \theta. \tag{10.63}$$

From Proposition 10.2.2 it also follows that

$$\begin{aligned}
&\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (1-x_0)(1-y_{0,0})\, Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) \\
&\quad= 1-\theta-\theta+\theta^2+\lim_{n\to\infty}\left(A_0^n-\frac{E_0^2}{E_0c_0+e_0+E_0K_0e_0}\right)(\mathcal{F}^{(n+1)}g)(\vartheta_n) \\
&\quad= 1-\theta.
\end{aligned} \tag{10.64}$$

On the other hand,

$$\begin{aligned}
&\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (1-x_0)(1-y_{0,0})\, Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) \\
&= \lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (1-x_0)(1-y_{0,0})\, \Big(\mathbf{1}_{\{1\}}(x_0)+\mathbf{1}_{\{0\}}(x_0)+\mathbf{1}_{\{(0,1)\}}(x_0)\Big) \\
&\qquad\qquad \times Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) \\
&= \lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} (1-y_{0,0})\,\mathbf{1}_{\{0\}}(x_0)\, Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) \\
&= 1-\theta.
\end{aligned} \tag{10.65}$$

Since $y\in[0,1]$ and

$$\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} \mathbf{1}_{\{0\}}(x_0)\, Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) = 1-\theta, \tag{10.66}$$

it follows that

$$\lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} y_{0,0}\mathbf{1}_{\{0\}}(x_0)\, Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) = 0. \tag{10.67}$$

This implies that if $x_0=0$, then $Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0)$ puts all its mass on $y_{0,0}=0$ in the limit as $n\to\infty$. Hence

$$\begin{aligned}
&\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)},\{(x_0,y_{0,0})=(0,0)\}\big) = 1-\theta,\\
&\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)},\{(x_0,y_{0,0})=(1,1)\}\big) = \theta.
\end{aligned} \tag{10.68}$$

Now assume (10.46). Then (10.56) still holds. On the other hand, also (10.57) still holds by Proposition 10.2.1. Therefore we obtain (10.58). On the other hand, by (10.46)

$$\lim_{n\to\infty}\mathcal{F}^{(n+1)}g = \lim_{n\to\infty}\int_{[0,1]\times[0,1]^{\mathbb{N}_0}} g(x_0)Q^{(n)}(\bar\vartheta^{(n)},\mathrm{d}z_0) = 0. \tag{10.69}$$

Hence (10.47) holds. □

A similar clustering property holds for the kernels associated with blocks.

Proposition 10.3.2 (Clustering: blocks). *Assume*

$$\lim_{n\to\infty} \vartheta_n = \theta. \tag{10.70}$$

Then

$$\begin{aligned} &\lim_{n\to\infty} Q_m^{(n)}\big(\bar\vartheta^{(n)}, \{(x_m, y_{(m,m)}) = (0,0)\}\big) = 1-\theta,\\ &\lim_{n\to\infty} Q_m^{(n)}\big(\bar\vartheta^{(n)}, \{(x_m, y_{(m,m)}) = (1,1)\}\big) = \theta, \end{aligned} \tag{10.71}$$

if and only if

$$\lim_{n\to\infty} A_m^n = \infty. \tag{10.72}$$

Consequently,

$$\lim_{n\to\infty} Q_m^{(n)}\big(\bar\vartheta^{(n)}, \{(x_m, y_{m,m})\} \notin \{(0,0),(1,1)\}\big) = 0. \tag{10.73}$$

Proof. We can proceed exactly as in the proof of Proposition 10.3.1. □

Finally, we can prove Theorem 4.5.1.

Proof. Note that (10.47) implies (10.72). Recall that single colonies of deep seed-banks that have already interacted and reached their quasi-equilibrium equal the block average of the level on which they interact (see Theorem 4.4.4). It follows that, for $m \in \mathbb{N}_0$,

$$\begin{aligned} &\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, (x_0, y_{0,m}) = (1,1)\big) = \theta,\\ &\lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, (x_0, y_{0,m}) = (0,0)\big) = 1-\theta. \end{aligned} \tag{10.74}$$

Therefore, for $N \in \mathbb{N}_0$,

$$\begin{aligned} &\lim_{n\to\infty} Q^{(n)}\left(\bar\vartheta^{(n)}, \bigcap_{m=0}^{N} \{(x_0, y_{0,m}) = (1,1) \text{ or } (x_0, y_{0,m}) = (0,0)\}\right)\\ &\quad= 1 - \lim_{n\to\infty} Q^{(n)}\left(\bar\vartheta^{(n)}, \bigcup_{m=0}^{N} \{(x_0, y_{0,m}) \in [0,1]^2\backslash\{(0,0),(1,1)\}\}\right)\\ &\quad\geq 1 - \lim_{n\to\infty} \sum_{m=0}^{N} Q^{(n)}\Big(\bar\vartheta^{(n)}, \big\{(x_0, y_{0,m}) \in [0,1]^2\backslash\{(0,0),(1,1)\}\big\}\Big) = 1-0 = 1. \end{aligned} \tag{10.75}$$

Note that

$$\begin{aligned} &\lim_{n\to\infty} Q^{(n)}\left(\bar\vartheta^{(n)}, \bigcap_{m=0}^{N} \big\{(x_0, y_{0,m}) = (1,1) \text{ or } (x_0, y_{0,m}) = (0,0)\big\}\right)\\ &\quad= \lim_{n\to\infty} Q^{(n)}\Big(\bar\vartheta^{(n)}, \big\{(x_0, (y_{0,m})_{0\leq m\leq N}) = (0, 0^{N+1})\\ &\qquad\qquad \text{or } (x_0, (y_{0,m})_{0\leq m\leq N}) = (1, 1^{N+1})\big\}\Big) = 1. \end{aligned} \tag{10.76}$$

On the other hand,

$$\begin{aligned} &\lim_{n\to\infty} Q^{(n)}\Big(\bar\vartheta^{(n)}, \{(x_0, (y_{0,m})_{0\leq m\leq N}) = (1, 1^{N+1})\}\Big)\\ &\quad\leq \lim_{n\to\infty} Q^{(n)}\big(\bar\vartheta^{(n)}, \{(x_0, y_{0,0}) = (1,1)\}\big) = \theta \end{aligned} \tag{10.77}$$

and

$$\lim_{n\to\infty} Q^{(n)}\Big(\bar{\vartheta}^{(n)}, \{(x_0,(y_{0,m})_{0\leq m\leq N}) = (0,0^{N+1})\}\Big) \\ \leq \lim_{n\to\infty} Q^{(n)}\big(\bar{\vartheta}^{(n)}, \{(x_0,y_{0,0}) = (0,0)\}\big) = 1-\theta. \tag{10.78}$$

Hence we conclude that

$$\lim_{n\to\infty} Q^{(n)}\Big(\bar{\vartheta}^{(n)}, \{(x_0,(y_{0,m})_{0\leq m\leq N}) = (1,1^{N+1})\}\Big) = \theta \tag{10.79}$$

and

$$\lim_{n\to\infty} Q^{(n)}\Big(\bar{\vartheta}^{(n)}, \{(x_0,(y_{0,m})_{0\leq m\leq N}) = (0,0^{N+1})\}\Big) = 1-\theta. \tag{10.80}$$

We can do the same for all finite-dimensional distributions. Since $[0,1]\times[0,1]^{\mathbb{N}_0}$ is compact, the process $z_0 = (x_0,(y_{0,m})_{m\in\mathbb{N}_0})$ is tight. Therefore, by (10.79)–(10.80) we find for every converging subsequence

$$\lim_{k\to\infty} Q^{(n_k)}\Big(\bar{\vartheta}^{(n_k)},\cdot\Big) = (1-\theta)\,\delta_{(0,0^{\mathbb{N}_0})} + \theta\,\delta_{(1,1^{\mathbb{N}_0})}. \tag{10.81}$$

We conclude that

$$\lim_{n\to\infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \mathrm{d}z_0) = (1-\theta)\,\delta_{(0,0^{\mathbb{N}_0})} + \theta\,\delta_{(1,1^{\mathbb{N}_0})}, \tag{10.82}$$

which is the claim in (4.137). □

§10.4 Dichotomy finite versus infinite seed-bank

In this section we prove Theorem 4.5.3.

Proof. We investigate for what choices of the sequences $\underline{c},\underline{K},\underline{e}$ defined in (4.5) and (4.10) we meet the *clustering criterion* $\lim_{n\to\infty} A_n \to \infty$ in (4.138). Recall from (4.64) and (4.136) that

$$A_n = \frac{1}{2}\sum_{k=0}^{n-1} \frac{E_k}{c_k}\,\frac{(E_kc_k+e_k)}{(E_kc_k+e_k)+E_kK_ke_k}, \qquad E_k = \frac{1}{1+\sum_{m=0}^{k-1}K_m}. \tag{10.83}$$

We distinguish between three regimes as $k\to\infty$:

(a) $E_kc_k+e_k \gg E_kK_ke_k$.

(b) $E_kc_k+e_k \asymp E_kK_ke_k$.

(c) $E_kc_k+e_k \ll E_kK_ke_k$.

These regimes correspond to the following scaling for A_n as $n\to\infty$:

(a) $A_n \sim \frac{1}{2}\sum_{k=0}^{n-1}\frac{E_k}{c_k}$.

(b) $A_n \asymp \sum_{k=0}^{n-1}\frac{E_k}{c_k}$.

(c) $A_n \sim \frac{1}{2}\sum_{k=0}^{n-1}\frac{E_kc_k+e_k}{c_kK_ke_k}$.

Recall from that (4.14) that

$$\rho = \sum_{m\in\mathbb{N}_0} K_m. \tag{10.84}$$

Different behaviour shows up for finite seed-bank ($\rho<\infty$) and infinite seed-bank ($\rho=\infty$).

(I) $\rho < \infty$**.** Note that $k \mapsto E_k$ is non-increasing and converges to $1/(1+\rho) > 0$. Since $E_k \leq 1$, we have $\lim_{k\to\infty} E_k K_k = 0$ and hence we are in regime 1. Therefore

$$A_n \sim \frac{1}{2(1+\rho)} \sum_{k=0}^{n-1} \frac{1}{c_k}, \qquad n \to \infty, \tag{10.85}$$

and clustering occurs if and only if $\sum_{k\in\mathbb{N}_0} \frac{1}{c_k} = \infty$, which is the same criterion as for the system without seed-bank.

(II) $\rho = \infty$**.** We focus on the settings in (4.52) and (4.53), which fall in regimes 1 and 2.

Asymptotically polynomial. Suppose that

$$K_k \sim A k^{-\alpha}, \quad k \to \infty,\ A \in (0,\infty),\ \alpha \in (-\infty, 1). \tag{10.86}$$

Then

$$E_k \sim \frac{1-\alpha}{A} k^{-(1-\alpha)}, \qquad E_k K_k \sim \frac{1-\alpha}{A} k^{-1}, \qquad k \to \infty. \tag{10.87}$$

Hence we are in regime 1. Suppose that

$$c_k \sim F k^{-\phi}, \quad k \to \infty,\ F \in (0,\infty),\ \phi \in \mathbb{R}. \tag{10.88}$$

Then

$$A_n \sim \frac{1-\alpha}{2AF} \sum_{k=1}^{n-1} k^{-1+\alpha+\phi}, \qquad n \to \infty, \tag{10.89}$$

and clustering occurs if and only if $-\phi \leq \alpha < 1$. In this case

$$\begin{aligned} -\phi < \alpha\colon &\quad A_n \sim \frac{1-\alpha}{2AF(\alpha+\phi)}\, n^{\alpha+\phi},\\ -\phi = \alpha\colon &\quad A_n \sim \frac{1-\alpha}{2AF}\, \log n. \end{aligned} \tag{10.90}$$

The case $\alpha = 1$ can be included. Then (10.87) becomes

$$E_k \sim \frac{1}{A\log k}, \qquad E_k K_k \sim \frac{1}{k \log k}, \qquad k \to \infty, \tag{10.91}$$

so that we are again in regime 1. Now (10.89) becomes

$$A_n \sim \tfrac{1}{2AF} \sum_{k=1}^{n-1} \frac{k^{\phi}}{\log k}, \qquad n \to \infty, \tag{10.92}$$

and clustering occurs if and only if $-\phi \leq 1$. In this case

$$\begin{aligned} -\phi < 1\colon &\quad A_n \sim \frac{1}{2AF(1+\phi)} \frac{n^{1+\phi}}{\log n},\\ -\phi = 1\colon &\quad A_n \sim \frac{1}{2AF}\, \log\log n. \end{aligned} \tag{10.93}$$

Pure exponential. Suppose that

$$K_k = K^k, \quad k \in \mathbb{N}_0,\, K \in (1,\infty). \tag{10.94}$$

Then

$$E_k = \frac{1}{1+\sum_{m=0}^{k-1} K^m} = \frac{1}{1+\frac{K^k-1}{K-1}} = \frac{K-1}{K^k+K-2}. \tag{10.95}$$

Suppose that

$$e_k = e^k, \quad c_k = c^k, \quad k \in \mathbb{N}_0,\, e,c \in (0,\infty). \tag{10.96}$$

Then

$$E_k c_k + e_k = \frac{K-1}{K^k+K-2} c^k + e^k, \qquad E_k K_k e_k = \frac{K-1}{K^k+K-2} K^k e^k, \tag{10.97}$$

and so

$$E_k c_k + e_k \sim (K-1)\left(\frac{c}{K}\right)^k + e^k, \qquad E_k K_k e_k \sim (K-1)e^k, \qquad k \to \infty. \tag{10.98}$$

For $c \leq Ke$ we are in regime 2, and hence

$$A_n \sim \frac{1}{2}\sum_{k=0}^{n-1} \frac{K-1}{(Kc)^k}\, \frac{(K-1)(\frac{c}{K})^k + e^k}{(K-1)(\frac{c}{K})^k + Ke^k}, \qquad n \to \infty, \tag{10.99}$$

which simplifies to

$$\begin{aligned} c < Ke\colon \quad & A_n \sim \frac{1}{2K}\sum_{k=0}^{n-1} \frac{K-1}{(Kc)^k}, \\ c = Ke\colon \quad & A_n \sim \frac{K-1}{2(2K-1)}\sum_{k=0}^{n-1} \frac{K-1}{(Kc)^k}. \end{aligned} \tag{10.100}$$

Clustering occurs if and only if $Kc \leq 1$. In this case

$$\begin{aligned} Kc < 1\colon \quad & \sum_{k=0}^{n-1} \frac{1}{(Kc)^k} \sim \frac{1}{1-Kc}\,(Kc)^{-(n-1)}, \\ Kc = 1\colon \quad & \sum_{k=0}^{n-1} \frac{1}{(Kc)^k} \sim n. \end{aligned} \tag{10.101}$$

For $c > Ke$, on the other hand, we are in regime 1, and hence

$$A_n \sim \frac{1}{2}\sum_{k=0}^{n-1} \frac{K-1}{(Kc)^k}, \qquad n \to \infty, \tag{10.102}$$

for which we can again use (10.101).

The case $K = 1$ can be included. Then $E_k = 1/k$ and (10.98) becomes

$$E_k c_k + e_k \sim \frac{1}{k} c^k + e^k, \qquad E_k K_k e_k \sim \frac{1}{k} e^k, \qquad k \to \infty, \tag{10.103}$$

we are again in regime 1. Hence

$$A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{kc^k}, \qquad n \to \infty, \tag{10.104}$$

and clustering occurs if and only if $c \leq 1$. In that case

$$\begin{aligned} c < 1: &\quad A_n \sim \frac{1}{2(1-c)} \frac{1}{n} c^{-(n-1)}, \\ c = 1: &\quad A_n \sim \frac{1}{2} \log n. \end{aligned} \tag{10.105}$$

□

In the above computations, only regimes 1 and 2 arise. Regime 3 arises, for instance, when

$$\lim_{k\to\infty} \frac{1}{k} \log K_k = \infty, \qquad \lim_{k\to\infty} K_k e_k / c_k = \infty. \tag{10.106}$$

Indeed, the first condition implies that $E_k \sim 1/K_{k-1}$ and $E_k K_k \gg 1$, while the second implies that $E_k K_k e_k \gg E_k c_k$. There are two subcases:

$$\begin{aligned} K_{k-1} e_k \ll c_k: &\quad A_n \sim \tfrac{1}{2} \sum_{k=0}^{n-1} \frac{1}{K_k K_{k-1} e_k}, \\ K_{k-1} e_k \gg c_k: &\quad A_n \sim \tfrac{1}{2} \sum_{k=0}^{n-1} \frac{1}{K_k c_k}. \end{aligned} \tag{10.107}$$

By picking, for instance, $e_k = 1/K_k K_{k-1}$, we find that $A_n \sim \frac{1}{2} n$ in the first subcase and $A_n \gg n$ in the second subcase. By picking, alternatively, $c_k = 1/K_k$, we find that $A_n \gg n$ in the first subcase and $A_n \sim \frac{1}{2} n$ in the second subcase.

APPENDIX B
Appendix Part II

§B.1 Computation of scaling coefficients

In Appendices B.1.1–B.1.2 we spell out a technical computation for the tail of the wake-up time defined in (4.40)–(4.41) in the two parameter regimes given by (4.52)–(4.53). In Appendix B.1.3 we carry out a computation that is needed in Section 5.1.

§B.1.1 Regularly varying coefficients

In (4.40), note that for large t in the sum over m only small values of e_m/N^m contribute, which means large m. Hence, by the Euler-MacLaurin approximation formula, we have

$$P(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m} \,\mathrm{e}^{-(e_m/N^m)t} \sim \frac{1}{\chi} \int_c^\infty \mathrm{d}m \, K_m \frac{e_m}{N^m} \,\mathrm{e}^{-(e_m/N^m)t}, \tag{B.1}$$

where c is a constant that identifies from which value of m onward terms contribute significantly. Make the change of variable $s = \frac{e_m}{N^m}$. Since $e_m \sim Bm^{-\beta}$ and $K_m \sim Am^{-\alpha}$ as $m \to \infty$, we have

$$s \sim Bm^{-\beta} N^{-m} \tag{B.2}$$

and hence

$$\begin{aligned} &\log s \sim \log B - \beta \log m - m \log N, \\ &\log \frac{1}{s} = m \log N \left(-\frac{B}{m \log N} + \frac{\beta \log m}{m \log N} + 1 \right) = [1 + o(1)]\, m \log N, \end{aligned} \tag{B.3}$$

which gives

$$m = [1 + o(1)]\, \frac{\log(\frac{1}{s})}{\log N}. \tag{B.4}$$

Thus,

$$\frac{1}{s} \frac{\mathrm{d}s}{\mathrm{d}m} = -\log N - \frac{\beta}{m} = -[1 + o(1)]\, \log N, \tag{B.5}$$

which implies

$$\frac{\mathrm{d}s}{\mathrm{d}m} = -[1 + o(1)]\, s \log N, \tag{B.6}$$

so that $s(m)$ is asymptotically decreasing in m, and

$$\frac{\mathrm{d}m}{\mathrm{d}s} = -[1+o(1)]\ (s\log N)^{-1}. \tag{B.7}$$

Note that if $c \le m < \infty$, then asymptotically $0 < m^{-\beta}N^{-m} < c^{-\beta}N^{-c} = C_2$. Doing the substitution, we get

$$\begin{aligned}\mathbb{P}(\tau > t) &\sim \frac{1}{\chi}\int_0^{C_2} \mathrm{d}s\, K_m s\,(s\log N)^{-1}\ \mathrm{e}^{-st} \\ &\sim \frac{1}{\chi}\int_0^{C_2} \mathrm{d}s\, A m^{-\alpha}\,(\log N)^{-1}\ \mathrm{e}^{-st} \\ &\sim \frac{1}{\chi}\int_0^{C_2} \mathrm{d}s\, A\left(\frac{\log(\frac{1}{s})}{\log N}\right)^{-\alpha}\ (\log N)^{-1}\ \mathrm{e}^{-st} \\ &\sim \frac{A}{\chi}\left(\frac{1}{\log N}\right)^{-\alpha+1}\int_0^{C_2}\mathrm{d}s\,\log\left(\tfrac{1}{s}\right)^{-\alpha}\ \mathrm{e}^{-st}.\end{aligned} \tag{B.8}$$

Next, put $st = u$, so $s = \frac{u}{t}$ and $\frac{\mathrm{d}s}{\mathrm{d}u} = \frac{1}{t}$ and $0 < u < tC_2$. Then

$$\mathbb{P}(\tau > t) \sim \frac{A}{\chi}\left(\frac{1}{\log N}\right)^{-\alpha+1}\frac{1}{t}\int_0^{C_2 t}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha}\ \mathrm{e}^{-u}. \tag{B.9}$$

We will show that

$$\frac{A}{\chi}\left(\frac{1}{\log N}\right)^{-\alpha+1}\frac{1}{t}\int_0^{C_2 t}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha}\ \mathrm{e}^{-u} \asymp \frac{A}{\chi}\left(\frac{1}{\log N}\right)^{-\alpha+1}\frac{1}{t}\int_0^{C_2 t}\mathrm{d}u\,\log t^{-\alpha}\ \mathrm{e}^{-u}. \tag{B.10}$$

For $\alpha = 0$ this claim is immediate. For $\alpha \in (-\infty, 0)$, note that $\log\left(\frac{t}{u}\right)^{-\alpha}$ is a decreasing function on $(0, C_2 t)$. Therefore we can reason as follows:

$$\begin{aligned}&\int_0^{C_2 t}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha}\ \mathrm{e}^{-u} \\ &= \int_0^{1}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha}\ \mathrm{e}^{-u} + \int_1^{C_2 t}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha}\ \mathrm{e}^{-u} \\ &\le \int_0^{1}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha} + \int_1^{C_2 t}\mathrm{d}u\,\log t^{-\alpha}\mathrm{e}^{-u} \\ &\le 2^{-\alpha}\int_0^{\frac{1}{t}}\mathrm{d}u\,\log\left(\tfrac{1}{u}\right)^{-\alpha} + 2^{-\alpha}\int_{\frac{1}{t}}^{1}\mathrm{d}u\,\log t^{-\alpha} + \log t^{-\alpha}\left[1-\mathrm{e}^{-1}\right] \\ &\le 2^{-\alpha}\Gamma(-\alpha+1) + 2^{-\alpha}\log t^{-\alpha}\left[1-\tfrac{1}{t}\right] + \log t^{-\alpha}\left[1-\mathrm{e}^{-1}\right] \\ &= \log t^{-\alpha}\left[2^{-\alpha}\tfrac{\Gamma(-\alpha+1)}{\log t^{-\alpha}} + 2^{-\alpha}\left[1-\tfrac{1}{t}\right] + \left[1-\mathrm{e}^{-1}\right]\right] \\ &\asymp \log t^{-\alpha}.\end{aligned} \tag{B.11}$$

For the lower bound, note that

$$\begin{aligned}\int_0^{C_2 t}\mathrm{d}u\,\log\left(\tfrac{t}{u}\right)^{-\alpha}\mathrm{e}^{-u} &\ge \log(t)^{-\alpha}\int_0^1\mathrm{d}u\,\mathrm{e}^{-u} + \log(\tfrac{1}{C_2})^{-\alpha}\int_1^{C_2 t}\mathrm{d}u\,\mathrm{e}^{-u} \\ &= \log t^{-\alpha}\left[1 - e^{-1} + \frac{\log(\frac{1}{C_2})^{-\alpha}}{\log t^{-\alpha}}\right] \asymp \log t^{-\alpha}.\end{aligned} \tag{B.12}$$

For $\alpha \in (0,1]$, note that the function $\log\left(\frac{t}{u}\right)^{-\alpha}$ is increasing in u. For the lower bound estimate

$$\begin{aligned}\int_0^{C_2 t} \mathrm{d}u \, \log\left(\tfrac{t}{u}\right)^{-\alpha} \mathrm{e}^{-u} &\geq \lim_{u\to 0} \log\left(\tfrac{t}{u}\right)^{-\alpha}[1-\mathrm{e}^{-1}] + \log t^{-\alpha}[\mathrm{e}^{-1} - e^{-C_2 t}] \\ &= \log t^{-\alpha}\left[0 + \mathrm{e}^{-1} - \mathrm{e}^{-C_2 t}\right] \asymp \log t^{-\alpha}.\end{aligned} \tag{B.13}$$

For the upper bound estimate

$$\begin{aligned}&\int_0^{C_2 t} \mathrm{d}u \, \log\left(\tfrac{t}{u}\right)^{-\alpha} \mathrm{e}^{-u} \\ &\leq \log t^{-\alpha}[1-\mathrm{e}^{-1}] + \log\left(\frac{t}{\sqrt{C_2 t}}\right)^{-\alpha} \int_1^{\sqrt{C_2 t}} \mathrm{d}u \, e^{-u} + \log\left(\frac{1}{C_2}\right)^{-\alpha} \int_{\sqrt{C_2 t}}^{C_2 t} \mathrm{d}u \, \mathrm{e}^{-u} \\ &= \log t^{-\alpha}[1-\mathrm{e}^{-1}] + \left(\tfrac{1}{2}\right)^{-\alpha} \log\left(\frac{t}{C_2}\right)^{-\alpha}\left[\mathrm{e}^{-1} - \mathrm{e}^{-\sqrt{C_2 t}}\right] \\ &\quad + \log\left(\frac{1}{C_2}\right)^{-\alpha}\left[\mathrm{e}^{-\sqrt{C_2 t}} - \mathrm{e}^{-C_2 t}\right] \\ &= \log t^{-\alpha}\left[1 - \mathrm{e}^{-1} + \left(\tfrac{1}{2}\right)^{-\alpha}\left(\frac{\log t - \log C_2}{\log t}\right)^{-\alpha}\left[\mathrm{e}^{-1} - \mathrm{e}^{-\sqrt{C_2 t}}\right]\right. \\ &\qquad \left. + \log\left(\frac{1}{C_2}\right)^{-\alpha} \frac{\left[\mathrm{e}^{-\sqrt{C_2 t}} - \mathrm{e}^{-C_2 t}\right]}{\log t^{-\alpha}}\right] \asymp \log t^{-\alpha}.\end{aligned} \tag{B.14}$$

§B.1.2 Pure exponential coefficients

In order to satisfy condition in (4.12), we must assume that $Ke < N$. Since $K \geq 1$ for $\rho = \infty$, we also have $e < N$. We again use that for large t only large m contribute to the sum. Hence, again by the Euler-MacLaurin approximation formula, we have

$$P(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m} \mathrm{e}^{-(e_m/N^m)t} \sim \int_M^\infty \mathrm{d}m \, K_m \frac{e_m}{N^m} \mathrm{e}^{-(e_m/N^m)t}. \tag{B.15}$$

Again we put $s = \frac{e^m}{N^m}$. Hence

$$\log s = m \log\left(\frac{e}{N}\right), \qquad m = \frac{\log s}{\log \frac{e}{N}}, \qquad \frac{\mathrm{d}m}{\mathrm{d}s} = \frac{1}{s \log \frac{e}{N}}, \tag{B.16}$$

and

$$K_m \sim K^m \sim \mathrm{e}^{\log s \frac{\log K}{\log \frac{e}{N}}} \sim s^{\frac{\log K}{\log \frac{e}{N}}}. \tag{B.17}$$

Since $s(m)$ is decreasing in m, putting $C = \left(\frac{e}{N}\right)^M$ we obtain

$$\mathbb{P}(\tau > t) \sim \int_0^C \mathrm{d}s \, K_m \frac{s}{s \log \frac{e}{N}} \mathrm{e}^{-st} \sim \frac{1}{\log \frac{e}{N}} \int_0^C \mathrm{d}s \, s^{\frac{\log K}{\log \frac{e}{N}}} \mathrm{e}^{-st}. \tag{B.18}$$

Substitute $u = st$, i.e., $\frac{u}{t} = s$, to get

$$\begin{aligned}\mathbb{P}(\tau > t) &\sim \frac{1}{\log \frac{e}{N}} t^{-1-\frac{\log K}{\log \frac{e}{N}}} \int_0^{Ct} \mathrm{d}u \, u^{\frac{\log K}{\log \frac{e}{N}}} \, \mathrm{e}^{-u} \\ &\sim \frac{1}{\log \frac{e}{N}} t^{\frac{-\log(\frac{e}{N})-\log K}{\log \frac{e}{N}}} \int_0^{Ct} \mathrm{d}u \, u^{\frac{\log K}{\log \frac{e}{N}}} \, \mathrm{e}^{-u} \sim \frac{1}{\log \frac{e}{N}} t^{-\frac{\log(\frac{N}{Ke})}{\log \frac{N}{e}}} \int_0^{Ct} \mathrm{d}u \, u^{\frac{\log K}{\log \frac{e}{N}}} \, \mathrm{e}^{-u}.\end{aligned} \tag{B.19}$$

The last integral converges because $\frac{\log K}{\log(\frac{e}{N})} > -1$, and

$$\int_0^{Ct} \mathrm{d}u \, u^{\frac{\log K}{\log \frac{e}{N}}} \, \mathrm{e}^{-u} \leq \int_0^{\infty} \mathrm{d}u \, u^{\frac{\log K}{\log \frac{e}{N}}} \, \mathrm{e}^{-u} = \Gamma\left(\frac{\log K}{\log(\frac{e}{N})} + 1\right). \tag{B.20}$$

§B.1.3 Slowly varying functions

Return to Section 5.1. Note that $t(s) = \varphi(s)^{-1}s^{\gamma}$. Since this is the total time two lineages are active up to time s, $t(s)$ must be smaller than s. By (4.49), we have

$$\frac{\varphi(t)}{\varphi(s)} = \exp\left[-\int_{t(s)}^{s} \frac{\mathrm{d}u}{u}\psi(u)\right]. \tag{B.21}$$

Since we are interested in $s \to \infty$, we may assume that $s \gg 1$ and $t(s) > 1$, and estimate

$$\begin{aligned}\frac{\varphi(t)}{\varphi(s)} &\leq \exp\left[\int_{t(s)}^{s} \frac{\mathrm{d}u}{u}\frac{C}{\log u}\right] = \exp\left[C(\log\log s - \log\log t(s))\right] \\ &= \exp\left[C \log\left(\frac{\log s}{\log\left(\varphi(s)^{-1}s^{\gamma}\right)}\right)\right] = \exp\left[-C \log\left(\frac{\gamma \log s - \log \varphi(s)}{\log s}\right)\right].\end{aligned} \tag{B.22}$$

A similar lower bound holds with the sign reversed. Using that $\lim_{s\to\infty} \frac{\log \varphi(s)}{\log s} = 0$, we get

$$\gamma^{C} \leq \liminf_{s\to\infty} \frac{\varphi(t)}{\varphi(s)} \leq \limsup_{s\to\infty} \frac{\varphi(t)}{\varphi(s)} \leq \gamma^{-C}. \tag{B.23}$$

Both bounds above are positive, so indeed $\frac{\varphi(t)}{\varphi(s)} \asymp 1$.

§B.2 Meyer-Zheng topology

§B.2.1 Basic facts about the Meyer-Zheng topology

In the Meyer-Zheng topology we assign to each real-valued Borel measurable function $(w(t))_{t\geq 0}$ a probability law on $[0,\infty] \times \bar{\mathbb{R}}$ that is called the pseudopath ψ_w. Note that the Borel-σ algebra on $[0,\infty] \times \bar{\mathbb{R}}$ is generated by sets of the form $[a,b] \times B$ for $B \in \mathcal{B}$ and $0 < a < b$. For $A = [a,b] \times B$, set

$$\psi_w(A) = \int 1_A(t, w(t))\mathrm{e}^{-t}\mathrm{d}t = \int_a^b 1_B(w(t))e^{-t}\mathrm{d}t, \tag{B.24}$$

i.e., ψ_w is the image measure of the mapping $t \to (t, w(t))$ under the measure $\lambda(\mathrm{d}t) = \mathrm{e}^{-t}\mathrm{d}t$. The set of all pseudopaths is denoted by Ψ. Note that a pseudopath corresponding to $(w(t))_{t>0}$ is simply its occupation measure. The following important facts are stated in [59]:

- If two paths w_1 and w_2 are the same Lebesgue a.e., then $\psi_{w_1} = \psi_{w_2}$.
- Denote by $\mathbf{D}$ the space of càdlàg paths on $[0,\infty] \times \mathbb{R}$. The mapping $\psi \colon \mathbf{D} \to \Psi$, $w \mapsto \psi_w$ is one-to-one on $\mathbf{D}$ and hence gives an embedding of $\mathbf{D}$ into the compact space $\bar{\mathcal{P}}$, the space of probability measures on $[0,\infty] \times \bar{\mathbb{R}}$.
- Note if f is a function on $[0,\infty] \times \mathbb{R}$ and $w \in \mathbf{D}$, then

  $$\psi_w(f) = \int_0^\infty f(t, w(t))\, \mathrm{e}^{-t}\mathrm{d}t. \tag{B.25}$$

  Therefore we say that the sequence of pseudopaths induced by $(w_n) \subset \mathbf{D}$ converges to a pseudopath w if, for all continuous bounded function $f(t, w(t))$ on $[0,\infty] \times \bar{\mathbb{R}}$,

  $$\lim_{n\to\infty} \int_0^\infty f(t, w_n(t))\, \mathrm{e}^{-t}\mathrm{d}t = \int_0^\infty f(t, w(t))\, \mathrm{e}^{-t}\mathrm{d}t. \tag{B.26}$$

  Since a pseudopath is a measure, convergence of pseudopaths is convergence of measures.
- $\mathbf{D}$ endowed with the pseudopath topology is *not* a Polish space. Ψ endowed with the pseudopath topology is a Polish space.
- According to [59][Lemma 1], the pseudopath topology on Ψ is convergence in Lebesgue measure on $\mathbf{D}$.

§B.2.2 Pseudopaths of stochastic processes on a general metric separable space

In [53] the results of [59] on state space $\mathbb{R}$ are generalised to a general metric separable space E. Let $(Z(t))_{t>0}$ be a stochastic process with state space E. Then we assign a random pseudopath to $(Z(t))$ as follows: for $\omega \in \Omega$ and $A = [a,b] \times B$, $0 \le a < b$ and $B \in \mathcal{B}(E)$,

$$\psi_{(Z(t,\omega))_{t\ge 0}}(A) = \int_a^b 1_B(Z(t,\omega))\, \mathrm{e}^{-t}\mathrm{d}t. \tag{B.27}$$

Hence $\psi_{(Z(t))_{t\ge 0}}$ is a random variable with state space Ψ, i.e., $\psi_{(Z(t))_{t\ge 0}} \in \mathcal{M}(\Psi)$, the set of probability measures on pseudopaths. Note that

$$\mathbb{E}\left[\psi_{(Z(t))_{t\ge 0}} f\right] = \mathbb{E}\left[\int_0^\infty f(t, Z(t,\omega))\, \mathrm{e}^{-t}\mathrm{d}t\right] = \mathbb{E}\left[\int_0^\infty f(t, Z(t))\, \mathrm{e}^{-t}\mathrm{d}t\right]. \tag{B.28}$$

Weak convergence in the Meyer-Zheng topology. Let $(Z_n(t))_{t\geq 0}$ and $(Z(t))_{t\geq 0}$ be stochastic processes with state-space E. We say that

$$\mathcal{L}\left[(Z_n(t))_{t\geq 0}\right] = \mathcal{L}\left[(Z(t))_{t\geq 0}\right] \text{ in the Meyer-Zheng topology} \tag{B.29}$$

if, for all $f \in \mathcal{C}_b(\Psi)$,

$$\lim_{n\to\infty} \mathbb{E}[f(\psi_{(Z^n(t))_{t\geq 0}})] = \mathbb{E}[f(\psi_{(Z(t))_{t\geq 0}})]. \tag{B.30}$$

Let $\mathcal{C}_m([0,\infty)\times E) \subset \mathcal{C}_b([0,\infty)\times E)$ be the set of functions of the form

$$\begin{aligned}\mathcal{C}_m([0,\infty)\times E) = \Big\{&F \in \mathcal{C}_b([0,\infty)\times E) : F(t,x(t)) = \prod_{i=1}^m \int_0^{T_i} f_i(t,x(t))\mathrm{d}t,\\ &m\in\mathbb{N},\ \forall 1\leq i\leq m,\ f_i \in \mathcal{C}_b([0,\infty)\times E),\ T_i>0\Big\}.\end{aligned} \tag{B.31}$$

Note that $\mathcal{C}_m$ is an algebra. Let $M_E[0,\infty)$ be the space of measurable processes from $[0,\infty)$ to $\mathbb{E}$, so $\mathbf{D} \subset M_E[0,\infty)$. Note that $\mathcal{C}_m$ separates points in $M_E[0,\infty)$. By [53][Proposition 4.5], the set $\mathcal{C}_m$ is separating in the set of measures on $M_E[0,\infty)$. This means that if two stochastic processes $(Z_1(t))_{t>0}$ and $(Z_2(t))_{t\geq 0}$ satisfy

$$\mathbb{E}[F(Z_1)] = \mathbb{E}[F(Z_2)] \qquad \forall F \in \mathcal{C}_m, \tag{B.32}$$

then $\mathcal{L}[Z_1] = \mathcal{L}[Z_2]$.

Define

$$F(\psi) = \int \mathrm{d}\psi \prod_{i=1}^m \int_0^{T_i} f_i(t,x(t))\,\mathrm{d}t. \tag{B.33}$$

Recall that a pseudopath ψ is associated with a path $w \in M_E[0,\infty)$. Hence this becomes

$$F(\psi_w) = \prod_{i=1}^m \int_0^{T_i} f_i(t,w(t))\,\mathrm{d}t. \tag{B.34}$$

Since each pseudopath $\psi\in\Psi$ is associated with a path in $M_E[0\infty)$, $\mathcal{C}_m$ also separates points on Ψ and hence $\mathcal{C}_m$ separates measures on Ψ. This implies that if

$$\mathbb{E}[F(\psi_{Z_1})] = \mathbb{E}[F(\psi_{Z_2})] \qquad \forall F \in \mathcal{C}_m, \tag{B.35}$$

then $\mathcal{L}[\psi_{Z_1}] = \mathcal{L}[\psi_{Z_2}]$. Therefore $\mathcal{L}[Z_1] = \mathcal{L}[Z_2]$ if and only if $\mathcal{L}[\psi_{Z_1}] = \mathcal{L}[\psi_{Z_2}]$.

The Meyer-Zheng topology is a weaker than the Skohorod topology.

Lemma B.2.1. *Let $(Z_n(t))_{t\geq 0}$ $n\in\mathbb{N}$ and $(Z(t))_{t\geq 0}$ be stochastic processes with Polish state-space E. If*

$$\lim_{n\to\infty} \mathcal{L}\left[(Z_n(t))_{t\geq 0}\right] = \mathcal{L}\left[(Z(t))_{t\geq 0}\right] \text{ in the Skohorod topology,} \tag{B.36}$$

then

$$\lim_{n\to\infty} \mathcal{L}\left[(Z_n(t))_{t\geq 0}\right] = \mathcal{L}\left[(Z(t))_{t\geq 0}\right] \text{ in the Meyer-Zheng topology.} \tag{B.37}$$

Proof. Since we do not know whether Ψ is compact, the set $\mathcal{C}_m$ does not have to be convergence determining. Therefore, via Skorohod's theorem we construct the process $\tilde{Z}^n$ and $\tilde{Z}$ on one probability space, such that $\mathcal{L}[\tilde{Z}^n] = \mathcal{L}[Z^n]$ and $\mathcal{L}[\tilde{Z}] = \mathcal{L}[Z]$, and

$$\lim_{n\to\infty} \tilde{Z}_n = \tilde{Z} \qquad a.s. \tag{B.38}$$

This implies

$$\lim_{n\to\infty} \psi_{\tilde{Z}^n} = \psi_{\tilde{Z}} \qquad a.s. \tag{B.39}$$

Consequently, for all $f \in \mathcal{C}_b(\Psi)$,

$$\lim_{n\to\infty} \mathbb{E}[f(\psi_{\tilde{Z}^n})] = \mathbb{E}[f(\psi_{\tilde{Z}})]. \tag{B.40}$$

Note that, since $\mathcal{L}[\tilde{Z}^n] = \mathcal{L}[Z^n]$ and $\mathcal{L}[\tilde{Z}] = \mathcal{L}[Z]$, we can use (B.32) and (B.35) to see that the latter implies $\mathcal{L}[\psi_{Z^n}] = \mathcal{L}[\psi_{\tilde{Z}^n}]$ and $\mathcal{L}[\psi_Z] = \mathcal{L}[\psi_{\tilde{Z}}]$. Hence (B.40) indeed implies that

$$\lim_{n\to\infty} \mathcal{L}[\psi_{Z^n}] = \mathcal{L}[\psi_Z]. \tag{B.41}$$

□

Convergence in probability in the Meyer-Zheng topology. Let (S, d) be a metric space, $\mathcal{B}(S)$ denote the Borel-σ algebra on S, and $\mathcal{P}(S)$ the set of probability measures on $\mathcal{B}(S)$. Recall (see e.g. [32, Chapter 3]) that the Prohorov metric d_P on the space $\mathcal{P}(S)$ is given by

$$d_P(\mathbb{P}, \mathbb{Q}) = \inf \{\epsilon > 0 \colon \mathbb{P}(A) \le \mathbb{Q}(A^\epsilon) + \epsilon \ \forall\, A \in \mathcal{C}\}, \tag{B.42}$$

where $\mathcal{C} \subset \mathcal{B}(S)$ is the set of all closed sets in S and $A^\epsilon = \{x \in S\colon \inf_{y\in A} d(x,y) < \epsilon\}$.

Recall the following theorem (see e.g.[[32, Theorem 3.1.2]])

Theorem B.2.2. *Let (S, d) be separable and let $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(S)$. Define $\mathcal{M}(\mathbb{P}, \mathbb{Q})$ to be the set of all $\mu \in \mathcal{P}(S \times S)$ with marginals $\mathbb{P}$ and $\mathbb{Q}$, i.e., $\mu(A \times S) = \mathbb{P}(A)$ and $\mu(S \times A) = \mathbb{Q}(A)$ for all $A \in \mathcal{B}(S)$. Then*

$$d_P(\mathbb{P}, \mathbb{Q}) = \inf_{\mu\in\mathcal{M}(\mathbb{P},\mathbb{Q})} \inf\{\epsilon > 0\colon\ \mu(\{(x, y)\colon\ d(x, y) \ge \epsilon\}) \le \epsilon\}. \tag{B.43}$$

Moreover, [32, Theorem 3.3.1] states that convergence of measures in the Prohorov distance, $\lim_{n\to\infty} d_P(\mathbb{P}_n, \mathbb{P}) = 0$, is the same as weak convergence $\mathbb{P}_n \Rightarrow \mathbb{P}$. Hence, since convergence of pseudopaths is weak convergence, we can endow the space of pseudopaths Ψ with the metric d_P.

Let (Ψ, d_P) be the pseudopath space metrized by the Prohorov distance. Let $(Z^n(t))_{t>0}$, $(Z(t))_{t>0}$ be stochastic processes on the state space E, where E is endowed with the metric $d(\cdot,\cdot)$. Note that convergence in probability in the Meyer-Zheng topology means that

APPENDIX B

$$\forall\, \delta > 0\colon \qquad \lim_{n\to\infty} \mathbb{P}\left[d_P\left(\psi_{Z^n}, \psi_Z\right) > \delta\right] = 0. \tag{B.44}$$

Tightness. Define the *conditional variation* for an $\mathbb{R}$-valued process $(U(t))_{t\geq 0}$ with natural filtration $(\mathcal{F}(t))_{t\geq 0}$ as follows. For a subdivision
$\tau\colon\ 0 = t_0 < t_1 < \cdots < t_n = \infty$, set

$$V_\tau(U) = \sum_{0\leq i<n} \mathbb{E}\Big[\big|\mathbb{E}[U(t_{i+1}) - U(t_i) \mid F(t_i)]\big|\Big] \tag{B.45}$$

(with $U(\infty) = 0$) and

$$V(U) = \sup_\tau V_\tau(U). \tag{B.46}$$

If $V(U) < \infty$, then U is called a *quasi-martingale.* Note that we can always stop the process at some finite time and work with compact time intervals.

Lemma B.2.3 (Tightness in the Meyer-Zheng topology).
If $(P_n)_{n\in\mathbb{N}}$ is a sequence of probability laws on $D([0,T],\mathbb{R})$ such that under P_n the coordinate process $(U(t))_{t\geq 0}$ is a quasi-martingale with a conditional variation $V_n(U)$ that is bounded uniformly in n, then there exists a subsequence $(P_{n_k})_{k\in\mathbb{N}}$ that converges weakly in the Meyer-Zheng topology on $D([0,T],\mathbb{R})$ to a probability law P, and $(U(t))_{t\geq 0}$ is a quasi-martingale under P.

(See [59, Theorem 7] for the identification of the limiting semi-martingale.)

§B.2.3 Proof of key lemmas

- **Proof of Lemma 6.2.19.**

Proof. Fix $\delta > 0$. Then

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}\left[d_P(\psi_{Z_n},\psi_Z) > \delta\right]\\
&= \lim_{n\to\infty} \mathbb{P}\left[\inf_{\mu\in\mathcal{M}(\psi_{Z_n},\psi_Z)} \inf\{\epsilon>0\colon\ \mu(\{(x,y)\colon\ d(x,y)\geq\epsilon\})\leq\epsilon\} > \delta\right]\\
&= \lim_{n\to\infty} \mathbb{P}\left[\forall\mu\in\mathcal{M}(\psi_{Z_n},\psi_Z),\ \inf\{\epsilon>0\colon\ \mu(\{(x,y)\colon\ d(x,y)\geq\epsilon\})\leq\epsilon\} > \delta\right]\\
&= \lim_{n\to\infty} \mathbb{P}\left[\forall\mu\in\mathcal{M}(\psi_{Z_n},\psi_Z),\ \mu(\{(x,y)\colon\ d(x,y)\geq\delta\}) > \delta\right].
\end{aligned} \tag{B.47}$$

Let $\mu_n \in \mathcal{P}(([0,\infty]\times E)^2)$ be the measure defined by

$$\mu_n(A) = \int_0^\infty 1_A\left((t,Z_n(t)),(t,Z(t))\right)\,\mathrm{e}^{-t}\mathrm{d}t, \qquad A\in\mathcal{B}(([0,\infty]\times E)^2), \tag{B.48}$$

such that, for $B\in\mathcal{B}([0,\infty]\times E)$,

$$\mu_n(B\times S) = \int_0^\infty 1_B(t,Z_n(t))1_S((t,Z(t))\,\mathrm{e}^{-t}\mathrm{d}t = \psi_{Z^n}(B), \tag{B.49}$$

and similarly $\mu_n(S\times B) = \psi_Z(B)$. Hence $\mu_n\in\mathcal{M}(\psi_{Z_n},\psi_Z)$ for all $n\in\mathbb{N}$, and we obtain from (B.47) that

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}\left[d_P(\psi_{Z_n},\psi_Z) > \delta\right] \\
&\le \lim_{n\to\infty} \mathbb{P}\left[\mu_n(\{(x,y) : d(x,y) \ge \delta\}) > \delta\right] \\
&\le \lim_{n\to\infty} \mathbb{P}\left[\int_0^\infty 1_{\{(x,y):d(x,y)\ge\delta\}}\left((t,Z_n(t)),(t,Z(t))\right)\mathrm{e}^{-t}\mathrm{d}t > \delta\right] \\
&\le \lim_{n\to\infty} \mathbb{P}\left[\int_0^\infty 1_{\{d(Z_n(t),Z(t))\ge\delta\}}\,\mathrm{e}^{-t}\mathrm{d}t > \delta\right] \\
&\le \lim_{n\to\infty} \frac{1}{\delta}\mathbb{E}\left[\int_0^\infty d(Z_n(t),Z(t))\,\mathrm{e}^{-t}\mathrm{d}t\right] \\
&= \lim_{n\to\infty} \frac{1}{\delta}\int_0^\infty \mathbb{E}\left[d(Z_n(t),Z(t))\right]\,\mathrm{e}^{-t}\mathrm{d}t = 0.
\end{aligned} \tag{B.50}$$

□

• **Proof of Lemma 6.2.20.**

Proof. We have to show that

$$\lim_{n\to\infty} \mathcal{L}\left[\psi_{(X_n,Y_n)}\right] = \mathcal{L}\left[\psi_{(X,c)}\right]. \tag{B.51}$$

Hence we must show that, for all $f \in \mathcal{C}_b(\Psi)$,

$$\lim_{n\to\infty} \mathbb{E}[f(\psi_{(X_n,Y_n)})] = \mathbb{E}[f(\psi_{(X,c)})]. \tag{B.52}$$

We can write

$$\begin{aligned}
&|\mathbb{E}[f(\psi_{(X_n,Y_n)}) - f(\psi_{(X,c)})]| \\
&\le |\mathbb{E}[f(\psi_{(X_n,Y_n)}) - f(\psi_{(X_n,c)})]| + |\mathbb{E}[f(\psi_{(X_n,c)}) - f(\psi_{(X,c)})]|.
\end{aligned} \tag{B.53}$$

Since $\lim_{n\to\infty}\mathbb{E}[d(Y_n(t),c)] = 0$ implies $\lim_{n\to\infty}\mathbb{E}[d((X_n(t),Y_n(t)),(X_n(t),c))] = 0$, it follows from Lemma 6.2.19 that, for all $\delta > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left[d_P\left(\psi_{(X_n,Y_n)},\psi_{(X_n,c)}\right)\right] = 0. \tag{B.54}$$

Hence, for all $f \in \mathcal{C}_b(\Psi)$,

$$\lim_{n\to\infty} |\mathbb{E}[f(\psi_{(X_n,Y_n)}) - f(\psi_{(X_n,c)})]| = 0. \tag{B.55}$$

To see that the second term in the right-hand side of (B.53) tends to zero, note that we can define

$$\tilde{f}(\psi_x) = f(\psi_{x,c}). \tag{B.56}$$

We show that $\tilde{f}$ is continuous.

Recall that convergence in the Meyer-Zheng topology is simply convergence in Lebesgue measure. Hence, for two paths $(t,x_n(t))$ and $(t,x(t)) \in M_E[0\infty)$ we have $\psi_{x_n} \to \psi_x$ if and only if, for all $\delta > 0$,

$$\lim_{n\to\infty} \int_0^\infty 1_{\{d(x_n(t),x(t))>\delta\}}\,\mathrm{e}^{-t}\mathrm{d}t = 0. \tag{B.57}$$

Appendix B

Therefore $\psi_{x_n} \to \psi_x$ implies that, for all $\delta > 0$,

$$\lim_{n\to\infty} \int_0^\infty 1_{\{d((x_n(t),c),(x(t),c))>\delta\}}\, \mathrm{e}^{-t} \mathrm{d}t = 0, \tag{B.58}$$

and hence $\psi_{x_n,c} \to \psi_{x,c}$. Therefore

$$\lim_{n\to\infty} \tilde{f}(\psi_{x_n}) = \lim_{n\to\infty} f(\psi_{(x_n,c)}) = f(\psi_{(x,c)}) = \tilde{f}(\psi_x) \tag{B.59}$$

and we conclude that $f \in \mathcal{C}_b(\Psi)$. Since $\mathcal{L}[X_n] = \mathcal{L}[X]$ in the Meyer-Zheng topology, we have, for all $f \in \mathcal{C}_b(\Psi)$,

$$\lim_{n\to\infty} |\mathbb{E}[f(\psi_{(X_n,c)}) - f(\psi_{(X,c)})]| = \lim_{n\to\infty} |\mathbb{E}[\tilde{f}(\psi_{(X_n)}) - \tilde{f}(\psi_{(X)})]| = 0. \tag{B.60}$$

Therefore also the second term on the right-hand side of (B.53) tends to 0. □

• **Proof of Lemma 6.2.21.**

Proof. For part (a), suppose that $\lim_{n\to\infty} \psi_{x_n} = \psi_x$. Then, since convergence in pseudopath space is convergence in measure, we have, for all $\delta > 0$,

$$\lim_{n\to\infty} \int_0^\infty 1_{\{d(x_n(t),x(t))>\delta\}}\, \mathrm{e}^{-t} \mathrm{d}t = 0. \tag{B.61}$$

Since f is a continuous function, this implies that, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \int_0^\infty 1_{\{d(f(x_n(t)),f(x(t)))>\epsilon\}}\, \mathrm{e}^{-t} \mathrm{d}t = 0. \tag{B.62}$$

and we conclude that $\lim_{n\to\infty} \psi_{f(x_n)} = \psi_{f(x)}$. Hence h is indeed continuous.

For part (b), recall that

$$\lim_{n\to\infty} \mathcal{L}[X_n] = \mathcal{L}[X] \text{ in the Meyer-Zheng topology} \tag{B.63}$$

implies that, for all $g \in \mathcal{C}_b(\Psi)$,

$$\lim_{n\to\infty} \mathbb{E}[g(\psi_{X_n})] = \mathbb{E}[g(\psi_X)]. \tag{B.64}$$

Since $h\colon \Psi \to \Psi$ is continuous, we have for all $g \in \mathcal{C}_b(\Psi)$ that $g \circ h \in \mathcal{C}_b(\Psi)$. Hence

$$\lim_{n\to\infty} \mathbb{E}[g(\psi_{f(X_n)}] = \lim_{n\to\infty} \mathbb{E}[g \circ h(\psi_{X_n})] = \mathbb{E}[g \circ h(\psi_X)] = \mathbb{E}[g(\psi_{f(X)})]. \tag{B.65}$$

We conclude that

$$\lim_{n\to\infty} \mathcal{L}[f(X_n)] = \mathcal{L}[f(X)] \text{ in the Meyer-Zheng topology.} \tag{B.66}$$

□

- **Proof of Lemma 7.2.14.**

Proof. Suppose that $\lim_{n\to\infty}\psi_{(x_n,y_n)} = \psi_{(x,y)}$. Since convergence of pseudopaths is convergence in Lebesgue measure, we have

$$\lim_{n\to\infty}\int_0^\infty 1_{\{d[(x_n,y_n),(x,y)]>\delta\}}\mathrm{e}^{-t}\mathrm{d}t = 0 \tag{B.67}$$

and, consequently,

$$\lim_{n\to\infty}\int_0^\infty 1_{\{d[x_n,x]>\delta\}}\mathrm{e}^{-t}\mathrm{d}t = 0. \tag{B.68}$$

Therefore $\lim_{n\to\infty}\psi_{x_n} = \psi_x$. Suppose that $f\in\mathcal{C}_b(\Psi(E))$, so f is bounded continuous function on the space of pseudopaths on $[0,\infty]\times E$. Define the function $\tilde{f}$ on the space of pseudopaths on $[0,\infty]\times E^2$, i.e., $\tilde{f}$ is a function on $\Psi(E^2)$, by

$$\tilde{f}(\psi_{(x,y)}) = f(\psi_x). \tag{B.69}$$

Then $\tilde{f}\in\mathcal{C}_b(\Psi(E^2))$ and

$$\lim_{n\to\infty}\tilde{f}(\psi_{(x_n,y_n)}) = \lim_{n\to\infty}f(\psi_{x_n}) = f(\psi_x) = \tilde{f}(\psi_{(x_n,y_n)}). \tag{B.70}$$

Hence $\tilde{f}$ is indeed a continuous function on $\Psi(E^2)$. Moreover, since f is bounded, it follows that $\tilde{f}$ is bounded and we conclude that $\tilde{f}\in\mathcal{C}_b(\Psi(E^2))$.

Therefore, if X_n, Y_n are continuous-time stochastic processes on E and

$$\lim_{n\to\infty}\mathcal{L}\left[(X_n(s),Y_n(s))_{s>0}\right] = \mathcal{L}\left[(X(s),Y(s))_{s>0}\right] \text{ in Meyer Zheng topology,} \tag{B.71}$$

then for all $f\in\mathcal{C}_b(\Psi(E^2))$ we have

$$\lim_{n\to\infty}\mathbb{E}[f(\psi_{(X_n,Y_n)})] = \mathbb{E}[f(\psi_{(X,Y)})]. \tag{B.72}$$

Since for each $f\in\mathcal{C}_b(\Psi(E))$ we can construct a function $\tilde{f}\in\mathcal{C}_b(\Psi(E^2))$ as in (B.69), we obtain for all $f\in\mathcal{C}_b(\Psi(E))$ that

$$\lim_{n\to\infty}\mathbb{E}[f(\psi_{(X_n)})] = \lim_{n\to\infty}\mathbb{E}[\tilde{f}(\psi_{(X_n,Y_n)})] = \mathbb{E}[\tilde{f}(\psi_{(X,Y)})] = \mathbb{E}[\tilde{f}(\psi_X)]. \tag{B.73}$$

We conclude that

$$\lim_{n\to\infty}\mathcal{L}[(X_n(s))] = \mathcal{L}[(X(s))_{s>0}] \text{ in Meyer-Zheng topology} \tag{B.74}$$

and, similarly,

$$\lim_{n\to\infty}\mathcal{L}[(Y_n(s))] = \mathcal{L}[(Y(s))_{s>0}] \text{ in Meyer-Zheng topology.} \tag{B.75}$$

□

Bibliography

[1] K. Alexander and Q. Berger. Local asymptotics for the first intersection of two independent renewals. *Electr. J. Probab.*, 21:paper no. 66, 20 pp., 2016.

[2] R. Arratia. *Coalescing Brownian Motions on the Line.* PhD thesis, University of Wisconsin-Madison, USA, 1979.

[3] S. Asmussen. *Applied Probability and Queues*, volume 51 of *Applications of Mathematics.* Springer-Verlag, New York, USA, 2003.

[4] E. Baake and A. Wakolbinger, editors. *Probabilistic Structures in Evolution.* European Mathematical Society Publishing House, 2021.

[5] J.-B. Baillon, P. Clement, A. Greven, and F. den Hollander. On the attracting orbit of a non-linear transformation arising from hierarchically interacting diffusions. Part I: The compact case. *Can. J. Math.*, 47:3–27, 1995.

[6] J.-B. Baillon, P. Clement, A. Greven, and F. den Hollander. On the attracting orbit of a non-linear transformation arising from hierarchically interacting diffusions. Part II: The non-compact case. *J. Funct. Anal.*, 146:236–298, 1997.

[7] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.

[8] N. Bingham, C. Goldie, and J. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 1987.

[9] J. Blath, E. Buzzoni, A. Gonzales Casanova, and M. Wilke-Berenguer. Structural properties of the seed bank and the two island diffusion. *J. Math. Biology*, 79:369–392, 2019.

[10] J. Blath, A. Gonzales Casanova, B. Eldon, and N. Kurt. Genealogy of a Wright-Fisher model with strong seed-bank component. *Birkhäuser, Progress in Probability*, 69:81–100, 2015.

[11] J. Blath, A. Gonzales Casanova, N. Kurt, and D. Spano. The ancestral process of long-range seed bank models. *J. Appl. Prob.*, 50:741–759, 2013.

[12] J. Blath, A. Gonzales Casanova, N. Kurt, and M. Wilke-Berenguer. A new coalescent for seed-bank models. *Ann. Appl. Probab.*, 26:857–891, 2016.

[13] L. Breiman. *Probability.* Addison-Wesley, Reading, Massachusetts, 1968.

[14] J. Cox and A. Greven. Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.*, 22:833–853, 1994.

[15] J. Cox, A. Greven, and T. Shiga. Finite and infinite systems of interacting diffusions. *Probab. Theory Rel. Fields*, 102:165–197, 1995.

[16] J. Cox and D. Griffeath. Diffusive clustering in the two-dimensional voter model. *Ann. Probab.*, 14:347–370, 1986.

[17] D. Dawson. *Measure-Valued Markov Processes*, volume 1541 of *Lecture Notes in Mathematics.* Springer-Verlag, New York, 1993.

[18] D. Dawson, L. Gorostiza, and A. Wakolbinger. Hierarchical random walks. In *Asymptotic Methods in Stochastics*, volume 44 of *Fields Inst. Commun.*, pages 173–193. Amer. Math. Soc., Providence, RI, 2004.

[19] D. Dawson, L. Gorostiza, and A. Wakolbinger. Degrees of transience and recurrence and hierarchical random walks. *Potential Anal.*, 22:305–350, 2005.

[20] D. Dawson and A. Greven. Hierarchical models of interacting diffusions: multiple times scales, phase transitions and cluster formation. *Probab. Theory Relat. Fields*, 96:435–473, 1993.

[21] D. Dawson and A. Greven. Multiple scale analysis of interacting diffusions. *Probab. Theory Relat. Fiels*, 95:467–508, 1993.

[22] D. Dawson and A. Greven. Multiple space-time scale analysis for interacting branching models. *Electron. J. Probab.*, 1:paper no. 14, 1–84, 1996.

[23] D. Dawson and A. Greven. *Spatial Fleming-Viot Models with Selection and Mutation*, volume 2092 of *Springer Lecture Notes in Math.* Springer Verlag, 2014.

[24] D. Dawson, A. Greven, F. den Hollander, R. Sun, and J. Swart. The renormalization transformation of two-type branching models. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44:1038–1077, 2008.

[25] D. Dawson, A. Greven, and J. Vaillancourt. Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc.*, 347:2277–2360, 1995.

[26] F. den Hollander. Renormalization of interacting diffusions. In *Complex Stochastic Systems (eds. O.E. Barndorff-Nielsen, D.R. Cox, C. Kluppelberg)*, volume 87 of *Monographs on Statistics and Applied Probability*, pages 219–233. Chapman and Hall, Boca Raton, 2001.

[27] F. den Hollander. Renormalization of interacting diffusions: a program and four examples. In *Partial Differential Equations and Functional Analysis*, volume 168 of *Operator Theory, Advances and Applications*, pages 123–136. Birkhauser, Basel, 2006.

[28] F. den Hollander and G. Pederzani. Multi-colony Wright-Fisher with seed-bank. *Indag. Math.*, 28:637–669, 2017.

[29] F. den Hollander and J. Swart. Renormalization of hierarchically interacting isotropic diffusions. *J. Stat. Phys.*, 93:243–291, 1998.

[30] R. Durret. *Probability Models for DNA Sequence Evolution (2nd ed)*. Springer, 2008.

[31] N. Ercolani, J. Jansen, and D. Ueltschi. Singularity analysis for heavy-tailed random variables. *J. Theor. Probab.*, 32:1–46, 2019.

[32] S. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence.* John Wiley, New York, 1986.

[33] S. Evans and K. Fleischmann. Cluster formation in a stepping stone model with continuous, hierarchically structured sites. *Ann. Probab.*, 24:1926–1952, 1996.

[34] W. Feller. *An Introduction to Probability Theory and Its Applications (2nd ed).* John Wiley, New York, 1971.

[35] K. Fleischmann and A. Greven. Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Rel. Fields*, 98:517–566, 1994.

[36] K. Fleischmann and A. Greven. Time-space analysis of the cluster formation in interacting diffusions. *Electr. J. Probab.*, 1:1–46, 1996.

[37] G. Folland. *Fourier Analysis and its Applications.* Wadsworth & Brooks/Cole, Belmont, CA, USA, 1992.

[38] J. Gartner. On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, 137:197–248, 1988.

[39] S. Goldstein. Maximal coupling. *Z. Wahrsch. Verw. Gebiete*, 46:193–204, 1978/79.

[40] A. Gonzales Casanova, E. Aguirre-von Wobeser, G. Espin, N. Servin-Gonzalez, L. Kurt, D. Spano, J. Blath, and G. Soberon-Chavez. Strong seed-bank effects in bacterial evolution. *J. Theor. Biol.*, 356:62–70, 2014.

[41] A. Greven, F. den Hollander, S. Kliem, and A. Klimovsky. Renormalisation of hierarchically interacting Cannings processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11:43–140, 2014.

[42] A. Greven, F. den Hollander, and A. Klimovsky. The hierarchical Cannings process in random environment. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 15:295–351, 2018.

[43] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: well-posedness, duality and equilibrium. *Preprint*, 2020.

[44] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: finite-systems scheme. *In Preparation*, 2021.

[45] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: renormalisation on the hierarchical group. *Preprint*, 2021.

[46] I. Herbst and L. Pitt. Diffusion equation techniques in stochastic monotonicity and positive correlations. *Probab. Theory Relat. Fields*, 87:275–312, 1991.

[47] B. Hughes. *Random Walk in Random Environment*, volume I. Clarendon Press, 1995.

[48] S. Jansen and N. Kurt. On the notion(s) of duality for Markov processes. *Probab. Surv.*, 11:59–120, 2014.

[49] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.*, 18(1):20–65, 1986.

[50] I. Kaj, S. Krone, and M. Lascoux. Coalescent theory for seed bank models. *J. Appl. Prob.*, 38:285–300, 2001.

[51] A. Klenke. Different clustering regimes in hierarchically interacting diffusions. *Ann. Probab.*, 24:660–697, 1996.

[52] A. Klenke. Multiscale analysis of clusters in spatial branching models. *Ann. Probab.*, 25:1670–1711, 1997.

[53] T. G. Kurtz. Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Ann. Probab.*, 19(3):1010–1034, 1991.

[54] J. Lennon, W. den Hollander, M. Wilke-Berenguer, and J. Blath. Principles of seed-banks: complexity emerging from dormancy. *Nature Communication*, 2021.

[55] J. Lennon and S. Jones. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9:119–130, 2011.

[56] T. Liggett. *Interacting Particle Systems.* Springer, New York, 1985.

[57] L. Lindenstrauss. Pointwise theorems for amenable groups. *Electron. Res. Announc. AMS*, 5:82–90, 1999.

[58] L. Lindvall. *Lectures on the Coupling Method.* Wiley, New York, 1992.

[59] P. Meyer and W. Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré, Prob. Stat.*, 20:353–372, 1984.

[60] T. Mojet. Scaling limits of moran and wright-fisher genetic drift models with seedbank components. Bachelor thesis, 2018.

[61] C. Newman and A. Wright. An invariance principle for certain dependent sequences. *Ann. Probab.*, 9:671–675, 1981.

[62] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, New York, 1999.

[63] L. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus.* Cambridge University Press, 2000. 494 pages.

[64] W. Rudin. *Fourier Analysis on Groups*, volume 12 of *Interscience Tracts in Pure and Applied Mathematics*. Interscience Publishers, New York, 1962.

[65] S. Sawyer and J. Felsenstein. Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.*, 20:1–10, 1983.

[66] T. Shiga. An interacting system in population genetics. *J. Math. Kyoto Univ.*, 20:213–242, 1980.

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

[68] W. Smith. Regenerative stochastic processes. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 232:6–31, 1955.

[69] F. Spitzer. *Principles of Random Walk*, volume 34. D. Van Nostrand Company, Princeton, NJ, 1964.

[70] A. Tellier, S. J. Y. Laurent, H. Lainer, P. Pavlidis, and W. Stephan. Inference of seed bank parameters in two wild tomato species using ecological and genetic data. *Proceedings of the National Academy of Sciences*, 108(41):17052–17057, 2011.

[71] A. Winter. Multiple scale analysis of branching processes under the palm distribution. *Electr. J. Prob.*, 7:paper no. 13, 72 pp, 2002.

[72] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971.

Samenvatting

Stel je een weide voor met rode en paarse klaprozen. Je kunt je afvragen of er een moment komt dat er alleen nog maar rode klaprozen zijn of juist alleen nog maar paarse klaprozen, of dat er altijd zowel rode als paarse klaprozen zullen zijn. En als er na vele generaties alleen nog één kleur klaprozen overblijft, met welke kans zijn dit dan de rode klaprozen en met welke kans zijn dit de paarse klaprozen?

Bovenstaande vragen gaan over de genetische evolutie van een populatie klaprozen. Als er na een aantal generaties nog maar één kleur klaprozen overblijft, zeggen we dat de genetische diversiteit van de populatie verloren is gegaan. In dit proefschrift bestuderen we genetische evolutie in populaties. Een wiskundig model om de genetische evolutie in een populatie te beschrijven is het Fisher-Wright model. Het Fisher-Wright model kan uitgebreid worden met verschillende evolutiemechanismen, zoals mutaties van genen of de selectie van sterkere genen. Ook bestaat er een model met meerdere koloniën, waarbij individuen in verschillende koloniën leven die elk door het Fisher-Wright model beschreven worden en waarbij de individuen kunnen migreren tussen de verschillende koloniën.

Een recente uitbreiding van het Fisher-Wright model is de toevoeging van een zaadbank aan de populatie. In een populatie met een zaadbank kan een individu voor een zekere tijd stoppen met zichzelf voort te planten. We zeggen dan dat het individu gaat "slapen". Alle slapende individuen samen vormen de zaadbank. Na een aantal generaties wordt het individu wakker en begint zich weer voor te planten. De klaproos is een soort die een zaadbank heeft. Als een zaadje van een klaproos in de grond belandt, maar door een storm met een dikke laag zand bedekt wordt, dan kan het gebeuren dat dit zaadje niet de volgende lente ontkiemt, maar pas vijf lentes later. Het zaadbank fenomeen wordt ook bij bacteriën waargenomen. Wanneer de omstandigheden voor een bacterie slecht zijn, bijvoorbeeld door te weinig voedsel of te lage temperaturen, dan kan een bacterie zich omvormen tot een zogeheten endospore. Een endospore kan overleven in moeilijke omstandigheden, maar kan zich niet voortplanten. Als de omstandigheden verbeteren dan kan de endospore weer terug transformeren in een bacterie en opnieuw zichzelf voortplanten.

In dit proefschrift bestuderen we het Fisher-Wright model waarbij individuen een van twee (gen)types kunnen zijn, tussen verschillende koloniën kunnen migreren, en in een zaadbank kunnen gaan slapen voor een bepaalde tijd. We nemen aan dat in elke kolonie oneindig veel individuen leven. Het bijbehorende wiskundige model heet het "multi-kolonie continuüm Fisher-Wright model met zaadbank". Doel van het proefschrift is om te bepalen of na zeer lange tijd de diversiteit in de populatie behouden blijft (in het geval van de klaprozen of er altijd rode en paarse klaprozen blijven bestaan) of dat de diversiteit verloren gaat (of na lange tijd er alleen nog maar rode of alleen nog maar paarse klaprozen bestaan). In het bijzonder proberen we het effect van de zaadbank op de genetische diversiteit vast te stellen.

We bestuderen dit vraagstuk in verschillende geografische ruimtes. In het eerste deel van het proefschrift zijn de koloniën waarin de individuen leven geplaatst in een algemene geografische ruimte. Het belangrijkste voorbeeld van zo'n geografische ruimte is het rooster $\mathbb{Z}^d$. In het tweede deel van het proefschrift zijn de koloniën waarin de individuen leven geplaatst volgens de hiërarchische groep. Intuïtief kan de hiërarchische groep als volgt geïnterpreteerd worden. Elke kolonie is een huis, een aantal huizen samen vormen een straat, een aantal straten samen vormen een stad, een aantal steden samen een provincie enzovoort. Individuen bewegen (migreren) veel vaker tussen de huizen in de straat, dan tussen de verschillende straten in stad. Een niveau hoger bewegen individuen veel vaker binnen hun eigen stad dan tussen de verschillende steden binnen de provincie, en nog minder vaak tussen de verschillende provincies. Deze structuur komt vaak voor in de ecologie. De hiërarchische groep beschrijft deze structuur op een wiskundige manier.

Het proefschrift bestaat uit twee delen voorafgegaan door een introductie in het vakgebied in Hoofdstuk 1. In deel I van het proefschrift bestuderen we drie varianten van "het multi-kolonie continuüm Fisher-Wright model met zaadbank". In elk van deze drie varianten zijn er twee (gen)type individuen, aangeduid met type $\heartsuit$ en type $\diamondsuit$. In het eerste model bestaat de populatie in elke kolonie uit actieve individuen en slapende individuen. De slapende individuen samen vormen de zaadbank. De actieve individuen kunnen zich voortplanten, migreren naar een andere kolonie, en gaan slapen. De slapende individuen kunnen alleen wakker worden: ze planten zich niet voort en migreren ook niet. Het tweede model is een uitbreiding van het eerste model waarbij we de zaadbank een interne structuur geven. Door de interne structuur kan een individu op verschillende manieren gaan slapen, langer of korter. In het derde model heeft de zaadbank dezelfde interne structuur als in het tweede model, maar kunnen individuen gaan slapen in verschillende koloniën. In elk van de drie modellen kan de voortplantingssnelheid afhankelijk zijn van de diversiteit in de populatie.

In hoofdstuk 2 stellen we voor elk van de drie modellen een stelsel van stochastische differentiaalvergelijkingen op. Deze stochastische differentiaalvergelijkingen beschrijven de frequentie van gentype $\heartsuit$ in de populatie. We tonen aan dat de processen die deze differentiaalvergelijkingen beschrijven goed gedefinieerd zijn en de Markov eigenschap hebben. Door de interne structuur van de zaadbank in het tweede en derde model heeft de tijd die een individu in de zaadbank doorbrengt een verdeling met een dikke staart, maar gaat de Markoveigenschap van het model niet verloren.

In hoofdstuk 2 tonen we ook aan dat, in het bijzondere geval dat de diffusie functie in de stochastische differentiaalvergelijkingen de Fisher-Wright diffusie is, er een duaal proces bestaat. Dit duale process stelt ons in staat voor elk model het lange termijn gedrag van het stelsel van stochastische differentiaalvergelijkingen te analyseren.

Tenslotte bepalen we in hoofdstuk 2 voor elk van de drie modellen of op de lange termijn de diversiteit in de populatie verloren gaat of dat deze altijd behouden blijft. Het blijkt dat in het eerste model enkel de manier waarop individuen migreren bepaalt of de diversiteit in de populatie verloren gaat of niet. In het tweede model blijkt dat als individuen lang genoeg slapen, zowel de manier waarop individuen migreren als de manier waarop individuen gaan slapen bepaalt of de genetische diversiteit behouden blijft. De interne structuur van de zaadbank zorgt ervoor dat de zaadbank kan voorkomen dat de diversiteit in de populatie verdwijnt. Als in het tweede model de individuen heel erg lang slapen dan blijkt dat de genetische diversiteit altijd behouden zal blijven, onafhankelijk van de manier waarop de individuen migreren. Voor het derde model vinden we dezelfde resultaten als voor het tweede model. Hoofdstuk 3 bevat de bewijzen van de stellingen in hoofdstuk 2.

In deel II van dit proefschrift bekijken we het tweede model in het specifieke geval dat de geografische ruimte de hiërarchische groep is. Het eerste doel van deel II is om de resultaten uit deel I toe te passen. Het tweede doel is het analyseren van de genetische diversiteit in de zogeheten “hiërarchische gemiddelde veld limiet”. In hoofdstuk 4 geven we een formele beschrijving van het model op de hiërarchische groep. Daarnaast geven we de resultaten volgend uit deel I van het proefschrift en beschrijven we de “hiërarchische gemiddelde veld limiet”. Hoofdstukken 5 tot en met 10 bevatten de bewijzen van de stellingen in hoofdstuk 4. Opnieuw blijkt dat als de individuen lang genoeg slapen de zaadbank kan voorkomen dat de diversiteit in de populatie verdwijnt.

Summary

Imagine a meadow with red and purple poppies. You might wonder whether there will be a moment when there are only red poppies left or only purple poppies left, or whether there will always be both red and purple poppies. And if, after multiple generations, only one colour is left, then what is the probability that there are only red poppies left and what is the probability there are only purple poppies left?

The above questions concern the genetic evolution of a poppy population. If after a number of generations only one colour of poppies is left, then we say that the genetic diversity in the population is lost. In this thesis we study genetic evolution in populations in a broader setting. A mathematical model that describes the genetic evolution in a population is the Fisher-Wright model. Different evolutionary mechanisms can be incorporated into the Fisher-Wright model, for example, mutation of genes or selection of stronger genes. There also exists a model where individuals live in multiple colonies, each evolving according to the Fisher-Wright model, and individuals are allowed to migrate between different colonies.

A recent extension of the Fisher-Wright model is the addition of a seed-bank to the population. In a population with seed-bank, individuals can stop reproducing themselves for awhile. We then say that these individuals become dormant. The repository of the dormant individuals is called the seed-bank. After a number of generations individuals resuscitate and reprise reproduction. The poppy is a species that has a seed-bank. If a poppy seed drops on the soil, but due to a storm gets covered with a thick layer of sand, then it may happen that the seed does not germinate next spring, but only five springs later. The seed-bank phenomenon is also observed in bacteria. When environmental conditions are bad, for example, the nutrition level is low or the temperature is low, a bacteria can produce an endospore. An endospore can survive difficult conditions, but cannot reproduce itself. When the environmental conditions improve, the endospore can transform itself back into a bacteria and reprise reproduction.

In this thesis we study the Fisher-Wright model with seed-bank in which individuals carry one of two gene types, the individuals can migrate between different colonies, and can become dormant in a seed-bank for a certain amount of time. We assume that in each colony there are infinitely many individuals. The corresponding mathematical model is called the “multi-colony continuum Fisher-Wright model with seed-bank”. The goal of this thesis is to determine whether on the long term the genetic diversity will be maintained (in case of the poppies, there will always be both red and purple poppies) or whether genetic diversity is lost (there are after a long time only red or only purple poppies). In particular, we try to determine the effect of the seed-bank on the genetic diversity in the population.

We study this question in different geographical spaces. In the first part of the thesis the colonies in which individuals live are placed on a general geographical space.

The most important example of such a space is the lattice $\mathbb{Z}^d$. In the second part of the thesis the colonies are placed on the hierarchical group. Intuitively, the hierarchical group may be interpreted in the following way. Each colony is a house, a couple of houses together forms a street, a couple of streets together forms a city and a couple cities together form a province, and so on. Individuals move more often between the houses in the street, than between the different streets in the city. One level up, the individuals move more often between the streets in their own city than between the different cities in the province, and even less often between different provinces. In ecology for instance, this multi-layer structure is natural. The hierarchical group describes this structure in a mathematical way.

The thesis consists of two parts, preceded by an introduction to population genetics in Chapter 1. In part I of the thesis we consider three versions of "the multi-colony continuum Fisher-Wright model with seed-bank". In each of these three versions individuals carry one of two gene types, denoted by type $\heartsuit$ and type $\diamondsuit$. In the first model the population in each colony consists of active and dormant individuals. The dormant individuals together form the seed-bank. The active individuals reproduce themselves, migrate between different colonies, and become dormant. The dormant individuals can only wake up: they do not reproduce themselves and also do not migrate. The second model is an extension of the first model in which the seed-bank has extra structure, which allows individuals become dormant in different ways, so that they can sleep shorter or longer. In the third model the seed-bank has the same structure as in the second model, but when an individual becomes dormant it is allowed to do so in a different colony then where it resides. In each of the three models the reproduction rate is dependent on the genetic diversity of the population.

In Chapter 2 we set up a system of stochastic differential equations for each of the three models. These systems describe the frequency of the gene type $\heartsuit$ in the population. We show that the processes described by these systems are well defined and have the Markov property. Due to the internal structure of the seed-bank in the second and third model, the time an individual spends in the seed-bank can be fat-tailed, while the Markov property of the system is maintained.

In Chapter 2 we also show that, in the special case where the diffusion function in the system of stochastic differential equations is the Fisher-Wright diffusion, there exists a dual. This dual enables us to describe the long-time behaviour of the system of stochastic differential equations.

Finally, in Chapter 2 we determine for each of the three models whether the genetic diversity is lost on the long term or not. It turns out that in the first model only the way in which the individuals migrate determines whether genetic diversity is preserved or lost. It turns out that in the second model if individuals are dormant long enough, then both the way in which the individuals migrate as well was the amount of time individuals are dormant determine whether genetic diversity is preserved or lost. The internal structure of the seed-bank can prevent the loss of genetic diversity. If in the second model individuals become dormant for a very long term, then the genetic diversity in the population is always preserved, independently of the way in which the individuals migrate. In the third model we find similar results as for the second model. Chapter 3 contains the proofs of the theorems stated in Chapter 2.

In part II of the thesis we consider the second model in the specific case where

the geographical space is the hierarchical group. The first goal of Part II is to apply the results obtained in Part I. The second goal is to analyse the genetic diversity in the so-called "hierarchical mean-field limit". In Chapter 4 we formally describe the model on the hierarchical group, state the results that follow from Part I and describe the "hierarchical mean-field limit". Chapters 5–10 contain the proofs of the theorems stated in Chapter 4. Again, it turns out that if individuals can become dormant for a long enough period, then the seed-bank can prevent the loss of genetic diversity within the population.

Acknowledgements

First of all, thank you Frank den Hollander and Andreas Greven. You always made plenty of time for me and always thought ahead about the problems arising in our project. It was a real experience for me spending long hours in the office working together with you. You formed a perfect team of supervisors: Frank always pushing the research forward even when serious problems arose, Andreas making sure that we never stopped before the picture was complete. Also, thank you both for creating opportunities for me to present our work at various workshops. Last but not least, I would like to thank you for your patience and support during all those challenging moments in the last five years.

Thanks to my colleagues at the Mathematical Institute in Leiden for helping me with all kinds of aspects during my PhD. In particular I like to thank Benthen, Conrado and Rangel for the co-organisation of the iPOD seminar. Mia and Bart, it was a pleasure to co-organise the PhD-colloquium together with you. There are too many people to name here as well, but I like to thank everyone for the enjoyable days working at the Mathematical Institute including the nice conversations during the much needed lunch and coffee breaks, which I missed so much during the corona pandemic last year.

Also, thanks a lot to the research group in Erlangen, which I met three times a year when I was visiting Andreas. In particular thanks to Andrej Depperschmidt and Max Grieshammer for their company, and thanks to Nicole Güthlein for arranging my visits to Erlangen so well.

Thanks to the NETWORKS programme and community. Many interesting workshops and training weeks were organized, but even more important, it was a real pleasure to be part of such an enthusiastic and open community. Also thanks to NETWORKS for the funding of my visits to Andreas in Erlangen.

Thanks to the DFG-SPP project Probabilistic Structures in Evolution, which enabled me to travel to many interesting meetings, where I met many inspiring people, and in particular the nice community of fellow PhD-students.

Thank you Thomas: in your bachelor thesis you did a lot of computations with great care, which helped me in my research.

Tenslotte wil ik ook graag Mayke, Erik, Kaj, Nienke, Barbera en Swinda bedanken voor de broodnodige afleiding met gezellige avondjes samen eten, spelletjes, picknicks, stranddagjes, escape rooms en wat al niet meer. Roos, Steffie en Clara heel erg bedankt voor onze gezellige kwartet avonden, dagen en weekendjes.

Dankjewel Coen, Elly en Remi, ik had me geen betere schoonfamilie kunnen voorstellen. Heel erg bedankt dat jullie altijd zo voor mij klaar staan.

Heleen en Katrien, bedankt voor jullie luisterend oor en, op zijn tijd, gedeelde verontwaardiging die je soms tijdens het schrijven van een proefschrift zo hard nodig hebt. Mama en papa, dit boekje was er zeker niet geweest als jullie niet vanaf jongs af aan hadden benadrukt hoe leuk het is om te leren en er alles aan hadden gedaan om te zorgen dat ik zo comfortabel kon studeren. En ook niet onbelangrijk voor dit proefschrift, dank voor het doorzettingsvermogen en de discipline die jullie mij al vroeg hebben bijgebracht.

Lieve Jori, bedankt voor jouw onvoorwaardelijke steun de afgelopen vijf jaar en het vertrouwen dat je altijd in mij hebt gehad. Bedankt voor al je geduld en je nooit aflatende pogingen om mij op te vrolijken als dat nodig was. Bedankt voor alle lol afgelopen anderhalf jaar aan het “keukentafelkantoor” en op de “gezelligstraat”. Lieve Jori, met jou is elke dag zo ontzettend veel leuker.

Curriculum Vitae

Margriet Oomen was born in Roosendaal in 1990. She completed her secondary education at the Gertrudiscollege in Roosendaal. In 2014 she obtained her Bachelor of Science in Mathematics and Physics at Universiteit Utrecht. In 2016 she obtained her Master of Science in Mathematics at Universiteit Utrecht. She wrote her master thesis entitled " Random Continued Fraction Expansions" under supervision of Dr. Karma Dajani. In 2016 she started her PhD-research project in the probability group of the Mathematical Institute in Leiden under supervision of Frank den Hollander and Andreas Greven.