

EURANDOM PREPRINT SERIES

2021-012

October 6, 2021

**Spatial populations with seed-bank:  
renormalisation on the hierarchical group**

A. Greven, F. den Hollander, M. Oomen  
ISSN 1389-2355

# Spatial populations with seed-bank: renormalisation on the hierarchical group

Andreas Greven<sup>1</sup>, Frank den Hollander<sup>2</sup>, Margriet Oomen<sup>3</sup>

6 October 2021

## Abstract

We consider a system of interacting diffusions labeled by a geographic space that is given by the *hierarchical group*  $\Omega_N$  of order  $N \in \mathbb{N}$ . Individuals live in colonies and are subject to resampling and migration as long as they are *active*. Each colony has a seed-bank into which individuals can retreat to become *dormant*, suspending their resampling and migration until they become active again. The migration kernel has a *hierarchical structure*: individuals hop between colonies at a rate that depends on the hierarchical distance between the colonies. The seed-bank has a *layered structure*: when individuals become dormant they acquire a colour that determines the rate at which they become active again. The latter allows us to model seed-banks whose wake-up times have a fat tail. We analyse a system of coupled stochastic differential equations that describes the population in the large-colony-size limit.

For fixed  $N \in \mathbb{N}$ , the system exhibits a dichotomy between *coexistence* (= locally multi-type equilibria) and *clustering* (= locally mono-type equilibria). To identify the *clustering regime*, i.e., the range of parameters controlling the migration and the seed-bank for which clustering prevails, we apply a necessary and sufficient criterion derived in [GHO20] that is valid for any geographic space given by a countable Abelian group endowed with the discrete topology. The diffusion function controlling the resampling does not play a role in this criterion.

We carry out a *multi-scale renormalisation analysis* in the *hierarchical mean-field limit*  $N \rightarrow \infty$ . In particular, we show that block averages on hierarchical space-time scale  $k \in \mathbb{N}$  perform a diffusion with a renormalised diffusion function that depends on  $k$ . In the clustering regime, after an appropriate scaling with  $k$ , this diffusion function converges to the Fisher-Wright diffusion function as  $k \rightarrow \infty$ , irrespective of the diffusion function controlling the resampling. Thus, the system exhibits *full universality* on large space-time scales in terms of the scaling limit. For several subclasses of parameters we identify the speed at which the scaled renormalised diffusion function converges to the Fisher-Wright diffusion as  $k \rightarrow \infty$ . This speed in turn determines the speed at which *mono-type clusters grow in space and time*. We show that the seed-bank reduces the speed compared to the model without seed-bank. This reduction is the result of a delicate interplay between migration and seed-bank, which in the limit as  $N \rightarrow \infty$  can be worked out explicitly.

*Keywords:* Fisher-Wright diffusion, resampling, migration, seed-bank, multi-scale renormalisation, universality.

*MSC 2010:* Primary 60J70, 60K35; Secondary 92D25.

*Acknowledgements:* AG was supported by the Deutsche Forschungsgemeinschaft (grant DFG-GR 876/6-1,2). FdH and MO were supported by the Netherlands Organisation for Scientific Research (NWO Gravitation Grant NETWORKS-024.002.003). FdH visited Bonn and Erlangen in the Fall of 2019 and 2020, supported by a Research Award of the Alexander von Humboldt Foundation. AG visited Leiden in February 2020, as Kloosterman Chair.

---

<sup>1</sup>) Department Mathematik, Universität Erlangen-Nürnberg, Cauerstr. 11, D-91058 Erlangen, Germany  
greven@mi.uni-erlangen.de

<sup>2</sup>) Mathematisch Instituut, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, NL  
denholla@math.leidenuniv.nl

<sup>3</sup>) Mathematisch Instituut, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, NL  
m.oomen@math.leidenuniv.nl

# Contents

<b>List of Symbols</b>	<b>4</b>
<b>1 Background, goals and outline</b>	<b>6</b>
1.1 Background . . . . .	6
1.2 Goals . . . . .	7
1.3 Outline . . . . .	7
<b>I MODEL AND MAIN RESULTS</b>	<b>8</b>
<b>2 Introduction of model and basic properties</b>	<b>8</b>
2.1 Model: geographic space $\Omega_N$ , hierachical group of order $N$ . . . . .	8
2.1.1 Hierarchical migration . . . . .	10
2.1.2 Layered seed-bank . . . . .	11
2.1.3 Resampling rate . . . . .	11
2.2 Evolution equations . . . . .	12
2.2.1 Evolution of single colonies . . . . .	12
2.2.2 Evolution of block averages . . . . .	13
2.3 Well-posedness . . . . .	14
2.4 Duality . . . . .	15
2.5 Clustering criterion . . . . .	18
2.5.1 Clustering criterion for Fisher-Wright diffusion function . . . . .	18
2.5.2 Clustering criterion for general diffusion function . . . . .	18
<b>3 Main results: <math>N &lt; \infty</math>, identification of clustering regime</b>	<b>18</b>
3.1 Finite mean wake-up time . . . . .	19
3.2 Infinite mean wake-up time . . . . .	20
3.3 Clustering regime . . . . .	21
<b>4 Main results: <math>N \rightarrow \infty</math>, renormalisation and multi-scale limit</b>	<b>22</b>
4.1 Intermezzo: Meyer-Zheng topology . . . . .	22
4.2 Main ingredients for the hierarchical multi-scale limit . . . . .	23
4.3 Hierarchical multi-scale limit theorems . . . . .	25
4.3.1 Effective process . . . . .	25
4.3.2 Full process . . . . .	27
4.4 Heuristics behind the multi-scale limit . . . . .	29
4.4.1 Evolution of the block-averages . . . . .	30
4.4.2 Limiting evolution of the block-averages . . . . .	31
4.4.3 Formation of the interaction chain . . . . .	35
<b>5 Main results <math>N \rightarrow \infty</math>: Orbit and cluster formation</b>	<b>37</b>
5.1 Orbit of renormalisation transformations . . . . .	37
5.2 Growth of mono-type clusters . . . . .	38
5.3 Rates of scaling for renormalised diffusion function . . . . .	42
<b>II PREPARATIONS AND PROOFS</b>	<b>43</b>
<b>6 Proofs: <math>N &lt; \infty</math>, identification of clustering regime</b>	<b>43</b>
6.1 Explanation of clustering criterion for infinite seed-bank . . . . .	43
6.2 Scaling of wake-up time and migration kernel for infinite seed-bank . . . . .	44
6.3 Hierarchical clustering . . . . .	46
<b>7 Preparation: <math>N \rightarrow \infty</math>, McKean-Vlasov process and mean-field system</b>	<b>47</b>
7.1 McKean-Vlasov process . . . . .	47
7.2 Mean-field system and McKean-Vlasov limit . . . . .	48
7.3 Proof of equilibrium and ergodicity . . . . .	49
7.4 Proof of McKean-Vlasov limit . . . . .	51

<b>8 Proofs: <math>N \rightarrow \infty</math>, mean-field finite-systems scheme</b>	<b>52</b>
8.1 Mean-field finite-systems scheme . . . . .	52
8.2 Abstract scheme behind finite-systems scheme . . . . .	55
<b>9 Proofs: <math>N \rightarrow \infty</math>, mean-field, proof of abstract scheme</b>	<b>60</b>
9.1 Proof of step 1. Equilibrium of the single components . . . . .	60
9.2 Proof of step 2. Convergence of the estimator . . . . .	70
9.3 Proof of step 3. Convergence of the 1-blocks in the Meyer-Zheng topology . . . . .	73
9.4 Proof of step 4. Mean-field finite-systems scheme . . . . .	74
<b>10 Preparation: <math>N \rightarrow \infty</math>, two-colour mean-field</b>	<b>76</b>
10.1 Two-colour mean-field finite-systems scheme . . . . .	77
10.2 Proof of the two-colour mean-field finite-systems scheme . . . . .	81
<b>11 Preparation: <math>N \rightarrow \infty</math>, two-level three-colour</b>	<b>95</b>
11.1 Two-level three-colour mean-field finite-systems scheme . . . . .	95
11.2 Scheme for the two-level three-colour mean-field analysis. . . . .	104
11.3 Proof of two-level three-colour mean-field finite-systems scheme . . . . .	105
11.3.1 Tightness of the 2-block estimators . . . . .	106
11.3.2 Stability of the 2-block estimators . . . . .	107
11.3.3 Tightness of the 1-block estimators . . . . .	107
11.3.4 Stability of the 1-block estimators . . . . .	108
11.3.5 Limiting evolution for the single components . . . . .	109
11.3.6 Limiting evolution of the 1-block estimator process . . . . .	113
11.3.7 Convergence of 2-block process . . . . .	123
11.3.8 State of the slow seed-banks . . . . .	125
11.3.9 Limiting evolution of the estimator processes . . . . .	134
11.3.10 Convergence in the Meyer-Zheng topology . . . . .	134
11.3.11 Proof of the two-level three-colour mean-field finite-systems scheme . . . . .	135
<b>12 Proofs: <math>N \rightarrow \infty</math>, hierarchical multi-scale limit theorems</b>	<b>135</b>
12.1 Finite-level mean-field finite-systems scheme and interaction chain . . . . .	135
12.2 Proof of the mean-field finite-systems scheme: finite-level . . . . .	140
12.3 Proof: of the hierarchical multi-scale limit theorems. . . . .	145
<b>13 Proofs: <math>N \rightarrow \infty</math>, orbit of renormalisation transformation</b>	<b>146</b>
13.1 Moment relations . . . . .	146
13.2 Iterate moment relations . . . . .	149
13.3 Clustering . . . . .	150
13.4 Dichotomy finite versus infinite seed-bank . . . . .	154
<b>A Computation of scaling coefficients</b>	<b>156</b>
A.1 Regularly varying coefficients . . . . .	157
A.2 Pure exponential coefficients . . . . .	158
A.3 Slowly varying functions . . . . .	159
<b>B Meyer-Zheng topology</b>	<b>159</b>
B.1 Basic facts about the Meyer-Zheng topology . . . . .	159
B.2 Pseudopaths of stochastic processes on a general metric separable space . . . . .	160
B.3 Proof of key lemmas . . . . .	162

## List of Symbols

$\Omega_N$	hierarchical group of order $N$ 9
$d_{\Omega_N}$	hierarchical distance on $\Omega_N$ 9
$\underline{c}$	sequence of migration rates $\underline{c} = (c_k)_{k \in \mathbb{N}_0}$ 10
$c_k$	migration rate within distance $k$ 10
$a^{\Omega_N}(\cdot, \cdot)$	migration kernel on $\Omega_N$ 10
$a_t^{\Omega_N}(\cdot, \cdot)$	time- $t$ migration kernel on $\Omega_N$ 10
$\delta$	degree of random walk 10
$\underline{K}$	sequence of relative seed-bank sizes $\underline{K} = (K_m)_{m \in \mathbb{N}_0}$ 11
$\underline{e}$	sequence of exchange parameters $\underline{e} = (e_m)_{m \in \mathbb{N}_0}$ 11
$e_m$	exchange parameter for colour $m$ 11
$K_m$	relative size of $m$ -dormant population 11
$\chi$	total exchange rate with dormant population 11
$\rho$	relative size of dormant population 11
$g$	diffusion function 11
$\mathcal{G}$	class of diffusion functions 11
$(w_\xi(t))_{t \geq 0}$	standard Brownian motion at site $\xi \in \Omega_N$ 12
$x_\xi(t)$	fraction of active individuals of type $\heartsuit$ at colony $\xi$ at time $t$ 12
$y_{\xi,m}(t)$	fraction of $m$ -dormant individuals of type $\heartsuit$ at colony $\xi$ at time $t$ 12
$S$	state space 12
$(Z^{\Omega_N}(t))_{t \geq 0}$	full process $(X^{\Omega_N}(t), Y^{\Omega_N}(t))_{t \geq 0}$ 12
$(X^{\Omega_N}(t), Y^{\Omega_N}(t))_{t \geq 0}$	full process $(x_\xi(t), (y_{\xi,m}(t))_{m \in \mathbb{N}_0})_{\xi \in \Omega_N, t \geq 0}$ 12
$\theta_x$	initial density parameter active population 12
$\theta_{y_m}$	initial density parameter $m$ -dormant population 12
$x_k^{\Omega_N}(t)$	active $k$ -block average at time $t$ 13
$y_{m,k}^{\Omega_N}(t)$	$m$ -dormant $k$ -block average at time $t$ 13
$C([0, \infty), S)$	space of continuous functions from $[0, \infty)$ to $S$ 14
$C_b([0, \infty), S)$	space of bounded continuous functions from $[0, \infty)$ to $S$ 14
$G$	generator 14
$\mathbf{F}$	domain of generator 14
$g_{\text{FW}}$	Fisher-Wright diffusion function 15
$(\mathcal{C}^{(n)}(t))_{t \geq 0}$	seed-bank coalescent 15
$b(\cdot, \cdot)$	transition kernel corresponding to lineage in dual 15
$(L(t))_{t \geq 0}$	block counting process 16
$H$	duality function 16
$z_{(\xi, R_\xi)}(t)$	fraction of individuals of type $\heartsuit$ in state $R_\xi$ in colony $\xi$ at time $t$ 17
$\sigma_k$	time lapse lineage in dual spends in active state after becoming active for $k$ -th time 17
$\tau_k$	time lapse lineage in dual spends in active state after becoming active for $k$ -th time 17
$\tau$	typical wake-up time of lineage in dual 18
$\varphi$	function that is slowly varying at infinity 20
$\gamma$	tail exponent of typical wake-up time 20
$D([0, \infty), E)$	space of càdlàg functions from $[0, \infty)$ to Polish space $E$ 22
$\psi_v$	pseudopath corresponding to measurable map $v : s \rightarrow v(s)$ from $[0, \infty)$ to $E$ 22
$\Psi$	space of all pseudopaths 22
$\vartheta_k$	initial density of $\heartsuit$ in active population and first $k + 1$ dormant populations 23
$E_k$	slowing-down constants 23
$(z_{l,(\theta, (y_{m,l})_{m \in \mathbb{N}_0})}(t))_{t \geq 0}$	full process on level $l$ with drift towards $\theta$ 23

$(z_{l,\theta}^{\text{eff}}(t))_{t \geq 0}$	effective process on level $l$ with drift towards $\theta$ 23
$\Theta^{(l),\Omega_N}(t)$	empirical average in $l$ -block at time $t$ 24
$\bar{\Theta}_x^{(l),\Omega_N}(t)$	empirical average in active $l$ -block at time $t$ 24
$\bar{\Theta}_{y_m}^{(l),\Omega_N}(t)$	empirical average in $m$ -dormant $l$ -block at time $t$ 24
$\Theta^{(l),\Omega_N}(t)$	full estimator of $l$ -block at time $t$ 24
$\Theta^{\text{eff},(l),\Omega_N}(t)$	effective estimator of the $l$ -block at time $t$ 24
$(L(N))_{N \in \mathbb{N}}$	sequence with $\lim_{N \rightarrow \infty} L(N) = \infty$ and $\lim_{N \rightarrow \infty} L(N)/N = 0$ 24
$\Gamma_{(\theta, y_l)}^{(l)}$	invariant measure for full $l$ -block average process with density $\theta$ 24
$\Gamma_{\theta}^{\text{eff},(l)}$	invariant measure for effective $l$ -block average process with density $\theta$ 24
$\mathcal{F}^{E_l, c_l, K_l, e_l}$	renormalisation transformation on level $l$ 24
$\mathcal{F}^{(k)}$	$k$ -th iterate of renormalisation transformation 24
$(M_{-l}^k)_{-l=-(k+1), -k, \dots, 0}$	interaction chain 25
$Q^{[l]}$	transition kernel of interaction chain on level $l$ 25
$(M_{-l}^{\text{eff},k})_{-l=-(k+1), -k, \dots, 0}$	effective interaction chain 25
$Q^{\text{eff},[l]}$	effective transition kernel of interaction chain on level $l$ 25
$\bar{x}_k^{\Omega_N}(t)$	first component of effective $k$ -block process 31
$Q^{(n)}$	$n$ -th composition of transition kernel of interaction chain 37
$\bar{y}^{(n)}$	initial state of level- $n$ full process 37
$A_n$	$n$ -th clustering coefficient 38
$A_m^n$	$n, m$ -th clustering coefficient 41
$[N]$	finite geographic space $\{0, 1, \dots, N-1\}$ 48
$\Delta_{\bar{\Theta}}^{[N]}(t)$	difference between active and dormant estimator at time $t$ 56
$\mu_N$	initial measure for infinite system obtained by periodic continuation of finite system 56
$(X^{\mu_N}(t), Y^{\mu_N}(t))_{t \geq 0}$	full infinite process $(x_{\xi}(t), (y_{\xi,m}(t))_{m \in \mathbb{N}_0})_{\xi \in \Omega_N, t \geq 0}$ starting from initial law $\mu_N$ 57
$L_t^0$	local time at 0 at time $t$ 64
$\bar{L}(N)$	specific sequence of times depending on $N$ 67
$G_{\dagger}^{[N]}$	$\mathcal{D}$ -semi martingale operator 75
$\Gamma_{\theta}^{\text{aux},(1)}$	equilibrium measure of auxiliary process with density $\theta$ at level 1 101
$\Omega_N^{k+1}$	hierarchical group truncated at level $k+1$ 135
$\Theta^{\text{aux},(l),\Omega_N^{k+1}}(t)$	auxiliary $l$ -block estimator process at time $t$ on level $k+1$ 138

# 1 Background, goals and outline

## 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 [BGCEK15] and [BGCKWB16] 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. [LJ11]). 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 [GHO20] 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 [GHO20] 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 [GHO20] 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 [GHO21] 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.

## 1.2 Goals

In the present paper we take as geographic space the *hierarchical group*  $\Omega_N$  of order  $N$ . The reason for this choice is that  $\Omega_N$  allows for more detailed computations. At the same time, migration on  $\Omega_N$  can be used to *approximate* migration on  $\mathbb{Z}^d$  in the *hierarchical mean-field limit*  $N \rightarrow \infty$ . 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  $N \rightarrow \infty$  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 [GHO20] to  $\Omega_N$  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  $N \rightarrow \infty$ . 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 [DG93b] and [DG93a] 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 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 [DGV95]), 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 [DG93a, DG93b, DGV95, DG96] and [BCGdH95, BCGdH97, dH01]. ■

## 1.3 Outline

The present paper consist of two parts:

- **Part I: Model and main results.** Sections 2–5 collect the main propositions and theorems. In Section 2 we define the hierarchical model and state some basic properties: the *well-posedness* of the associated martingale problem (Proposition 2.6), the *duality relation* (Proposition 2.7), and the *clustering criterion* via duality (Proposition 2.12). These properties were all derived in [GHO20]. In Section 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 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 3.3)). In Section 4 we state our main results for  $N \rightarrow \infty$ ,



the hierarchical mean-field limit. In particular, we introduce *block averages on successive hierarchical space-time scales*, analyse their limiting dynamics (Theorems 4.2 and 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 different hierarchical levels interact with each other. In Section 5 we identify the *orbit of the renormalisation transformation* in the clustering regime (Theorem 5.1), identify the rate of scaling for the renormalised diffusion function (Theorem 5.3), and link this scaling to the rate of growth of mono-type clusters.

- **Part II: Preparations and proofs.** Sections 6–13 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 Section 6 we prove our main results for  $N < \infty$ . In Sections 7–9 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 Sections 10–12 we consider extensions of the mean-field model (consisting of finitely hierarchies), which serve as further preparation. In Section 12 we use the results in Sections 7–12 to deal with the full hierarchical model (consisting of infinitely many hierarchies), and prove our main results for  $N \rightarrow \infty$ . In Section 13 we analyse the orbit of the renormalisation transformation controlling the multi-scaling. Appendix A contains a technical computation needed for the identification of the clustering regime. Appendix B 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.

## Part I

# MODEL AND MAIN RESULTS

## 2 Introduction of model and basic properties

Section 2.1 introduces the model ingredients, Section 2.2 gives the evolution equations, Section 2.3 states the well-posedness, Section 2.4 introduces the dual and states the duality relation, while Section 2.5 formulates the dichotomy between clustering versus coexistence in terms of the dual.

### 2.1 Model: geographic space $\Omega_N$ , hierarchical group of order $N$

**Single colony.** Our building block is the single-colony Fisher-Wright model with seed-bank defined in [BGCKWB16]. 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 active 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 become active, i.e., the active and the dormant population *exchange* individuals (see Fig. 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 dormant population evolve only through exchange with the active population. It was shown in [BGCKWB16] 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:

$$(2.1) \quad \begin{aligned} dx(t) &= Ke[y(t) - x(t)] dt + \sqrt{x(t)(1-x(t))} dw(t), \\ dy(t) &= e[x(t) - y(t)] dt. \end{aligned}$$

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. 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 [BGCKWB16] and in [GHO20, Appendix A].

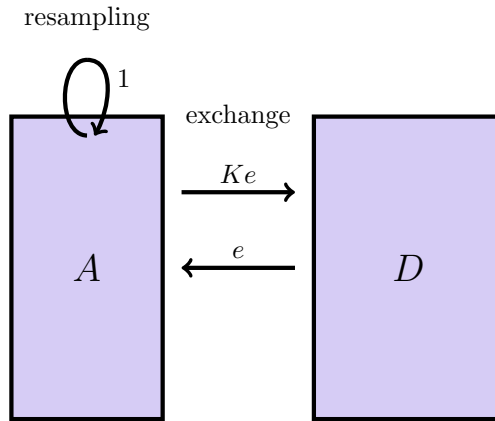


Figure 1: Active individuals resample at rate 1. Active and dormant individuals exchange at rate  $e$ . The extra factor  $K$  arises from the fact that the dormant population is  $K$  times as large as the active population. Dormant individuals suspend resampling.

**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  $(\mathbb{N}_0 = \mathbb{N} \cup \{0\})$

$$(2.2) \quad \Omega_N = \left\{ \xi = (\xi_k)_{k \in \mathbb{N}_0} : \xi_k \in \{0, 1, \dots, N-1\}, \sum_{k \in \mathbb{N}_0} \xi_k < \infty \right\},$$

which with addition modulo  $N$  becomes the hierarchical group of order  $N$  (see Fig. 2). The *hierarchical distance* on  $\Omega_N$  is defined by

$$(2.3) \quad d_{\Omega_N}(\xi, \eta) = d_{\Omega_N}(0, \xi - \eta) = \min \{k \in \mathbb{N}_0 : \xi_l = \eta_l \forall l \geq k\}, \quad \xi, \eta \in \Omega_N,$$

and is an ultra-metric, i.e.,

$$(2.4) \quad d_{\Omega_N}(\xi, \eta) \leq \max \{d_{\Omega_N}(\xi, \zeta), d_{\Omega_N}(\eta, \zeta)\} \quad \forall \xi, \eta, \zeta \in \Omega_N.$$

The choice of  $\Omega_N$  as geographic space plays an important role for population models, and was first exploited in [SF83] 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  $\xi$ :  $\xi_0$  is the ‘house’,  $\xi_1$  is the ‘street’,  $\xi_2$  is the ‘village’,  $\xi_3$  is the ‘province’,  $\xi_4$  is the ‘country’, and so on. To describe the system on the hierarchical group we need three ingredients:

1. Hierarchical migration (Section 2.1.1).
2. Layered seed-bank (Section 2.1.2).
3. Resampling rate (Section 2.1.3).

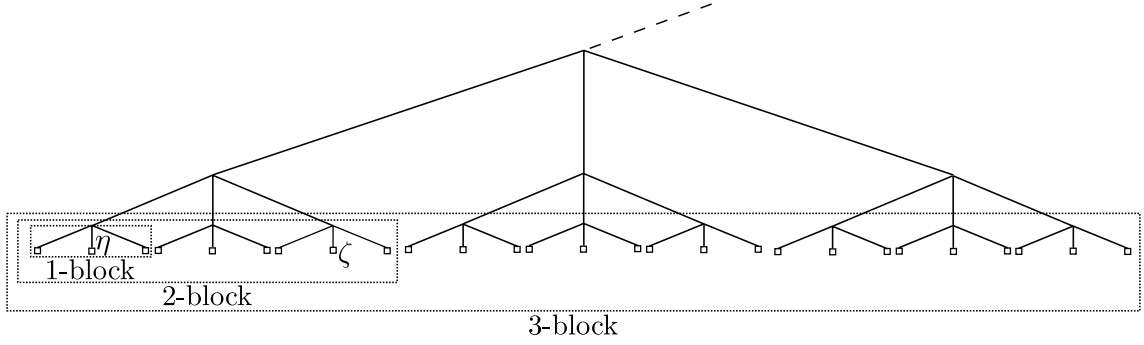


Figure 2: Close-ups of a 1-block, a 2-block and a 3-block in the hierarchical group of order  $N = 3$ . The elements of the group are the leaves of the tree (indicated by  $\square$ 's). The hierarchical distance between two elements in the group is the graph distance to the most recent common ancestor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

### 2.1.1 Hierarchical migration

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

$$(2.5) \quad \underline{c} = (c_k)_{k \in \mathbb{N}_0} \in (0, \infty)^{\mathbb{N}_0}$$

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  $a^{\Omega_N}(\cdot, \cdot)$  is therefore given by

$$(2.6) \quad a^{\Omega_N}(\eta, \xi) = \sum_{k \geq d_{\Omega_N}(\eta, \xi)} \frac{c_{k-1}}{N^{k-1}} \frac{1}{N^k}, \quad \eta, \xi \in \Omega_N, \eta \neq \xi, \quad a^{\Omega_N}(\eta, \eta) = 0, \quad \eta \in \Omega_N.$$

Throughout the paper, we assume that

$$(2.7) \quad \limsup_{k \rightarrow \infty} \frac{1}{k} \log c_k < \log N.$$

This guarantees that the total migration rate per individual is finite. Indeed, note that for every  $\eta \in \Omega_N$ ,

$$(2.8) \quad \sum_{\xi \in \Omega_N} a^{\Omega_N}(\eta, \xi) = \sum_{\xi \in \Omega_N} \sum_{k \geq d_{\Omega_N}(\eta, \xi)} \frac{c_{k-1}}{N^{2k-1}} = \sum_{k \in \mathbb{N}} \left[ \sum_{\xi \in \Omega_N} \mathbf{1}_{\{d_{\Omega_N}(\eta, \xi) \leq k\}} \right] \frac{c_{k-1}}{N^{2k-1}} = \sum_{k \in \mathbb{N}} \frac{c_{k-1}}{N^{k-1}},$$

which is finite because of (2.7).

**Remark 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 [DGW05])

$$(2.9) \quad \delta = \sup \left\{ \zeta \in (-1, \infty) : \int_0^\infty dt t^\zeta a_t^{\Omega_N}(0, 0) < \infty \right\}.$$

The degree is said to be  $\delta^+$ , respectively,  $\delta^-$  when the integral is finite, respectively, infinite at the degree. If  $\delta > 0$ , then  $\delta$  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 (2.9) that includes a slowly varying function in front of the power  $t^\zeta$ . However, such an extension appears not to have been explored in the literature.) ■

By playing with  $\underline{c}$  and letting  $N \rightarrow \infty$ , 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 [DGV95], [DG96], [DG14], [GdHKK14], [GdHK18].

### 2.1.2 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

$$(2.10) \quad \begin{aligned} \underline{K} &= (K_m)_{m \in \mathbb{N}_0} \in (0, \infty)^{\mathbb{N}_0}, \\ \underline{e} &= (e_m)_{m \in \mathbb{N}_0} \in (0, \infty)^{\mathbb{N}_0}, \end{aligned}$$

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.,

$$(2.11) \quad K_m = \frac{\text{size } m\text{-dormant population}}{\text{size active population}}.$$

- 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

$$(2.12) \quad \limsup_{m \rightarrow \infty} \frac{1}{m} \log(K_m e_m) < \log N.$$

This guarantees that the total rate of exchange per individual, given by

$$(2.13) \quad \chi = \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m},$$

is finite. On the other hand, the relative size of the dormant population with respect to the active population

$$(2.14) \quad \rho = \sum_{m \in \mathbb{N}_0} K_m$$

can be either finite or infinite. We will see that  $\rho < \infty$  and  $\rho = \infty$  represent two *different regimes*.

### 2.1.3 Resampling rate

To describe the resampling we use a diffusion function  $g$  that is taken from the set

$$(2.15) \quad \mathcal{G} = \left\{ g(x): [0, 1] \rightarrow [0, \infty): g(0) = g(1) = 0, g(x) > 0 \forall x \in (0, 1), g \text{ Lipschitz} \right\},$$

and think of  $h(x) = g(x)/x(1-x)$  as the rate of resampling at type frequency  $x$ . The choice  $g = dg_{\text{FW}}$ ,  $d \in (0, \infty)$ , with  $g_{\text{FW}}(x) = x(1-x)$ ,  $x \in [0, 1]$ , corresponds to Fisher-Wright resampling at rate  $d$ . We use a collection of independent Brownian motions

$$(2.16) \quad W = ((w_\xi(t))_{t \geq 0})_{\xi \in \Omega_N}$$

to describe the fluctuations of the type frequencies caused by the resampling in each colony.

## 2.2 Evolution equations

### 2.2.1 Evolution of single colonies

With the above three ingredients, we can now describe the evolution of the system. For  $\xi \in \Omega_N$ , define

$$(2.17) \quad \begin{aligned} x_\xi(t) &= \text{the fraction of active individuals of type } \heartsuit \text{ at colony } \xi \text{ at time } t, \\ y_{\xi,m}(t) &= \text{the fraction of } m\text{-dormant individuals of type } \heartsuit \text{ at colony } \xi \text{ and time } t. \end{aligned}$$

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

$$(2.18) \quad S = \mathfrak{s}^{\Omega_N}.$$

Our object of interest is the random process taking values in  $S$ , written

$$(2.19) \quad (Z^{\Omega_N}(t))_{t \geq 0} = (X^{\Omega_N}(t), Y^{\Omega_N}(t))_{t \geq 0}, \quad (X^{\Omega_N}(t), Y^{\Omega_N}(t)) = (x_\xi(t), (y_{\xi,m}(t))_{m \in \mathbb{N}_0})_{\xi \in \Omega_N},$$

whose components evolve according to the following SSDE (= system of stochastic differential equations):

$$(2.20) \quad \begin{aligned} dx_\xi(t) &= \sum_{\eta \in \Omega_N} a^{\Omega_N}(\xi, \eta) [x_\eta(t) - x_\xi(t)] dt + \sqrt{g(x_\xi(t))} dw_\xi(t) \\ &\quad + \sum_{m \in \mathbb{N}_0} \frac{K_m e_m}{N^m} [y_{\xi,m}(t) - x_\xi(t)] dt, \\ dy_{\xi,m}(t) &= \frac{e_m}{N^m} [x_\xi(t) - y_{\xi,m}(t)] dt, \quad m \in \mathbb{N}_0, \quad \xi \in \Omega_N. \end{aligned}$$

The first term in the first equation describes the evolution of the active population at colony  $\xi$  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  $\xi$  (see Fig. 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

$$(2.21) \quad \begin{aligned} \mathcal{L}(X^{\Omega_N}(0), Y^{\Omega_N}(0)) &= \mu^{\otimes \Omega_N} \\ \text{with } \mathbb{E}^\mu[x_\xi(0)] &= \theta_x, \mathbb{E}^\mu[y_{\xi,m}(0)] = \theta_{y_m} \text{ with } \lim_{m \rightarrow \infty} \theta_{y_m} = \theta \text{ for some } \theta \in [0, 1]. \end{aligned}$$

The last assumption in (2.21), which in [GHO20] was referred to as  $\mu$  being *colour regular*, guarantees that for finite  $N$  the system in (2.20) converges to an ergodic equilibrium.

**Remark 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. ■

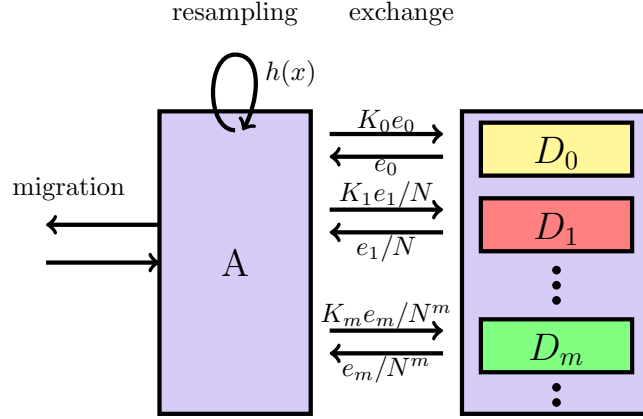


Figure 3: Active individuals ( $A$ ) are subject to migration, resampling and exchange with dormant individuals ( $D$ ). When active individuals become dormant they are assigned a colour ( $D_m$ ,  $m \in \mathbb{N}_0$ ), which they lose when they become active again. The resampling rate in the active state at type- $\heartsuit$  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).

### 2.2.2 Evolution of block averages

The choice of the migration kernel in (2.6) 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 2.3. [Block averages]** For  $k \in \mathbb{N}_0$ , let  $B_k(0) = \{\eta \in \Omega_N : d_{\Omega_N}(0, \eta) \leq k\}$  denote the  $k$ -block around 0. Define the  $k$ -block average around 0 at time  $N^k t$  by

$$(2.22) \quad \begin{aligned} x_k^{\Omega_N}(t) &= \frac{1}{N^k} \sum_{\eta \in B_k(0)} x_\eta(N^k t), \\ y_{m,k}^{\Omega_N}(t) &= \frac{1}{N^k} \sum_{\eta \in B_k(0)} y_{\eta,m}(N^k t), \quad m \in \mathbb{N}_0. \end{aligned}$$

The  $k$ -block average represents the dynamics of the system on space-time scale  $k$ . ■

By translation invariance of the SSDE in (2.20), 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 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  $\Omega_N$  for the block averages because they are functionals of the system of components (recall Remark 2.2). ■

Using Definition 2.3 and inserting the specific choice of the migration kernel defined in (2.6), we can rewrite (2.20) for  $\xi = 0$  as follows (0-blocks are single components):

$$(2.23) \quad \begin{aligned} dx_0^{\Omega_N}(t) &= \sum_{l \in \mathbb{N}} \frac{c_{l-1}}{N^{l-1}} [x_l^{\Omega_N}(N^{-l}t) - x_0^{\Omega_N}(t)] dt + \sqrt{g(x_0^{\Omega_N}(t))} dw(t) \\ &\quad + \sum_{m \in \mathbb{N}_0} \frac{K_m e_m}{N^m} [y_{m,0}^{\Omega_N}(t) - x_0^{\Omega_N}(t)] dt, \\ dy_{m,0}^{\Omega_N}(t) &= \frac{e_m}{N^m} [x_0^{\Omega_N}(t) - y_{m,0}^{\Omega_N}(t)] dt, \quad m \in \mathbb{N}_0. \end{aligned}$$

From (2.23) 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 (2.22)):

$$(2.24) \quad \begin{aligned} dx_k^{\Omega_N}(t) &= \sum_{l \in \mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} [x_{k+l}^{\Omega_N}(N^{-l}t) - x_k^{\Omega_N}(t)] dt + \sqrt{\frac{1}{N^k} \sum_{i \in B_k(0)} g(x_i(N^k t))} dw_k(t) \\ &+ \sum_{m \in \mathbb{N}_0} N^k \frac{K_m e_m}{N^m} [y_{m,k}^{\Omega_N}(t) - x_k^{\Omega_N}(t)] dt, \\ dy_{m,k}^{\Omega_N}(t) &= N^k \frac{e_m}{N^m} [x_k^{\Omega_N}(t) - y_{m,k}^{\Omega_N}(t)] dt, \quad m \in \mathbb{N}_0. \end{aligned}$$

To deduce these equations from (2.20), we sum over  $\xi \in B_k$ , speed up time by a factor  $N^k$ , insert the specific choice of the migration kernel in (2.6), and use the standard scaling properties of Brownian motion:  $w(ct) \stackrel{d}{=} \sqrt{c} w(t)$  and  $\sqrt{a}w(t) + \sqrt{b}w'(t) \stackrel{d}{=} \sqrt{a+b}w''(t)$ , with  $w(t)$  and  $w'(t)$  independent Brownian motions, and with  $\stackrel{d}{=}$  denoting equality in distribution. This computation is spelled out in Section 12.

**Remark 2.5.** The *block averages* and their evolution equations in (2.24) will be key objects in the analysis of the hierarchical mean-field limit  $N \rightarrow \infty$ . We will see that the limit  $N \rightarrow \infty$  brings about considerable simplifications. In Section 4 we discuss these simplifications in detail. In particular, a *complete separation of space-time scales* takes place, 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. ■

### 2.3 Well-posedness

Let  $C([0, \infty), S)$  denote the space of continuous function from  $[0, \infty)$  to the state space  $S$  and let  $C_b([0, \infty), S)$  be the space of bounded continuous functions from  $[0, \infty)$  to  $S$ . The generator of the system in (2.20) is given by

$$(2.25) \quad \begin{aligned} G &= \sum_{\xi \in \Omega_N} \left( \sum_{\eta \in \Omega_N} a^{\Omega_N}(\xi, \eta) [x_\eta(t) - x_\xi(t)] \frac{\partial}{\partial x_\xi} + \frac{1}{2} g(x_\xi(t)) \frac{\partial^2}{\partial x_\xi^2} \right. \\ &\quad \left. + \sum_{m \in \mathbb{N}_0} \left[ \frac{K_m e_m}{N^m} [y_{\xi,m}(t) - x_\xi(t)] \frac{\partial}{\partial x_\xi} + \frac{e_m}{N^m} [x_\xi(t) - y_{\xi,m}(t)] \frac{\partial}{\partial y_{\xi,m}} \right] \right). \end{aligned}$$

Let

$$(2.26) \quad \mathbf{F} = \left\{ f \in C_b([0, \infty), S) : f \text{ depends on finitely many components and is twice continuously differentiable in each component} \right\}.$$

#### Proposition 2.6. [Well-posedness]

- (a) *The SSDE in (2.20) has a unique strong solution in  $C([0, \infty), S)$ , whose law is the unique solution of the  $(G, \mathbf{F}, \delta_u)$ -martingale problem for all  $u \in S$ .*
- (b) *The process starting from  $u \in S$  is Feller and strong Markov. Consequently, the SSDE in (2.20) defines a unique Borel Markov process starting from any initial law on  $S$ .*

*Proof.* Comparing with what is called model 2 in [GHO20], we see that the Abelian group is chosen as in (2.2), the transition kernel is chosen as in (2.6), 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 [SS80], in the same way as shown in the proof of [GHO20, Theorem 2.1]. □

Henceforth we write  $\mathbb{P}$  and  $\mathbb{E}$  to denote probability and expectation with respect to the random process in (2.19).

## 2.4 Duality

If  $g = dg_{\text{FW}}$ , 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 [GHO20, 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 corresponding block-counting process counts the number of partition elements that are left when we travel backwards in time.

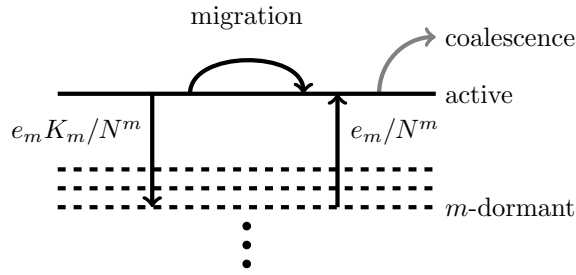


Figure 4: Transition scheme for an ancestral lineage in the dual, which moves according to the transition kernel  $b(\cdot, \cdot)$  defined in (2.30). Two active ancestral lineages that are at the same colony coalesce at rate  $d$ .

Formally, the spatial seed-bank coalescent is described as follows. Let  $\mathbb{S} = \Omega_N \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$  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  $\pi$ , where

$$(2.27) \quad \begin{aligned} \pi &= ((\pi_1, \eta_1), \dots, (\pi_{\bar{n}}, \eta_{\bar{n}})), \quad \bar{n} = |\pi|, \\ \pi_\ell &\subset \{1, \dots, n\}, \quad \{\pi_1, \dots, \pi_{\bar{n}}\} \text{ is a partition of } \{1, \dots, n\}, \\ \eta_\ell &\in \mathbb{S}, \quad \ell \in \{1, \dots, \bar{n}\}, \quad 1 \leq \bar{n} \leq n. \end{aligned}$$

A marked partition element  $(\pi_\ell, \eta_\ell)$  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

$$(2.28) \quad (\mathcal{C}^{(n)}(t))_{t \geq 0},$$

and starts from

$$(2.29) \quad \mathcal{C}^{(n)}(0) = \pi(0), \quad \pi(0) = \{(\{1\}, \eta_{\ell_1}), \dots, (\{n\}, \eta_{\ell_n})\}, \quad \eta_{\ell_1}, \dots, \eta_{\ell_n} \in \mathbb{S}.$$

The  $n$ -spatial seed-bank coalescent is the Markov process that evolves according to the following two rules (see Figs. 4–5).

1. Each partition element moves independently of all other partition elements according to the transition kernel  $b(\cdot, \cdot)$

$$(2.30) \quad b((\xi, R_\xi), (\eta, R_\eta)) = \begin{cases} a^{\Omega_N}(\xi, \eta), & \text{if } R_\xi = R_\eta = A, \\ K_m \frac{e_m}{N^m}, & \text{if } \xi = \eta, R_\xi = A, R_\eta = D_m, \text{ for some } m \in \mathbb{N}_0, \\ \frac{e_m}{N^m}, & \text{if } \xi = \eta, R_\xi = D_m, R_\eta = A, \text{ for some } m \in \mathbb{N}_0, \\ 0, & \text{otherwise,} \end{cases}$$

where  $a^{\Omega_N}(\cdot, \cdot)$  is the migration kernel defined in (2.6),  $K_m$ ,  $m \in \mathbb{N}_0$  are the relative sizes of the  $m$ -dormant population and the active population defined in (2.11), and  $e_m$ ,  $m \in \mathbb{N}_0$  are the



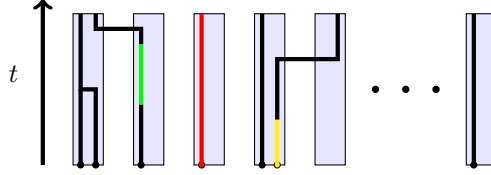


Figure 5: Picture of the evolution of lineages in the spatial coalescent. The purple blocks depict the colonies, the black lines the active lineages, and the coloured lines the dormant lineages. Blue lineages can migrate. Two black lineages can coalesce when they are at the same colony. Red dormant lineages 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*.

coefficients controlling the exchange between the active and the dormant population defined in (2.10). Thus, an active partition element migrates according to the transition kernel  $a^{\Omega_N}(\cdot, \cdot)$  and becomes  $m$ -dormant at rate  $K_m \frac{\epsilon_m}{N^m}$ , while an  $m$ -dormant partition element can only become active and does so at rate  $\frac{\epsilon_m}{N^m}$ .

- 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 gives a schematic overview of the possible transitions of a single lineage, while Fig. 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  $n \rightarrow \infty$ . This object is well-defined by Kolmogorov's extension theorem (see [BGCKWB16, 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.,

$$\begin{aligned}
 L(t) &= (L_{(\xi, A)}(t), (L_{(\xi, D_m)}(t))_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}, \\
 (2.31) \quad L_{(\xi, A)}(t) &= L_{(\xi, A)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar{n}} \mathbf{1}_{\{\eta_\ell(t) = (\xi, A)\}}, \\
 L_{(\xi, D_m)}(t) &= L_{(\xi, D_m)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar{n}} \mathbf{1}_{\{\eta_\ell(t) = (\xi, D_m)\}}, \quad m \in \mathbb{N}_0.
 \end{aligned}$$

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  $(\xi, R_\xi) \in \Omega_N \times \{A, (D_m)_{m \in \mathbb{N}_0}\} \setminus \{(\eta, R_\eta)\}$ , and 1 at the site  $(\eta, R_\eta)$ . From the evolution of  $\mathcal{C}^{(n)}(t)$  described below (2.28) we see that the block-counting process has the following transition kernel:

$$\begin{aligned}
 (2.32) \quad & (m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} \rightarrow \\
 & \begin{cases} (m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} - \delta_{(\eta, A)} + \delta_{(\zeta, A)}, & \text{at rate } m_\eta a(\eta, \zeta) \text{ for } \eta, \zeta \in \Omega_N, \\ (m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} - \delta_{(\eta, A)}, & \text{at rate } d \binom{m_\eta}{2} \text{ for } \eta \in \Omega_N, \\ (m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} - \delta_{(\eta, A)} + \delta_{(\eta, D_m)}, & \text{at rate } m_\eta K_m \frac{\epsilon_m}{N^m} \text{ for } \eta \in \Omega_N, \\ (m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} + \delta_{(\eta, A)} - \delta_{(\eta, D_m)}, & \text{at rate } n_{\eta, m} \frac{\epsilon_m}{N^m} \text{ for } \eta \in \Omega_N. \end{cases}
 \end{aligned}$$

The process  $(Z(t))_{t \geq 0}$  defined in (2.20) is dual to the block-counting process  $(L(t))_{t \geq 0}$  with *duality function*  $H: S \times S' \rightarrow \mathbb{R}$  defined by

$$(2.33) \quad H\left((x_\xi, (y_{\xi, m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}, (m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}\right) = \prod_{\xi \in \Omega_N} x_\xi^{m_\xi} \prod_{m \in \mathbb{N}_0} y_{\xi, m}^{n_{\xi, D_m}}.$$

**Proposition 2.7. [Duality relation].** *Let  $H$  be as in (2.33). Then, for all  $(x_\xi, (y_{\xi, m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}$*

$\in S$  and  $(m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} \in S'$ ,

$$(2.34) \quad \begin{aligned} & \mathbb{E}_{(x_\xi, (y_{\xi, m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}} \left[ H \left( (x_\xi(t), (y_{\xi, m}(t))_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}, (m_\xi, n_\xi)_{\xi \in \Omega_N} \right) \right] \\ &= \mathbb{E}_{(m_\xi, (n_{\xi, D_m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}} \left[ H \left( (x_\xi, (y_{\xi, m})_{m \in \mathbb{N}_0})_{\xi \in \Omega_N}, (L_{(\xi, A)}(t), (L_{(\xi, D_m)}(t))_{m \in \mathbb{N}_0})_{\xi \in \Omega_N} \right) \right] \end{aligned}$$

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

Proposition 2.7 was proved in [GHO20, Section 2.4]. Since the duality function  $H$  captures all the mixed moments of  $(Z(t))_{t \geq 0}$ , the duality relation is that of a *moment dual*.

**Remark 2.8. [Duality relation in terms of the effective geographic space]** Interpreting  $(Z(t))_{t \geq 0}$  as a process on the effective geographic space  $\mathbb{S} = \Omega_N \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$ , we can rewrite (2.20) as

$$(2.35) \quad \begin{aligned} dz_{(\xi, R_\xi)}(t) &= \sum_{(\xi, R_\xi) \in \mathbb{S}} b((\xi, R_\xi), (\eta, R_\eta)) [z_{(\eta, R_\eta)}(t) - z_{(\xi, R_\xi)}(t)] dt \\ &\quad + \mathbf{1}_{\{R_\xi = A\}} \sqrt{g(z_{(\xi, R_\xi)}(t))} dw_\xi(t), \quad (\xi, R_\xi) \in \mathbb{S}, \end{aligned}$$

where  $b(\cdot, \cdot)$  is the transition kernel defined in (2.30). If  $g = dg_{\text{FW}}$ , then we can write its dual process as follows. Let  $(L(t))_{t \geq 0} = (L(\mathcal{C}(t))_{t \geq 0})$  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.,

$$(2.36) \quad \begin{aligned} L(t) &= (L_{(\xi, R_\xi)}(t))_{(\xi, R_\xi) \in \mathbb{S}}, \\ L_{(\xi, R_\xi)}(t) &= L_{(\xi, R_\xi)}(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} \mathbf{1}_{\{\eta_\ell(t) = (\xi, R_\xi)\}}. \end{aligned}$$

Rewrite the duality function  $H$  in (2.33) as

$$(2.37) \quad H \left( (z_{(\xi, R_\xi)}, l_{(\xi, R_\xi)})_{(\xi, R_\xi) \in \mathbb{S}} \right) = \prod_{(\xi, R_\xi) \in \mathbb{S}} z_{(\xi, R_\xi)}^{l_{(\xi, R_\xi)}}.$$

Then, for  $z \in S$  and  $l \in S'$ , the duality relation in (2.34) reads

$$(2.38) \quad \mathbb{E}_{z_{(\xi, R_\xi)}} [H(z_{(\xi, R_\xi)}(t), l_{(\xi, R_\xi)})] = \mathbb{E}_{l_{(\xi, R_\xi)}} [H(z_{(\xi, R_\xi)}, L_{(\xi, R_\xi)}(t))].$$

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 (2.30). Interpreting the duality relation in terms of the geographic space  $\Omega_N$ , we see that an ancestral lineage is a random walk on  $\Omega_N$ , 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. ■

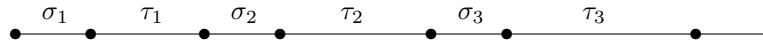


Figure 6: Renewal process induced by a lineage in the dual moving according to the transition kernel  $b(\cdot, \cdot)$ . For  $k \in \mathbb{N}$ ,  $\sigma_k$  denotes the  $k$ th active period and  $\tau_k$  the  $k$ th dormant period.

**Remark 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. 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 (2.13))

$$(2.39) \quad \mathbb{P}(\sigma_1 > t) = e^{-\chi t}, \quad \mathbb{P}(\tau_1 > t) = \sum_{m \in \mathbb{N}_0} \frac{K_m \frac{c_m}{N^m}}{\chi} e^{-\frac{c_m}{N^m} t}, \quad t \geq 0.$$

**Remark 2.10.** [Wake up times] The renewal process in Fig. 6 is key to understanding the long-time behaviour of the model (as we will see in Section 3). Note that

$$(2.40) \quad \tau = \tau_1$$

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 (2.14)), then  $\tau$  may have a fat-tail (examples are given in Section 3). In other words, *the internal structure of the seed-bank allows us to model fat-tailed wake-up times without losing the Markov property of the evolution*.  $\square$

Note that even when there is no dual, i.e.,  $g \in \mathcal{G}$  with  $g \neq dg_{\text{FW}}$ , we can still define  $\tau$  by (2.40), since  $\tau_1$  in (2.39) 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  $\tau$  as the typical wake-up time of an individual in the population.  $\blacksquare$

## 2.5 Clustering criterion

### 2.5.1 Clustering criterion for Fisher-Wright diffusion function

In [GHO20] 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 2.11.** [Colour regular initial measures] We say that a translation invariant initial measure  $\mu(0)$  is *colour regular* when

$$(2.41) \quad \lim_{m \rightarrow \infty} \mathbb{E}_{\mu(0)}[y_{0,m}] \quad \text{exists.}$$

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.  $\blacksquare$

The key clustering criterion is the following.

**Proposition 2.12.** [Clustering criterion] *Suppose that  $\mu(0)$  is translation invariant. If  $\rho = \infty$  (recall (2.14)), then additionally suppose that  $\mu(0)$  is colour regular. Let  $d \in (0, \infty)$ . Then the system with  $g = dg_{\text{FW}}$  clusters if and only if in the dual two partition elements coalesce with probability 1.*

The idea behind Theorem 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 [GHO20, Section 4.3]. The proof is valid for any geographic space given by a countable Abelian group endowed with the discrete topology, of which  $\Omega_N$  is an example.

### 2.5.2 Clustering criterion for general diffusion function

For  $g \in \mathcal{G}$  with  $g \neq dg_{\text{FW}}$  no dual is available and hence we cannot use the clustering criterion in Proposition 2.12. However, as shown in [GHO20], we can argue by duality comparison arguments (see [GHO20, Lemma 5.5 and Lemma 6.3]) that the system evolving according to (2.20) with  $g \in \mathcal{G}$  clusters if and only if the system with  $g = dg_{\text{FW}}$  for some  $d \in (0, \infty)$  clusters. In particular, for  $g = dg_{\text{FW}}$ ,  $d \in (0, \infty)$ , whether or not the system clusters does not depend on the resampling rate  $d$ .

## 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 2.12 is met. In [GHO20, 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  $\chi$  in (2.13),  $\rho$  in (2.14) and  $\tau$  in (2.40). From (2.39) it follows that

$$(3.1) \quad \mathbb{E}[\tau] = \sum_{m \in \mathbb{N}_0} \frac{K_m}{\chi} = \frac{\rho}{\chi},$$

and hence the mean wake-up time is finite if  $\rho < \infty$  and infinite if  $\rho = \infty$ . In Section 3.1 we look at  $\rho < \infty$  and in Section 3.2 at  $\rho = \infty$ . In Section 3.3 we summarise our findings and identify the clustering regime.

### 3.1 Finite mean wake-up time

Suppose that the system evolving according to (2.20) has a translation invariant initial measure  $\mu(0)$  with density  $\theta \in (0, 1)$ . Then [GHO20, Theorem 3.3] says that for  $\rho < \infty$  clustering occurs if and only if

$$(3.2) \quad \int_1^\infty dt a_t^{\Omega_N}(0, 0) = \infty.$$

It is known that (3.2) holds for the hierarchical migration defined in (2.6) if and only if [DGW05, Section 3]

$$(3.3) \quad \sum_{k \in \mathbb{N}_0} \frac{1}{c_k} = \infty.$$

Hence, for  $\rho < \infty$ , the clustering criterion depends on the migration kernel *only* and the seed-bank has no effect.

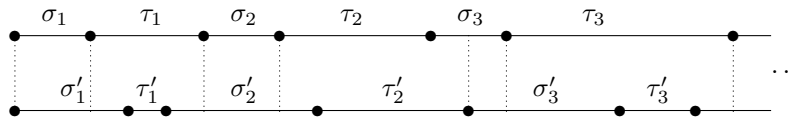


Figure 7: Successive periods during which the two random walks are active and dormant (recall (2.39) and Fig. 6). The time lapses between successive pairs of dotted lines represent periods of *joint activity*.

In view of Proposition 2.12, if  $g = dg_{FW}$ , 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 (3.2) 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 (2.30)). Looking at the renewal process induced by the dual process (recall Remark 2.9 and Fig. 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

$$(3.4) \quad \int_1^\infty dt \left( \frac{1}{1+\rho} \right)^2 a_{2\frac{1}{1+\rho}t}(0, 0),$$

By Polya's argument, if the integral in (3.4) is infinite, then two lineages in the dual meet infinitely often while being active. After a variable transformation, (3.4) becomes the integral in (3.2) up to a constant. (For a formal proof of the criterion in (3.2), we refer to [GHO20].) By the above argument, we can think of the integral in (3.2) 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 2.5.2.

In terms of the degree of the random walk (recall Remark 2.1), (3.2) corresponds to hierarchical migration with degree  $0^-$ . The same criterion as in (3.2) was found in [FG96] 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.

### 3.2 Infinite mean wake-up time

If  $\rho = \infty$ , then the seed-bank does affect the dichotomy. To apply the criterion in [GHO20, Theorem 3.3], we assume that the system evolving according to (2.20) 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 [GHO20] for  $\rho = \infty$  applies to wake-up times  $\tau$  (recall (2.40)) of the form

$$(3.5) \quad \frac{\mathbb{P}(\tau \in dt)}{dt} \sim \varphi(t) t^{-(1+\gamma)}, \quad t \rightarrow \infty, \quad \gamma \in (0, 1],$$

with  $\varphi$  slowly varying at infinity. Define

$$(3.6) \quad \hat{\varphi}(t) = \begin{cases} \varphi(t), & \gamma \in (0, 1), \\ \mathbb{E}[\tau \wedge t] & \gamma = 1. \end{cases}$$

As shown in [BGT87, Section 1.3], every slowly varying function  $\varphi$  may be assumed to be infinitely differentiable and to be represented by the integral

$$(3.7) \quad \varphi(t) = \exp \left[ \int_{(\cdot)}^t \frac{du}{u} \psi(u) \right]$$

for some  $\psi: [0, \infty) \rightarrow \mathbb{R}$  such that  $\lim_{u \rightarrow \infty} |\psi(u)| = 0$ . From (3.5) we see that  $\hat{\varphi}(t)$  is also slowly varying. If we assume that  $|\psi(u)| \leq C/\log u$  for some  $C < \infty$ , then the system clusters if and only if (see [GHO20, Section 3.2])

$$(3.8) \quad \int_{(\cdot)}^{\infty} dt \hat{\varphi}(t)^{-1/\gamma} t^{-(1-\gamma)/\gamma} \hat{a}_t(0, 0) = \infty.$$

Note that  $\gamma$  in (3.5) is the *tail exponent* of the typical wake-up time  $\tau$  (recall Remark 2.9) and depends on the sequences  $\underline{e}, \underline{K}$  in (2.10) governing the exchange with the seed-bank. If  $g = dg_{\text{FW}}$ , then in view of Theorem 2.12 the criterion in (3.8) determines whether two lineages in the dual coalesce with probability 1.

In Section 6 we will use the renewal process induced by the dual (recall Remark 2.9) to show that (3.8) indeed gives the *total hazard of coalescence* of two dual lineages. Therefore the integral in (3.8) is the counterpart of (3.2). The rate of coalescence again does not affect the dichotomy: in [GHO20] a duality comparison argument was used to show that (3.8) gives the clustering criterion also for  $g \in \mathcal{G}$  with  $g \neq dg_{\text{FW}}$ . The effect of the seed-bank on the dichotomy is embodied by the term  $\hat{\varphi}(t)^{-1/\gamma} t^{-(1-\gamma)/\gamma}$  in (3.8). The criterion in (3.8) 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 (3.8) says that (recall Remark 2.1)

$$(3.9) \quad \text{clustering} \iff \text{either } \delta^- \leq -\frac{1-\gamma}{\gamma} \text{ or } \delta^+ < -\frac{1-\gamma}{\gamma}.$$

Condition (3.8) 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 3.1.** In what follows we will focus on the following two specific parameter regimes:

- *Asymptotically polynomial*, i.e.,

$$(3.10) \quad K_k \sim Ak^{-\alpha}, \quad e_k \sim Bk^{-\beta}, \quad c_k \sim Fk^{-\phi}, \quad k \rightarrow \infty, \quad A, B, F \in (0, \infty), \quad \alpha, \beta, \phi \in \mathbb{R}.$$

- *Pure exponential*, i.e.,

$$(3.11) \quad K_k = K^k, \quad e_k = e^k, \quad c_k = c^k, \quad k \in \mathbb{N}_0, \quad K, e, c \in (0, \infty).$$

Note that both (2.7) and (2.12) are satisfied for  $N \rightarrow \infty$ . Also note that an infinite seed-bank corresponds to  $\alpha \in (-\infty, 1]$ , respectively,  $K \in [1, \infty)$ .  $\square$

The scaling of the wake-up time and the migration kernel in these parameter regimes are as follows.

**Theorem 3.2.** [Infinite seed-bank: Scaling of wake-up time and migration kernel] *Suppose that  $\rho = \infty$ . Then*

(a) *Subject to (3.10),*

$$(3.12) \quad \gamma = 1, \quad \varphi(t) \asymp (\log t)^{-\alpha}, \quad \hat{\varphi}(t) \asymp \begin{cases} (\log t)^{1-\alpha}, & \alpha \in (-\infty, 1), \\ \log \log t, & \alpha = 1, \end{cases}$$

$$a_t^{\Omega_N}(0, 0) \asymp t^{-1} \log^\phi t.$$

(b) *Subject to (3.11),*

$$(3.13) \quad \gamma = \gamma_{N,K,e} = \frac{\log(N/Ke)}{\log(N/e)}, \quad \varphi(t) \asymp 1, \quad \hat{\varphi}(t) \asymp \begin{cases} 1, & K \in (1, \infty), \\ \log t, & K = 1, \end{cases}$$

$$a_t^{\Omega_N}(0, 0) \asymp t^{-1-\delta_{N,c}},$$

where

$$(3.14) \quad \delta_{N,c} = \frac{\log c}{\log(N/c)}.$$

Theorem 3.2 will be proved in Section 6.

Note that, by (3.13),  $\gamma_{N,K,e} = 1$  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  $N \rightarrow \infty$ . Also note that, subject to (3.10), (3.12) says that the degree of the random walk is  $0^-$  for  $\phi \geq -1$  and  $0^+$  for  $\phi < -1$ , while subject to (3.11), by (3.13), the degree of the random walk is  $\delta_{N,c}^-$ , which is 0 for all  $N$  when  $c = 1$ , and tends to 0 as  $N \rightarrow \infty$  from above when  $c > 1$  and from below as  $c < 1$ . Thus, both (3.10) and (3.11) with  $N \rightarrow \infty$  correspond to a *critically recurrent migration* and a *critically infinite seed-bank*.

### 3.3 Clustering regime

Summarising the above discussion, we can now identify the clustering regime for both finite and infinite seed-banks.

**Theorem 3.3.** [Clustering regime]

(1) *If  $\rho < \infty$ , then clustering prevails if and only if*

$$(3.15) \quad \sum_{k \in \mathbb{N}_0} \frac{1}{c_k} = \infty.$$

(2) *If  $\rho = \infty$ , then clustering prevails for  $N$  large enough*

(a) *Subject to (3.10) if and only if*

$$(3.16) \quad -\phi \leq \alpha \leq 1.$$

(b) *Subject to (3.11) if and only if*

$$(3.17) \quad Kc \leq 1 \leq K.$$

Also Theorem 3.3 will be proved in Section 6.

Note that for  $\rho < \infty$  the clustering regime follows by combining (3.2) and (3.3), while for  $\rho = \infty$  the clustering regime follows by substituting into (3.8) the scaling of the wake-up times and the migration kernel stated in Theorem 3.2.

**Remark 3.4.** Note that subject to (3.10), respectively, (3.11),  $\rho < \infty$  implies that  $\alpha > 1$ , respectively,  $K < 1$ , and so the clustering regime is  $-\phi \leq 1$ , respectively,  $c \leq 1$  (recall (3.15)), which are less stringent than (3.16), respectively, (3.17).  $\blacksquare$

## 4 Main results: $N \rightarrow \infty$ , 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 2.3. In Section 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 4.2 we present the conceptual ingredients needed for our theorems. In Section 4.3 we state two versions of the hierarchical multi-scale limit (Theorems 4.2 and 4.4), and comment on how they are related to each other. In Section 4.4 we explain how they arise from a heuristic analysis of the SSDE in (2.20).

### 4.1 Intermezzo: Meyer-Zheng topology

Recall the block averages defined in (2.3) and their evolution equations in (2.24). In the limit as  $N \rightarrow \infty$ , some of the pre-factors in (2.24) 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 [DG14, pp. 792–794] and employ the so-called *Meyer-Zheng topology* on *pseudopaths*, (see [MZ84]), which is based on the following idea. Consider functions  $f: [0, \infty) \rightarrow E$ , 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

$$(4.1) \quad \lim_{n \rightarrow \infty} \int_a^b dt [1 \wedge d(f(t), f_n(t))] = 0 \quad \forall 0 \leq a < b < \infty.$$

However, the topology induced by the metric in (4.1) 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 (4.1) 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  $\psi_v$  is the *probability measure*  $\rho$  on  $[0, \infty) \times E$ , defined by

$$(4.2) \quad \rho((a, b) \times B) = \int_a^b ds e^{-s} 1_B(v(s)), \quad B \in \mathcal{B}(E),$$

where  $\mathcal{B}(E)$  is the Borel  $\sigma$ -algebra over  $E$ . Hence  $\psi_v$  is the image measure of  $e^{-t} dt$  under the mapping  $t \rightarrow (t, v(t))$ . 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 (4.1). 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  $\Psi$ .

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  $\lim_{n \rightarrow \infty} \psi_{v_n} = \psi_v$ , i.e.,  $\lim_{n \rightarrow \infty} \psi_{v_n} f = \psi_v f$  for all continuous bounded functions  $f$  on  $[0, \infty) \times E$ .

**Remark 4.1.** [**Pseudopaths**] The space  $\Psi$  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 [MZ84, p. 372]).  $\blacksquare$

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 we collect some basic facts about the Meyer-Zheng topology taken from [MZ84] and [Kur91].

## 4.2 Main ingredients for the hierarchical multi-scale limit

Recall the definition of  $\theta_x$  and  $\theta_{y_m}$  in (2.21). Define

$$(4.3) \quad \vartheta_k = \frac{\theta_x + \sum_{m=0}^k K_m \theta_{y_m}}{1 + \sum_{m=0}^k K_m}, \quad k \in \mathbb{N}_0.$$

For  $\rho < \infty$ , and for  $\rho = \infty$  under the additional assumption of colour regularity (recall Proposition 2.12), we have

$$(4.4) \quad \lim_{k \rightarrow \infty} \vartheta_k = \theta \quad \text{for some } \theta \in [0, 1].$$

Define the *slowing-down constants* ( $E_0 = 1$ )

$$(4.5) \quad E_k = \frac{1}{1 + \sum_{m=0}^{k-1} K_m}, \quad k \in \mathbb{N}_0.$$

For  $l \in \mathbb{N}_0$ , let

$$(4.6) \quad (\theta, (y_{m,l})_{m \in \mathbb{N}_0})$$

be a sequence of random variables taking values in  $[0, 1]$ , and let

$$(4.7) \quad \left( z_{l,(\theta, (y_{m,l})_{m \in \mathbb{N}_0})}(t) \right)_{t \geq 0} = \left( x_l(t), (y_{m,l}(t))_{m \in \mathbb{N}_0} \right)_{t \geq 0}$$

be the *full process* on level  $l$  evolving according to

$$(4.8) \quad \begin{aligned} dx_l(t) &= E_l \left[ c_l [\theta - x_l(t)] dt + \sqrt{(\mathcal{F}^{(l)}g)(x_l(t))} dw(t) + K_l e_l [y_{l,l}(t) - x_l(t)] dt \right], \\ y_{m,l}(t) &= x_l(t), & 0 \leq m < l, \\ dy_{l,l}(t) &= e_l [x_l(t) - y_{l,l}(t)] dt, & m = l, \\ y_{m,l}(t) &= y_{m,l}, & m > l. \end{aligned}$$

where  $\mathcal{F}^{(l)}g$  is an element of  $\mathcal{G}$ , (recall (2.15)), that will be defined in (4.17) below. By [YW71] the above SSDE has a unique solution for every initial measure. For  $l \in \mathbb{N}_0$ , let

$$(4.9) \quad (z_{i,\theta}^{\text{eff}}(t))_{t \geq 0} = (x_i^{\text{eff}}(t), y_{i,l}^{\text{eff}}(t))_{t \geq 0}$$

be the *effective process* evolving according to

$$(4.10) \quad \begin{aligned} dx_i^{\text{eff}}(t) &= E_l \left[ c_l [\theta - x_i^{\text{eff}}(t)] dt + \sqrt{(\mathcal{F}^{(l)}g)(x_i^{\text{eff}}(t))} dw(t) + K_l e_l [y_{i,l}^{\text{eff}}(t) - x_i^{\text{eff}}(t)] dt \right], \\ dy_{i,l}^{\text{eff}}(t) &= e_l [x_i^{\text{eff}}(t) - y_{i,l}^{\text{eff}}(t)] dt. \end{aligned}$$

Comparing (4.8) with (4.10), we see that the effective process looks at the non-trivial components of the full process.

Apart from (4.7) and (4.9), we need the following list of *ingredients* to formally state the multi-scale limit:



1. For  $l \in \mathbb{N}_0$  and  $t > 0$ , define the *estimators* for the finite system by

$$(4.11) \quad \begin{aligned} \bar{\Theta}^{(l),\Omega_N}(t) &= \frac{1}{N^l} \sum_{\xi \in B_l} \frac{x_\xi^{\Omega_N}(t) + \sum_{m=0}^{l-1} K_m y_{\xi,m}^{\Omega_N}(t)}{1 + \sum_{m=0}^{l-1} K_m}, \\ \bar{\Theta}_x^{(l),\Omega_N}(t) &= \frac{1}{N^l} \sum_{\xi \in B_l} x_\xi^{\Omega_N}(t), \\ \bar{\Theta}_{y_m}^{(l),\Omega_N}(t) &= \frac{1}{N^l} \sum_{\xi \in B_l} y_{\xi,m}^{\Omega_N}(t), \quad m \in \mathbb{N}_0, \end{aligned}$$

and put

$$(4.12) \quad \begin{aligned} \Theta^{(l),\Omega_N}(t) &= (\Theta_x^{(l),\Omega_N}(t), (\Theta_{y_m}^{(l),\Omega_N}(t))_{m \in \mathbb{N}_0}), \\ \Theta^{\text{eff},(l),\Omega_N}(t) &= (\bar{\Theta}^{(l),\Omega_N}(t), \Theta_{y_l}^{(l),\Omega_N}(t)). \end{aligned}$$

We call  $(\Theta^{(l),\Omega_N}(t))_{t \geq 0}$  the *full estimator process* and  $(\Theta^{\text{eff},(l),\Omega_N}(t))_{t \geq 0}$  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 (2.24), 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  $N \rightarrow \infty$ . 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  $N \rightarrow \infty$ .

2. For  $l \in \mathbb{N}_0$ , consider *time scales*  $N^l t_l$  such that

$$(4.13) \quad \mathcal{L}[\bar{\Theta}^{(l),\Omega_N}(N^l t_l - L(N)N^{l-1}) - \bar{\Theta}^{(l),\Omega_N}(N^l t_l)] = \delta_0$$

for all sequences  $(L(N))_{N \in \mathbb{N}}$  satisfying  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , 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.

3. For  $l \in \mathbb{N}_0$  the *invariant measure* for the limiting evolution of the full process on level  $l$  in (4.8) is denoted by

$$(4.14) \quad \Gamma_{(\theta, y_l)}^{(l)}, \quad y_l = (y_{m,l})_{m \in \mathbb{N}_0}.$$

(The existence of and convergence to this equilibrium will be proved in Section 12.2.) Note that  $\Gamma_{(\theta, y_l)}^{(l)}$  depends on choice of the rates  $E_l, c_l, K_l, e_l$  in (4.8). The *invariant measure* of the limiting evolution for the effective  $l$ -block process in (4.10) is denoted by

$$(4.15) \quad \Gamma_\theta^{\text{eff},(l)}.$$

Also  $\Gamma_\theta^{\text{eff},(l)}$  depends on the choice of the rates  $E_l, c_l, K_l, e_l$ .

4. 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

$$(4.16) \quad (\mathcal{F}^{E_l, c_l, K_l, e_l} g)(\theta) = \int_{[0,1]^2} g(x) \Gamma_\theta^{\text{eff},(l)}(dx), \quad \theta \in [0, 1].$$

(In Section 9 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

$$(4.17) \quad \mathcal{F}^{(k)} = \mathcal{F}^{E_{k-1}, c_{k-1}, K_{k-1}, e_{k-1}} \circ \dots \circ \mathcal{F}^{E_0, c_0, K_0, e_0}.$$

5. For  $k \in \mathbb{N}_0$ , define the *interaction chain* [DGV95]

$$(4.18) \quad (M_{-l}^k)_{-l=-(k+1), -k, \dots, 0}$$

as the *time-inhomogeneous* Markov chain on  $[0, 1] \times [0, 1]^{\mathbb{N}_0}$  with initial state

$$(4.19) \quad M_{-(k+1)}^k = (\vartheta_k, \overbrace{\vartheta_k, \dots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y_{k+1}}, \theta_{y_{k+2}}, \dots)$$

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

$$(4.20) \quad Q^{[l]}(u, dv) = \Gamma_u^{(l)}(dv).$$

(See Fig. 9.) For  $k \in \mathbb{N}_0$ , define the *effective interaction chain*

$$(4.21) \quad (M_{-l}^{\text{eff}, k})_{-l=-(k+1), -k, \dots, 0}$$

as the *time-inhomogeneous* Markov chain on  $[0, 1] \times [0, 1]$  with initial state

$$(4.22) \quad M_{-(k+1)}^{\text{eff}, k} = (\vartheta_k, \theta_{y_{k+1}})$$

that evolves from time  $-(l+1)$  to time  $-l$  according to the transition kernel  $Q^{\text{eff}, [l]}$  on  $[0, 1] \times [0, 1]$  given by

$$(4.23) \quad Q^{\text{eff}, [l]}(u, dv) = \Gamma_{u_x}^{\text{eff}, (l)}(dv),$$

where  $u_x$  denotes the first component of  $u = (u_x, u_y)$ . (See Fig. 8.) We denote the components of  $(M_{-l}^{\text{eff}, k})$  by

$$(4.24) \quad M_{-l}^{\text{eff}, k} = (M_{-l, x}^{\text{eff}, k}, M_{-l, y}^{\text{eff}, k}).$$

### 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.

#### 4.3.1 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 4.2.

**Theorem 4.2. [Hierarchical mean-field: the effective process]** *Suppose that the initial state of the hierarchical system is given by (2.21). Let  $L(N)$  be such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . For  $k \in \mathbb{N}$  and  $t_k, \dots, t_0 \in (0, \infty)$ , set  $\bar{t} = N^k L(N) + \sum_{n=0}^k N^n t_n$ .*

(a) For  $k \in \mathbb{N}$ ,

$$(4.25) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (l), \Omega_N}(\bar{t}) \right)_{l=k+1, k, \dots, 0} \right] = \mathcal{L} \left[ (M_{-l}^{\text{eff}, k})_{-l=-(k+1), -k, \dots, 0} \right].$$

(b) For  $k \in \mathbb{N}$ ,

$$(4.26) \quad \begin{aligned} l > k: & \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (l), \Omega_N}(\bar{t} + N^k t) \right)_{t > 0} \right] = \delta_{(\vartheta_l, \theta_{y_l})}, \\ l = k: & \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (l), \Omega_N}(\bar{t} + N^l t) \right)_{t > 0} \right] = \mathcal{L} \left[ \left( z_{k, M_{-(k+1), x}^{\text{eff}, k}}^{\text{eff}}(t) \right)_{t > 0} \right], \\ l < k: & \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (l), \Omega_N}(\bar{t} + N^l t) \right)_{t > 0} \right] = \mathcal{L} \left[ \left( z_{l, M_{-(l+1), x}^{\text{eff}, k}}^{\text{eff}}(t) \right)_{t > 0} \right], \end{aligned}$$

$$\begin{array}{c}
M_{-(k+1)}^{k,\text{eff}} = (\vartheta_k, \theta_{k+1}) \\
\downarrow Q^{[k]} \left( M_{-(k+1)}^{k,\text{eff}}, (du, dv) \right) = \Gamma_{\vartheta_k}^{\text{eff}}(du, dv) \\
M_{-k}^{k,\text{eff}} = (x_k(t_k), y_{k,k}(t_k)) \\
\downarrow Q^{[k-1]} \left( M_{-k}^{k,\text{eff}}, (du, dv) \right) = \Gamma_{x_k(t_k)}^{\text{eff}}(du, dv) \\
M_{-(k-1)}^{k,\text{eff}} = (x_{k-1}(t_{k-1}), y_{k-1,k-1}(t_{k-1})) \\
\downarrow Q^{[k-2]} \left( M_{-(k-1)}^{k,\text{eff}}, (du, dv) \right) = \Gamma_{x_{k-1}(t_{k-1})}^{\text{eff}}(du, dv) \\
\vdots \\
\downarrow \\
M_{-1}^{k,\text{eff}} = (x_1(t_1), y_{1,1}(t_1)) \\
\downarrow Q^{[0]} \left( M_{-1}^{k,\text{eff}}, (du, dv) \right) = \Gamma_{x_1(t_1)}^{\text{eff}}(du, dv) \\
M_0^{k,\text{eff}} = (x_0(t_0), y_{0,0}(t_0))
\end{array}$$

Figure 8: Effective interaction chain.

where the initial laws of the limiting processes are given by (see Fig. 8)

$$\begin{aligned}
(4.27) \quad & \mathcal{L} \left[ z_{k, M_{-(k+1),x}^k}^{\text{eff}}(0) \right] = \Gamma_{M_{-(k+1),x}^k}^{\text{eff},(k)}, \\
& \mathcal{L} \left[ z_{k, M_{-(k+1),x}^k}^{\text{eff}}(0) \right] = \Gamma_{M_{-(l+1),x}^k}^{\text{eff},(k)}, \\
& \Gamma_{M_{-(l+1)}^k}^{\text{eff},(l)} = \int_{[0,1]^2} \dots \int_{[0,1]^2} \Gamma_{M_{-(k+1)}^k}^{\text{eff},(k)}(du_k) \dots \Gamma_{u_{l+2}}^{\text{eff},(l+1)}(du_{l+1}) \Gamma_{u_{l+1}}^{\text{eff},(l)}
\end{aligned}$$

Theorem 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 (4.21) and depicted in Fig. 8. Note that the process  $(\Theta^{\text{eff},(l),\Omega_N}(\bar{t} + N^l t))_{t>0}$  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.,  $(\Theta^{\text{eff},(l),\Omega_N}(\bar{t} + N^k t))_{t>0}$  converges to the constant process taking the value  $(\vartheta_l, \theta_{y_l}) = \Theta^{\text{eff},(l),\Omega_N}(0)$ .
- 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  $\vartheta_k$ . Therefore migration between the  $k$ -blocks in the hierarchical mean-field limit is replaced by a drift towards  $\vartheta_k$ , 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 (4.16)), 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 [DG93b]. However, the seed-bank still quantitatively affects the migration and the resampling through the slowing-down factor  $E_k$ . (In Section 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 4.2 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 below. We will see later that Theorem 4.2 does describe all the *non-trivial* components of the system.

**Remark 4.3. [Quasi equilibria]** Note that Theorem 4.2 does not depend on the choice of  $t_k, \dots, t_0 \in (0, \infty)$ . 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.  $\blacksquare$

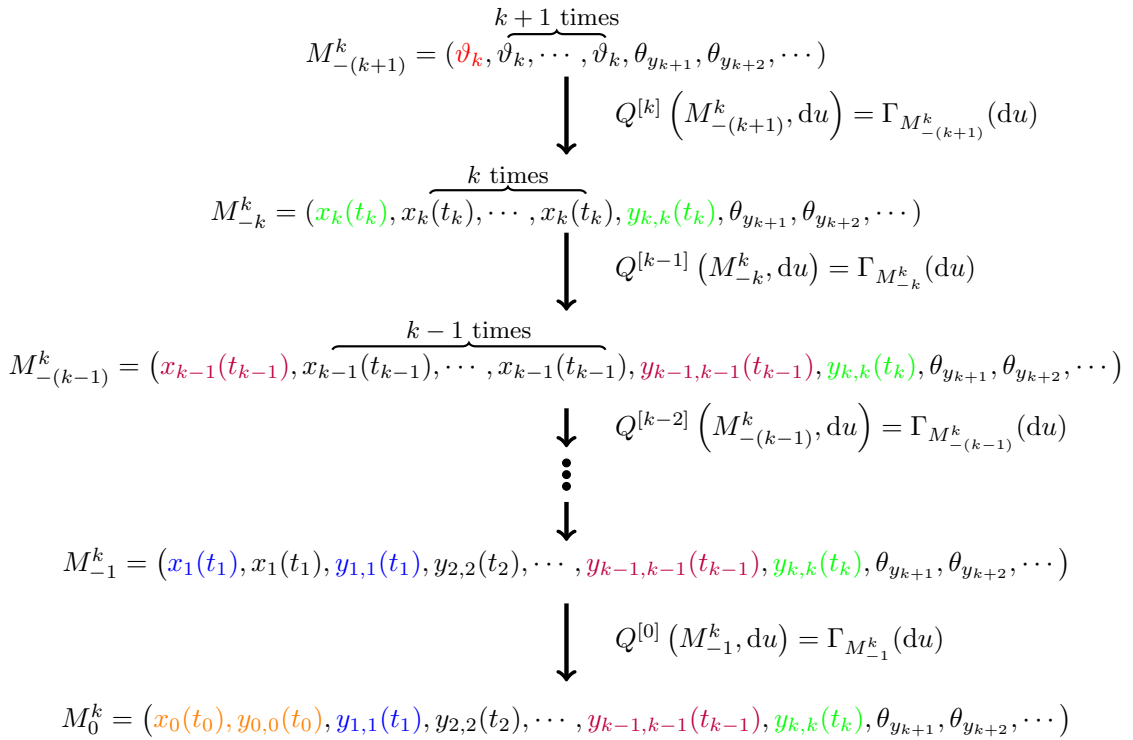


Figure 9: Full interaction chain.

### 4.3.2 Full process

To state our second main theorem, we will again use the process and the notation as defined in Section 4.2.

**Theorem 4.4. [Hierarchical mean-field: full process]** *Suppose that the initial state is given by (2.21). Let  $L(N)$  be such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . For  $k \in \mathbb{N}$  and  $t_k, \dots, t_0 \in (0, \infty)$ , set  $\bar{t} = N^k L(N) + \sum_{n=0}^k N^n t_n$ .*

(a) For  $k \in \mathbb{N}$ ,

$$(4.28) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(l), \Omega_N}(\bar{t}) \right)_{l=k+1, k, \dots, 0} \right] = \mathcal{L} \left[ (M_{-l}^k)_{-l=-(k+1), -k, \dots, 0} \right].$$

(b) For  $k \in \mathbb{N}$ ,

$$(4.29) \quad \begin{aligned} l > k: \quad & \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(l), \Omega_N}(\bar{t} + N^k t) \right)_{t > 0} \right] = \delta_{(M_{-(k+1)}^k)}, \\ l = k: \quad & \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(l), \Omega_N}(\bar{t} + N^l t) \right)_{t > 0} \right] = \mathcal{L} \left[ \left( z_{k, M_{-(k+1)}^k}(t) \right)_{t > 0} \right], \\ l < k: \quad & \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(l), \Omega_N}(\bar{t} + N^l t) \right)_{t > 0} \right] = \mathcal{L} \left[ \left( z_{l, M_{-(l+1)}^k}(t) \right)_{t > 0} \right], \end{aligned}$$

in the Meyer-Zheng topology,

where the initial laws of the limiting processes are given by (see Fig. 9)

$$(4.30) \quad \begin{aligned} \mathcal{L} \left[ z_{k, M_{-(k+1), x}^k}(0) \right] &= \Gamma_{M_{-(k+1)}^k}^{(k)}, \\ \mathcal{L} \left[ z_{k, M_{-(k+1), x}^k}(0) \right] &= \Gamma_{M_{-(l+1)}^k}^{(l)}, \\ \Gamma_{M_{-(l+1)}^k}^{(l)} &= \int_{\mathfrak{s}} \cdots \int_{\mathfrak{s}} \int_{\mathfrak{s}} \Gamma_{M_{-(k+1)}^k}^{(k)}(du_k) \Gamma_{u_k}^{(k-1)}(du_{k-1}) \cdots \Gamma_{l+2}^{(l+1)}(du_1) \Gamma_{u_{l+1}}^{(l)} \end{aligned}$$

**Remark 4.5. [convergence in the Meyer-Zheng topology]** Note that Theorem 4.4(b) is stated in the Meyer-Zheng topology. This topology is needed because at time-scales  $N^l t$  rates occur in (2.24) that tend to infinity as  $N \rightarrow \infty$ . In Section 4.4 we define the Meyer-Zheng topology and explain why it is needed.  $\blacksquare$

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 (4.18). 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.,  $(\Theta^{(l), \Omega_N}(\bar{t} + N^k t))_{t > 0}$  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  $\vartheta_k$  instead of their initial state  $\theta_{y_m}$ . However, on time scale  $\bar{t} + N^k t$   $l$ -block averages of  $m$ -dormant populations are not moving. (In Section 4.4 we explain how the shift from  $\theta_{y_m}$  to  $\vartheta_k$  occurs.) Also the  $l$ -block averages of the active population are in state  $\vartheta_k$ .
- 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 4.2. 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  $\vartheta_k$  (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 4.2. 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  $M_{-(l+1)}^k$ , 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  $M_{-(l+1)}^k$ . Hence the  $l$ -block averages with colours  $m > k$  are still in their initial states  $\theta_{y_m}$ , 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 were 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.)

**Remark 4.6. [Comparison to the hierarchical multi-scale limit without seed-bank]** Comparing Theorem 4.4 with the multi-scale limit theorems derived for the hierarchical system without seed-bank [DG93b], [DG93a], [DGV95], 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 4.7. [ $k \rightarrow \infty$  limit of the interaction chain]** The result in (4.26) raises the question how the hierarchical multi-scale limit behaves for large  $k$ . We find the following dichotomy:

$$(4.31) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ (M_{-(k+1), -k, \dots, 0}^k) \right] = \mathcal{L} \left[ (M_k^\infty)_{k \in \mathbb{Z}^-} \right],$$

where in the *clustering regime*

$$(4.32) \quad \mathcal{L} \left[ (M_k^\infty)_{k \in \mathbb{Z}^-} \right] = \theta \delta_{(1, 1^{\mathbb{N}_0})^{\mathbb{Z}^-}} + (1 - \theta) \delta_{(0, 0^{\mathbb{N}_0})^{\mathbb{Z}^-}}$$

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

$$(4.33) \quad \lim_{l \rightarrow \infty} M_{-l}^\infty = (\theta, \theta).$$

In the latter case,  $M^\infty$  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  $N \rightarrow \infty$  (see [DGV95, Proposition 6.2 and 6.3]). ■

**Remark 4.8. [Interaction field]** Theorem 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  $N \rightarrow \infty$ , 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  $\eta, \eta'$  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 [DGV95, 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 Heuristics behind the multi-scale limit

The proofs of Theorems 4.2 and 4.4 written out in Sections 7–12, are long and technical. In order to help the reader appreciate these proofs, we provide the heuristics in Sections 4.4.1–4.4.3.

#### 4.4.1 Evolution of the block-averages

Recall the block averages introduced in Definition 2.3 and their evolution defined in (2.24). In the limit as  $N \rightarrow \infty$ , we heuristically obtain from the SSDE in (2.24) the following results for the  $k$ -block process

$$(4.34) \quad (x_k^{\Omega_N}(t), (y_{m,k}^{\Omega_N}(t))_{m \in \mathbb{N}_0})_{t \geq 0}.$$

- **Migration.** Recall that the migration is captured by the first term of the first equation in (2.24), i.e., the first term of the evolution of the active part of the population. Letting  $N \rightarrow \infty$ , 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  $x_{k+1}^{\Omega_N}(0)$ , the initial state of the active  $(k+1)$ -block average. Note that the migration term in (2.24) can be written as

$$(4.35) \quad \sum_{l \in \mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} [x_{k+l}^{\Omega_N}(N^{-l}t) - x_k^{\Omega_N}(t)] dt = \sum_{l \in \mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} \left[ \frac{1}{N^l} \sum_{k=0}^{N^l-1} x_k^{\Omega_N}(t) - x_k^{\Omega_N}(t) \right] dt.$$

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  $N \rightarrow \infty$ , the latter can be approximated by  $\mathbb{E}[x_k(t)]$ . Effectively, as  $N \rightarrow \infty$ , the  $k$ -blocks become independent given the value of  $x_{k+1}^{\Omega_N}(0)$ , 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 rate of the single colonies. Therefore, in the limit  $N \rightarrow \infty$ , 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 exactly the renormalised diffusion function  $\mathcal{F}g$  (see (4.16)). 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  $N \rightarrow \infty$ , 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-equilibrium of the  $(k-1)$ -blocks. This yields another iteration of the renormalisation transformation (see (4.17)). 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 (2.24) and the second equation in (2.24) together describe the exchange of the active  $k$ -block with the  $m$ -dormant  $k$ -block. To describe the limiting behaviour as  $N \rightarrow \infty$ , we distinguish three cases:  $0 \leq m < k$ ,  $m = k$ ,  $m > k$ .

- If  $0 \leq m < k$ , then we see that the rate of exchange between the active  $k$ -block and the  $m$ -dormant  $k$ -block tends to infinity as  $N \rightarrow \infty$ . We therefore expect them to equalise, i.e.,

$$(4.36) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( x_k^{\Omega_N}(t) - y_{m,k}^{\Omega_N}(t) \right)_{t > 0} \right] = \delta_0,$$

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* [MZ84].)

- 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  $N \rightarrow \infty$ .

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*.

#### 4.4.2 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  $N \rightarrow \infty$  of (2.24), 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 (2.24), we look at the *effective  $k$ -block process* defined as

$$(4.37) \quad \left( \bar{x}_k^{\Omega_N}(t), y_{k,k}^{\Omega_N}(t) \right)_{t \geq 0},$$

where we abbreviate

$$(4.38) \quad \bar{x}_k^{\Omega_N}(t) = \frac{x_k^{\Omega_N}(t) + \sum_{m=0}^{k-1} K_m y_{m,k}^{\Omega_N}(t)}{1 + \sum_{m=0}^{k-1} K_m}.$$

By (4.36) and the heuristic discussion given above, the process in (4.37) equals  $(x_k^{\Omega_N}(t), y_{k,k}^{\Omega_N}(t))_{t \geq 0}$  in the limit as  $N \rightarrow \infty$ , 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 (2.24), we see that the process in (4.37) evolves according to the SSDE

$$(4.39) \quad \begin{aligned} d\bar{x}_k^{\Omega_N}(t) &= E_k \sum_{l \in \mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} [x_{k+l}^{\Omega_N}(N^{-l}t) - x_k^{\Omega_N}(t)] dt \\ &\quad + E_k \sqrt{\frac{1}{N^k} \sum_{\xi \in B_k(0)} g(x_\xi(N^k t))} dw_k(t) \\ &\quad + E_k \sum_{m=k}^{\infty} N^k \frac{K_m e_m}{N^m} [y_{m,k}^{\Omega_N}(t) - x_k^{\Omega_N}(t)] dt, \\ dy_{k,k}^{\Omega_N}(t) &= e_k [x_k^{\Omega_N}(t) - y_{k,k}^{\Omega_N}(t)] dt. \end{aligned}$$

In (4.39) no infinite rates appear anymore. In the limit as  $N \rightarrow \infty$ , by (4.36) we can approximate

$$(4.40) \quad x_k^{\Omega_N}(t) \approx y_{m,k}^{\Omega_N}(t), \quad 0 \leq m < k,$$

such that

$$(4.41) \quad x_k^{\Omega_N}(t) \approx \bar{x}_k^{\Omega_N}(t).$$

We can therefore approximate (4.39) by

$$(4.42) \quad \begin{aligned} d\bar{x}_k^{\Omega_N}(t) &= E_k \sum_{l \in \mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} [\bar{x}_{k+l}^{\Omega_N}(N^{-l}t) - \bar{x}_k^{\Omega_N}(t)] dt \\ &\quad + E_k \sqrt{\frac{1}{N^k} \sum_{\xi \in B_k(0)} g(\bar{x}_\xi(N^k t))} dw_k(t) \\ &\quad + E_k \sum_{m=k}^{\infty} N^k \frac{K_m e_m}{N^m} [y_{m,k}^{\Omega_N}(t) - \bar{x}_k^{\Omega_N}(t)] dt, \\ dy_{k,k}^{\Omega_N}(t) &= e_k [\bar{x}_k^{\Omega_N}(t) - y_{k,k}^{\Omega_N}(t)] dt. \end{aligned}$$

Hence, in the limit as  $N \rightarrow \infty$ , the process in (4.37) becomes autonomous. Moreover, assuming that  $\lim_{N \rightarrow \infty} \bar{x}_{l+1}^{\Omega_N}(0) = \vartheta_l$ , we see that (4.42) approaches the effective process defined in (4.10), with

$$(4.43) \quad \theta = \vartheta_l, \quad E = E_k, \quad c = c_k, \quad e = e_k, \quad K = K_k, \quad g = \mathcal{F}^{(k)}g.$$



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 (4.37). 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 7–12 to build on techniques developed for the hierarchical mean-field model without seed-bank in [DG93a], [DGV95]. 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 (4.40) and (4.41), 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

$$\begin{aligned}
(4.44) \quad d\bar{x}_k^{\Omega_N}(t) &\approx E_k \sum_{l \in \mathbb{N}} \frac{c_{k+l-1}}{N^{l-1}} [\bar{x}_{k+l}^{\Omega_N}(N^{-l}t) - \bar{x}_k^{\Omega_N}(t)] dt \\
&\quad + E_k \sqrt{\frac{1}{N^k} \sum_{\xi \in B_k(0)} g(\bar{x}_\xi(N^k t))} dw_k(t) \\
&\quad + E_k \sum_{m=k}^{\infty} N^k \frac{K_m e_m}{N^m} [y_{m,k}^{\Omega_N}(t) - \bar{x}_k^{\Omega_N}(t)] dt, \\
y_{k,k}^{\Omega_N}(t) &= \bar{x}_k^{\Omega_N}(t), \\
dy_{k,k}^{\Omega_N}(t) &= e_k [\bar{x}_k^{\Omega_N}(t) - y_{k,k}^{\Omega_N}(t)] dt, \\
y_{k,k}^{\Omega_N}(t) &= y_{m,k}^{\Omega_N}(0).
\end{aligned}$$

By the ergodic theorem for exchangeable measures, we can assume that  $\lim_{N \rightarrow \infty} \bar{x}_{k+1}^{\Omega_N}(0) = \vartheta_k$  a.s.. We expect that (4.44) approaches (4.8) with

$$(4.45) \quad \theta = \vartheta_k, \quad E = E_k, \quad c = c_k, \quad e = e_k, \quad K = K_k, \quad g = \mathcal{F}^{(k)}g.$$

To prove that

$$(4.46) \quad \lim_{N \rightarrow \infty} \mathcal{L}[y_{m,k}^{\Omega_N}(t)] = \lim_{N \rightarrow \infty} \mathcal{L}[\bar{x}_k^{\Omega_N}(t)], \quad 0 \leq m < k-1,$$

we need the *Meyer-Zheng topology* explained in Section 4.1. In the proof in Sections 8–12 we show how the above approximations can be made rigorous.

**Conserved quantities.** Note that, by (2.24) and (4.35), for each  $k \in \mathbb{N}_0$

$$(4.47) \quad \mathbb{E} \left[ \frac{x_k^{\Omega_N}(t) + \sum_{m \in \mathbb{N}_0} K_m y_{m,k}^{\Omega_N}(t)}{1 + \sum_{m \in \mathbb{N}_0} K_m} \right] = \theta, \quad t \geq 0,$$

is a conserved quantity. For each  $k \in \mathbb{N}$  we obtain that for  $l \geq k$

$$(4.48) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{x_k^{\Omega_N}(t) + \sum_{m=0}^l K_m y_{m,k}^{\Omega_N}(t)}{1 + \sum_{m=0}^l K_m} \right] = \vartheta_l,$$

is a conserved quantity. For the effective process, (4.42) implies that

$$\begin{aligned}
(4.49) \quad \frac{d}{dt} \mathbb{E}[\bar{x}_k^{\Omega_N}(t)] &= E_k K_k e_k \left( \mathbb{E}[y_{k,k}^{\Omega_N}(t)] - \mathbb{E}[\bar{x}_k^{\Omega_N}(t)] \right), \\
\frac{d}{dt} \mathbb{E}[y_{k,k}^{\Omega_N}(t)] &= e_k \left( \mathbb{E}[\bar{x}_k^{\Omega_N}(t)] - \mathbb{E}[y_{k,k}^{\Omega_N}(t)] \right).
\end{aligned}$$

Recall that

$$(4.50) \quad \mathbb{E}[\bar{x}_k^{\Omega_N}(0)] = \mathbb{E} \left[ \frac{x_k^{\Omega_N}(0) + \sum_{m=0}^{k-1} K_m y_{m,k}^{\Omega_N}(0)}{1 + \sum_{m=0}^{k-1} K_m} \right] = \vartheta_{k-1}, \quad \mathbb{E}[y_{k,k}^{\Omega_N}(0)] = \theta_{y_k}.$$

Therefore we can solve (4.49) explicitly as

$$(4.51) \quad \begin{aligned} \mathbb{E}[\bar{x}_k^{\Omega_N}(t)] &= \vartheta_k + \frac{E_k K_k}{1 + E_k K_k} (\vartheta_{k-1} - \theta_{y_k}) e^{-(E_k K_k + 1)e_k t}, \\ \mathbb{E}[y_{k,k}^{\Omega_N}(t)] &= \vartheta_k - \frac{1}{1 + E_k K_k} (\vartheta_{k-1} - \theta_{y_k}) e^{-(E_k K_k + 1)e_k t}. \end{aligned}$$

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  $\vartheta_{k-1}$  is pulled a bit towards  $\theta_{y_k}$ , 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  $(\theta_x, \vartheta_0, \vartheta_1, \vartheta_2, \dots)$  for the expectation of the active population on space-time scales  $\{0, 1, 2, 3, \dots\}$ . From (4.35) we see that the drift of  $\bar{x}_k^{\Omega_N}(t)$  is towards

$$(4.52) \quad \bar{x}_{k+1}^{\Omega_N}(N^{-1}t) = \frac{1}{N^k} \sum_{k=0}^{N^k-1} x_k^{\Omega_N}(t) \approx \mathbb{E}[\bar{x}_k^{\Omega_N}(t)],$$

where the last approximation can be made because the  $k$ -blocks decouple. Hence, in the limit as  $N \rightarrow \infty$ , 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  $(\Theta^{(l), \Omega_N}(t))_{t>0}$  defined in (4.12). Equation 2.23 implies that the evolution of the estimator process is given by

$$(4.53) \quad \begin{aligned} d\Theta_x^{(l), \Omega_N}(t) &= \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1}} [\Theta_x^{(n), \Omega_N}(t) - \Theta_x^{(l)}(t)] dt \\ &\quad + \sqrt{\frac{1}{N^{2l}} \sum_{\xi \in B_l} g(x_\xi(t))} dw(t) \\ &\quad + \sum_{m \in \mathbb{N}_0} \frac{K_m e_m}{N^m} [\Theta_{y_m}^{(l), \Omega_N}(t) - \Theta_x^{(1), \Omega_N}(t)] dt, \end{aligned}$$

$$d\Theta_{y_m}^{(l), \Omega_N}(t) = \frac{e_m}{N^m} [\Theta_x^{(l), \Omega_N}(t) - \Theta_{y_m}^{(l), \Omega_N}(t)] dt, \quad m \in \mathbb{N}_0.$$

Looking at the estimator process  $(\Theta^{(l), \Omega_N}(t))_{t>0}$  on time scale  $N^k t$ , we see that

$$(4.54) \quad \begin{aligned} d\Theta_x^{(l), \Omega_N}(N^k t) &= \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1-k}} [\Theta_x^{(n), \Omega_N}(N^k t) - \Theta_x^{(l), \Omega_N}(N^k t)] dt \\ &\quad + \sqrt{\frac{N^k}{N^{2l}} \sum_{\xi \in B_l} g(x_\xi(N^k t))} dw(t) \\ &\quad + \sum_{m \in \mathbb{N}_0} \frac{K_m e_m}{N^{m-k}} [\Theta_{y_m}^{(l), \Omega_N}(N^k t) - \Theta_x^{(l), \Omega_N}(N^k t)] dt, \end{aligned}$$

$$d\Theta_{y_m}^{(l), \Omega_N}(N^k t) = \frac{e_m}{N^{m-k}} [\Theta_x^{(l), \Omega_N}(N^k t) - \Theta_{y_m}^{(l), \Omega_N}(N^k t)] dt, \quad m \in \mathbb{N}_0.$$

From (4.54) we get that, on time scale  $N^k t$ , for all  $l \geq k + 1$ ,

$$(4.55) \quad \Theta^{(l), \Omega_N}(N^k t) = \frac{\Theta_x^{(l), \Omega_N}(N^k t) + \sum_{m=0}^{l-1} K_m \Theta_{y_m}^{(l), \Omega_N}(N^k t)}{1 + \sum_{m=0}^{l-1} K_m}, \quad t \geq 0,$$

is a conserved quantity in the limit as  $N \rightarrow \infty$ , and for  $t \geq 0$ ,

$$(4.56) \quad \lim_{N \rightarrow \infty} \Theta^{(l), \Omega_N}(N^k t) = \frac{\theta_x + \sum_{m=0}^{l-1} K_m \theta_{y_m}}{1 + \sum_{m=0}^{l-1} K_m} = \vartheta_l, \text{ in probability.}$$

For  $m > k$ , also  $\Theta_{y_m}^{(l), \Omega_N}(N^k t)$  is a conserved quantity, and

$$(4.57) \quad \lim_{N \rightarrow \infty} \Theta_{y_m}^{(l), \Omega_N}(N^k t) = \theta_{y_m}, \quad t \geq 0.$$

However, for  $l \geq k + 1$ ,  $\Theta_x^{(l), \Omega_N}(N^k t)$  and  $(\Theta_{y_m}^{(l), \Omega_N}(N^k t))_{m=0}^k$  are not conserved quantities in the limit as  $N \rightarrow \infty$ . Note that from 2.24 we heuristically see that the full  $l$ -block estimator process  $(\Theta^{(l), \Omega_N}(N^k t))_{t \geq 0}$  with  $l > k$  converges to the process

$$(4.58) \quad \left( \Theta_x^{(l)}(t), \left( \Theta_{y_m}^{(l)}(t) \right)_{m \in \mathbb{N}_0} \right)_{t > 0},$$

which evolves according to

$$(4.59) \quad \begin{aligned} d\Theta_x^{(l)}(t) &= E_{k-1} K_k e_k [\Theta_{y_k}^{(l)}(t) - \Theta_x^{(l), \Omega_N}(t)] dt, \\ \Theta_{y_m}^{(l)}(t) &= \Theta_x^{(l)}(t), & m < k, \\ d\Theta_{y_k}^{(l)}(t) &= e_k [\Theta_x^{(l)}(t) - \Theta_{y_k}^{(l)}(t)] dt, & m = k, \\ \Theta_{y_m}^{(l)}(t) &= \theta_{y_m}, & m > k. \end{aligned}$$

This system can be explicitly solved as

$$(4.60) \quad \begin{aligned} \Theta_x^{(l)}(t) &= \frac{\Theta_x^{(l)}(0) + E_{k-1} K_k \Theta_{y_k}^{(l)}(0)}{1 + E_{k-1} K_k} \\ &\quad + \frac{E_{k-1} K_k}{1 + E_{k-1} K_k} [\Theta_x^{(l), \Omega_N}(0) - \Theta_{y_k}^{(l)}(0)] e^{-(E_{k-1} K_k e_k + e_k)t}, \\ \Theta_{y_m}^{(l)}(t) &= \Theta_x^{(l)}(t), & m < k, \\ \Theta_{y_k}^{(l)}(t) &= \frac{\Theta_x^{(l)}(0) + E_{k-1} K_k \Theta_{y_k}^{(l)}(0)}{1 + E_{k-1} K_k} \\ &\quad - \frac{1}{1 + E_{k-1} K_k} [\Theta_x^{(l), \Omega_N}(0) - \Theta_{y_k}^{(l)}(0)] e^{-(E_{k-1} K_k e_k + e_k)t}, & m = k, \\ \Theta_{y_m}^{(l)}(t) &= \theta_{y_m}, & m > k. \end{aligned}$$

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  $\Theta_x^{(l)}(0)$  depends on the scaling we choose. To illustrate this, we note that

$$(4.61) \quad \begin{aligned} \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}(0) &= \theta_x, \\ \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}(L(N) + t) &= \vartheta_0, \\ \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}((L(N)N^n + N^n t)) &= \vartheta_n, \quad 0 \leq n \leq k. \end{aligned}$$

However, the scaling  $N^k t$  would imply that

$$(4.62) \quad \begin{aligned} \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}(0) &= \theta_x, \\ \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N} \left( N^k \frac{L(N) + t}{N^k} \right) &= \vartheta_0, \\ \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N} \left( N^k \frac{L(N)N^n + N^n t}{N^k} \right) &= \vartheta_n, \quad 0 \leq n \leq k. \end{aligned}$$

Hence, if  $(t(N))_{N \in \mathbb{N}}$  is a sequence such that  $\lim_{N \rightarrow \infty} N^k t(N) = 0$ , then the value of the limit

$$(4.63) \quad \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}(N^k t(N))$$

depends on  $(t(N))_{N \in \mathbb{N}}$ . Moreover, to obtain a limiting process for  $(\Theta^{(l), \Omega_N}(N^k t))_{t \geq 0}$  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 4.2.

From (4.60) it follows that, for  $l \geq k$ ,

$$(4.64) \quad \begin{aligned} \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}(\bar{t}) &= \vartheta_k, \text{ in probability,} \\ \lim_{N \rightarrow \infty} \Theta_{y_m}^{(l), \Omega_N}(\bar{t}) &= \vartheta_k, \quad m \leq k, \text{ in probability,} \\ \lim_{N \rightarrow \infty} \Theta_{y_m}^{(l), \Omega_N}(\bar{t}) &= \theta_{y_m}, \quad m > k \text{ in probability.} \end{aligned}$$

Note that the shifting of averages mentioned earlier is closely related to the conserved quantities discussed in Section 4.4.2 because, for large  $N$ ,

$$(4.65) \quad \lim_{N \rightarrow \infty} \Theta_x^{(l), \Omega_N}(\bar{t}) \approx \mathbb{E}[x_k^{\Omega_N}],$$

where the expectation is taken in the quasi-equilibrium the  $k$ -blocks have attained after scaling with time  $\bar{t}$ .

#### 4.4.3 Formation of the interaction chain

In Section 4.4.1 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 4.2. From (4.53) it follows that the process  $(\Theta^{(l), \Omega_N}(\bar{t} + N^l t))_{t > 0}$  evolves according to

$$(4.66) \quad \begin{aligned} d\Theta_x^{(l), \Omega_N}(\bar{t} + N^l t) &= \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1-l}} [\Theta_x^{(n), \Omega_N}(\bar{t} + N^l t) - \Theta_x^{(l)}(\bar{t} + N^l t)] dt \\ &\quad + \sqrt{\frac{1}{N^l} \sum_{\xi \in B_l} g(x_\xi(\bar{t} + N^l t))} dw(t) \\ &\quad + \sum_{m \in \mathbb{N}_0} \frac{K_m e_m}{N^{m-l}} [\Theta_{y_m}^{(l), \Omega_N}(\bar{t} + N^l t) - \Theta_x^{(1), \Omega_N}(\bar{t} + N^l t)] dt, \\ d\Theta_{y_m}^{(l), \Omega_N}(\bar{t} + N^l t) &= \frac{e_m}{N^{m-l}} [\Theta_x^{(l), \Omega_N}(\bar{t} + N^l t) - \Theta_{y_m}^{(l), \Omega_N}(\bar{t} + N^l t)] dt, \quad m \in \mathbb{N}_0. \end{aligned}$$

Therefore, in the limit as  $N \rightarrow \infty$ , 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

$$(4.67) \quad \lim_{N \rightarrow \infty} \mathcal{L}[\Theta^{(k+1), \Omega_N}(\bar{t} + N^k t)] = \lim_{N \rightarrow \infty} \mathcal{L}[\Theta^{(k+1), \Omega_N}(\bar{t})],$$

since the  $(k+1)$ -block has not yet started to move at time  $\bar{t} + N^k t$ . From (4.64) it follows that

$$(4.68) \quad \lim_{N \rightarrow \infty} \Theta_x^{(k+1), \Omega_N}(\bar{t} + N^k t) = \vartheta_k \text{ in probability.}$$

Therefore the drift of the active population  $\Theta_x^{(l), \Omega_N}(\bar{t} + N^l t)$  is towards  $\vartheta_k$ . Since  $\bar{t} > L(N)N^k$ , the process  $\Theta^{(k), \Omega_N}(\bar{t} + N^k t)$  has, in the limit  $N \rightarrow \infty$ , already reached its equilibrium, which is denoted by  $\Gamma_{\bar{\vartheta}_k}$ , where

$$(4.69) \quad \bar{\vartheta}_k = (\vartheta_k, \overbrace{\vartheta_k, \dots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y_{k+1}}, \theta_{y_{k+2}}, \dots),$$

so that we recognise  $(\bar{\vartheta}_k) = M_{-(k+1)}^k$ . From (4.64) with  $l = k+1$  we see that  $(\bar{\vartheta}_k) = M_{-(k+1)}^k$  represents the state of  $\Theta^{(k+1), \Omega_N}(\bar{t})$ .

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  $\Gamma_{\vartheta_k}^{(k)}$ . Thus, the drift of the active  $(k-1)$ -block average is towards the instantaneous state of the active  $k$ -block average, which has distribution  $\Gamma_{\vartheta_k}^{(k)}$ . 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 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  $(\Theta^{\text{eff},(l),\Omega_N}(t))_{t>0}$  instead of the full process  $(\Theta^{(l),\Omega_N}(t))_{t>0}$ , we can consider in (4.53) 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  $\Gamma_{\vartheta_k}^{(k)}$  by  $\Gamma_{\vartheta_k}^{\text{eff},(k)}$  and  $\Gamma_u^{(l)}$  by  $\Gamma_u^{\text{eff},(l)}$ . Hence we find the effective interaction chain defined in (4.21) and depicted in Figure 8.

**Slow seed-banks.** From (4.64) 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  $\vartheta_k$ , and all  $l$ -blocks of seed-banks with colour  $m > k$  equal their initial values  $\theta_{y_m}$ . 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. 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, (4.66) 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  $\vartheta_k$ , 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

$$(4.70) \quad y_{m,0}(\bar{t}) = \int_0^{\bar{t}} ds \frac{e_m}{N^m} [x_0(s) - y_{m,0}(s)].$$

Hence, for large  $N$ ,

$$(4.71) \quad y_{m,0}(\bar{t}) \approx \int_0^{L(N)N^k + N^k t_k + \dots + N^m t_m} ds \frac{e_m}{N^m} [x_0(s) - y_{m,0}(s)].$$

Similarly, for the  $l$ -block average with colour  $m$  we have

$$(4.72) \quad \begin{aligned} \Theta_{y_m}^{(l),\Omega_N}(\bar{t}) &= \int_0^{\bar{t}} ds \frac{e_m}{N^m} [\Theta_x^{(l),\Omega_N}(s) - \Theta_{y_m}^{(l),\Omega_N}(s)] \\ &\approx \int_0^{L(N)N^k + N^k t_k + \dots + N^m t_m} ds \frac{e_m}{N^m} [\Theta_x^{(l),\Omega_N}(s) - \Theta_{y_m}^{(l),\Omega_N}(s)]. \end{aligned}$$

Thus, we see that the state of  $\Theta_{y_m}^{(l),\Omega_N}(\bar{t})$  is completely determined at time  $L(N)N^k + N^k t_k + \dots + N^m t_m$ , 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  $L(N)N^k + N^k t_k + \dots + N^m t_m$ , 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  $l < m$ ,

$$(4.73) \quad \begin{aligned} \Theta_{y_m}^{(l), \Omega_N}(\bar{t} + N^l t) &\approx \Theta_{y_m}^{(l), \Omega_N}(L(N)N^k + N^k t_k + \cdots + N^m t_m) \\ &\approx \Theta_{y_m}^{(m), \Omega_N}(L(N)N^k + N^k t_k + \cdots + N^m t_m). \end{aligned}$$

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 4.4.1–4.4.3, we see how Theorems 4.2 and Theorems 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 7–9 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 10–12 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 (2.19) and each capturing a key feature:

1. Two-colour mean-field finite-systems scheme (Section 10.1).
2. Two-level hierarchical mean-field system (Section 11.1).
3. Finite-level mean-field system (Section 12.1).

In Section 12 we put the pieces together to prove the multi-scaling for the infinite-level system given in Theorems 4.2 and 4.4.

## 5 Main results $N \rightarrow \infty$ : 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 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  $k \rightarrow \infty$  to the Fisher-Wright diffusion function, irrespective of the choice of  $g$ . In Section 5.2 we show how the scaling factors  $A_k$  are connected to the *growth of mono-type clusters*. In Section 5.3 we identify the asymptotics of  $A_k$  as  $k \rightarrow \infty$  in terms of the model parameters.

### 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 (4.20) that describe the connection between subsequent hierarchical levels in the interaction chain. Define the following composition of kernels (see Fig. 10):

$$(5.1) \quad Q^{(n)} = Q^{[n]} \circ \cdots \circ Q^{[0]}, \quad n \in \mathbb{N}.$$

In words,  $Q^{(n)}(z_n, dz_0)$  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 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 (4.17)) scales with  $n$ . We identify the scaling and show that it does *not* depend on  $g$  (see Fig. 11).

To state the clustering result, abbreviate

$$(5.2) \quad \bar{\vartheta}^{(n)} = (\vartheta_n, \overbrace{\vartheta_n, \dots, \vartheta_n}^{n+1 \text{ times}}, \vartheta_{y_{n+1}}, \vartheta_{y_{n+2}}, \dots).$$

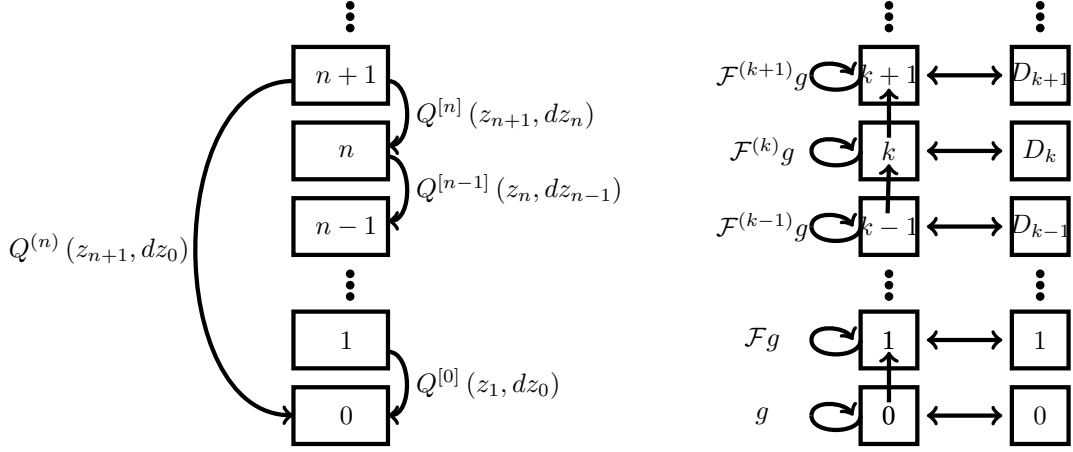


Figure 10: *Left*: The interaction chain that connects successive hierarchical levels downwards. The arrows on the right correspond to (4.20), the arrow on the left corresponds to (5.1). *Right*: The renormalisation transformation that connects successive hierarchical levels upwards. The vertical arrows correspond to (4.17). 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.

**Theorem 5.1. [Renormalised scaling]** *Let  $c_k$  be as in (2.5),  $e_k$  and  $K_k$  as in (2.10) and  $E_k$  as in (4.5). Define*

$$(5.3) \quad A_n = \frac{1}{2} \sum_{k=0}^{n-1} \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}, \quad n \in \mathbb{N}.$$

Then

$$(5.4) \quad \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \cdot) = (1 - \theta) \delta_{(0,0^{\mathbb{N}_0})} + \theta \delta_{(1,1^{\mathbb{N}_0})}$$

if and only if

$$(5.5) \quad \lim_{n \rightarrow \infty} A_n = \infty.$$

Moreover, if (5.5) holds, then for all  $g \in \mathcal{G}$ ,

$$(5.6) \quad \lim_{n \rightarrow \infty} A_n \mathcal{F}^{(n)} g = g_{\text{FW}} \quad \text{pointwise,}$$

with  $g_{\text{FW}}(x) = x(1 - x)$ ,  $x \in [0, 1]$ .

The proof of Theorem 5.1 is given in Section 13. The scaling factors  $A_n$  can be interpreted as clustering coefficients: in Section 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 (5.4) corresponds to the *clustering regime*. According to (5.6), even though  $A_n$  depends on the choice of the sequences  $\underline{K}$ ,  $\underline{e}$ ,  $\underline{c}$  in (2.5) and (2.10), the limit  $A_n \mathcal{F}^{(n)} g$  as  $n \rightarrow \infty$  is *universal*: irrespective of the choice of  $g \in \mathcal{G}$ , the limit is the standard Fisher-Wright diffusion function  $g_{\text{FW}}$ . Thus,  $g_{\text{FW}}$  is the *global attractor of the renormalisation transformation* (see Fig. 11).

## 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  $\Omega_N$  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  $\Omega_N$  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.

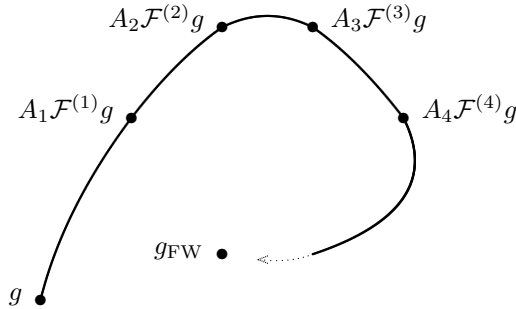


Figure 11: Flow of the iterates  $\mathcal{F}^{(n)}g$ ,  $n \in \mathbb{N}_0$ , of the renormalisation transformation acting on the class  $\mathcal{G}$ . After multiplication by  $A_n$ , the flow is globally attracted by  $g_{\text{FW}}$ .

- 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  $g \in \mathcal{G}$  [EF96], [Kle96]. 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  $t \rightarrow \infty$ , 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  $t \rightarrow \infty$  with  $(V(s))_{s \geq 0}$  the standard Fisher-Wright diffusion, irrespective of the choice of  $g \in \mathcal{G}$  [FG94], [FG96].
- For more general migration it is possible that mono-type clusters grow slower than any positive power of  $t$  as  $t \rightarrow \infty$ . 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 [DGV95].

From the perspective of explaining *universality* in  $g \in \mathcal{G}$  in the hierarchical mean-field limit  $N \rightarrow \infty$ , the above type of behaviour has been studied in detail in [DGV95] and [GdHKK14] 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  $\Omega_N$  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  $N \rightarrow \infty$ . To that end, we look at the interaction chain  $M_{-l(k)}^k$  for  $k \rightarrow \infty$ , where the *level scaling function*  $l: \mathbb{N}_0 \rightarrow \mathbb{N}_0$  is non-decreasing with  $\lim_{k \rightarrow \infty} l(k) = \infty$  and is suitably chosen such that we obtain a *non-trivial clustering limiting law*, i.e.,

$$(5.7) \quad \lim_{k \rightarrow \infty} \mathcal{L}[M_{-l(k)}^k] = \mathcal{L}[\hat{\theta}],$$

where the limiting sequence of random frequencies  $\hat{\theta}$  satisfies

$$(5.8) \quad 0 < \mathbb{P}(\hat{\theta} \in \{0^{\mathbb{N}_0}, 1^{\mathbb{N}_0}\}) < 1.$$

In line with [DGV95] and [DG96], 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  $l_\chi: \mathbb{N}_0 \rightarrow \mathbb{N}_0$ ,  $\chi \in I \subseteq [0, \infty)$ , called the *cluster scales*, satisfying (5.7)–(5.8):



- (1) **Fast clustering:**  $\lim_{k \rightarrow \infty} l_\chi(k)/k = 1$  for all  $\chi \in I$ .
- (2) **Diffusive clustering:**  $\lim_{k \rightarrow \infty} l_\chi(k)/k = \kappa(\chi)$  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) **Slow clustering:**  $\lim_{k \rightarrow \infty} l_\chi(k)/k = 0$  for all  $\chi \in I$ . (This regime borders with the regime of coexistence.)

We write  $(M_\chi^\infty)_{\chi \in I}$  with  $M_\chi^\infty = \lim_{k \rightarrow \infty} M_{-l_\chi(k)}^k$  to denote the *cluster process*.

**Remark 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  $\chi$  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.  $\blacksquare$

Recall (5.3). 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 5.3 below shows that, subject to (3.10) and (3.11), 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 [Arr79], [CG86] (voter model on  $\mathbb{Z}$ , respectively,  $\mathbb{Z}^2$ ) and in [DG93a], [FG94], [DGV95], [Kle96] (interacting Fleming-Viot processes on  $\Omega_N$  with  $N < \infty$ , respectively,  $N \rightarrow \infty$ ), all types of clustering in [DG96], [Kle97], [Win02] (interacting Feller diffusions on  $\Omega_N$  with  $N \rightarrow \infty$ ) and in [GdHKK14], [GdHKK18] (interacting Cannings processes on  $\Omega_N$  in non-random and random environment with  $N \rightarrow \infty$ ).

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  $(M_{-l_\chi(k)}^k)_{\chi \in I}$  converges as  $k \rightarrow \infty$  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

$$(5.9) \quad V_{-l}^k = \text{the conditional variance of the active part of } M_{-l}^k \text{ given } M_{-(l+1)}^k$$

and consider the sum of random variables  $A_{k,n} = \sum_{-(k+1) \leq -l \leq -n} V_{-l}^k$ ,  $n \in \mathbb{N}_0$ . In order for the system to cluster, we must have  $\lim_{k \rightarrow \infty} A_{k,n} = 0$  for every  $n \in \mathbb{N}_0$ . The *volatility profile* is given by

$$(5.10) \quad (p_\chi(k))_{\chi \in [0,1]}, \quad p_\chi(k) = A_{k,l_\chi(k)} / A_{k,0}.$$

This profile is a *random variable* that depends on the interaction chain up to  $M_{-l_\chi(k)}^k$ . 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

$$(5.11) \quad \lim_{k \rightarrow \infty} p_\chi(k) = 1 - \kappa(\chi), \quad \chi \in [0, 1],$$

while for fast clustering the limit is 0 and for slow clustering the limit is 1. From (5.6) we know that the scaled renormalised diffusion function  $A_n \mathcal{F}^{(n)}(g)$  tends to the standard Fisher-Wright diffusion function as  $n \rightarrow \infty$ . 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 5.1 allow us to compute the first and second moments of all the block averages, which will be done in Section 13.1 (Propositions 13.1–13.4). 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  $N \rightarrow \infty$ , we find that the variance of the active component  $x_m^n$  of  $M_{-m}^n$  equals

$$(5.12) \quad \text{Var}(x_m^n) = \mathbb{E}[(x_m^n - \vartheta_n)^2] = A_m^n (\mathcal{F}^{(n+1)}g)(\vartheta_n),$$

where

$$(5.13) \quad A_m^n = \frac{1}{2} \sum_{k=m}^n \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}.$$

(Note that  $A_n = A_0^{n-1}$ .) On the other hand, since  $x_m^n \in (0, 1)$  we have  $\text{Var}(x_m^n) \in (0, 1)$  and  $A_m^n (\mathcal{F}^{(n+1)}g)(\vartheta_n) \in (0, 1)$ . Taking  $m = 0$ , we get  $(\mathcal{F}^{(n+1)}g)(\vartheta_n) \in (0, \frac{1}{A_0^n})$ . This implies that

$$(5.14) \quad (\mathcal{F}^{(n+1)}g)(\vartheta_n) = \int_{[0,1]^2} (\mathcal{F}^m g)(x_m) Q_m^{(n)}((\vartheta_n, \theta_{y,n}), dz_m) \in \left(0, \frac{1}{A_0^n}\right).$$

Since  $\lim_{n \rightarrow \infty} A_n (\mathcal{F}^{(n+1)}g) = g_{\text{FW}}$ , for large enough  $m, n$  we can approximate  $\mathcal{F}^{(m)}g \approx g_{\text{FW}}/A_0^m$ . Therefore

$$(5.15) \quad \int \frac{g_{\text{FW}}}{A_0^m}(x_m) Q_m^{(n)}((\vartheta_n, \theta_{y,n}), dz_m) \in \left(0, \frac{1}{A_0^n}\right),$$

or, equivalently,

$$(5.16) \quad \int_{[0,1]^2} x_m(1-x_m) Q_m^{(n)}((\vartheta_n, \theta_{y,n}), dz_m) \in \left(0, \frac{A_0^m}{A_0^n}\right).$$

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  $m, n \rightarrow \infty$  says something about the speed at which monotype clusters form.

To capture the cluster growth, we must decide how we let  $m, n \rightarrow \infty$ . For this we look for clusters of radius  $l_\chi(n)$  with  $\chi \in I$ . Put

$$(5.17) \quad f^n(l_\chi(n)) = \frac{A_0^{l_\chi(n)}}{A_0^n},$$

and define, for  $\epsilon > 0$ ,

$$(5.18) \quad \mathcal{X}_\epsilon^n = \inf\{\chi \in I: f^n(l_\chi(n)) < \epsilon\}.$$

Then the three types of clustering correspond to:

- (1) **Fast clustering:**  $\lim_{n \rightarrow \infty} l_{\mathcal{X}_\epsilon^n}(n)/n = 1$ .
- (2) **Diffusive clustering:**  $\lim_{n \rightarrow \infty} l_{\mathcal{X}_\epsilon^n}(n)/n = R$  for some random variable  $R$  taking values in  $(0, 1)$ .
- (3) **Slow clustering:**  $\lim_{n \rightarrow \infty} l_{\mathcal{X}_\epsilon^n}(n)/n = 0$ .

In terms of the interaction chain starting from  $-k$  with  $k \rightarrow \infty$ , in view of (5.12) this corresponds to the variance in the interaction chain being concentrated near the beginning, being spread out or being concentrated near the end.

### 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

$$(5.19) \quad A_n = \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{c_k}$$

and (5.4) holds if and only if

$$(5.20) \quad \sum_{k \in \mathbb{N}_0} \frac{1}{c_k} = \infty.$$

Various subcases were analysed in [BCGdH95]. For the system with seed-bank because  $E_0 = 1$  and  $E_k < 1$  (see (4.5)), it follows from (5.3) that

$$(5.21) \quad A_n < \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{c_k}.$$

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 (3.10), or are *pure exponential*, as in (3.11). 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 5.3. [Rates of scaling for diffusion function]** *Let  $\rho$  be as defined in (2.14).*

(I) *If  $\rho < \infty$ , then (5.5) holds if and only if (5.20) hold, and*

$$(5.22) \quad A_n \sim \frac{1}{2(1+\rho)} \sum_{m=0}^{k-1} \frac{1}{c_m}.$$

(II) *If  $\rho = \infty$ , then (5.5) holds in the following cases:*

- *Subject to (3.10) if and only if  $-\phi \leq \alpha \leq 1$ , with*

$$(5.23) \quad \begin{aligned} -\phi < \alpha < 1: & \quad A_n \sim C_1 n^{\alpha+\phi}, \\ -\phi = \alpha < 1: & \quad A_n \sim C_2 \log n, \\ -\phi < \alpha = 1: & \quad A_n \sim C_3 \frac{n^{1+\phi}}{\log n}, \\ -\phi = \alpha = 1: & \quad A_n \sim C_4 \log \log n, \end{aligned}$$

where

$$(5.24) \quad C_1 = \frac{1}{2AF} \frac{1-\alpha}{\alpha+\phi}, \quad C_2 = \frac{1}{2AF} (1-\alpha), \quad C_3 = \frac{1}{2AF} \frac{1}{1+\phi}, \quad C_4 = \frac{1}{2AF}.$$

*The values of  $B, \beta$  play no role for the clustering, nor for the asymptotics.*

- *Subject to (3.11) if and only if  $Kc \leq 1 \leq K$ , with*

$$(5.25) \quad \begin{aligned} c < Ke, Kc < 1: & \quad A_n \sim \hat{C}_1 (Kc)^{-(n-1)}, \\ c < Ke, Kc = 1: & \quad A_n \sim \bar{C}_1 n, \\ c = Ke, Kc < 1: & \quad A_n \sim \hat{C}_2 (Kc)^{-(n-1)}, \\ c = Ke, Kc = 1: & \quad A_n \sim \bar{C}_2 n, \\ c > Ke, Kc < 1: & \quad A_n \sim \hat{C}_3 (Kc)^{-(n-1)}, \\ c > Ke, Kc = 1: & \quad A_n \sim \bar{C}_3 n, \\ c < 1 = K: & \quad A_n \sim \tilde{C}_1 n^{-1} c^{-(n-1)}, \\ c = 1 = K: & \quad A_n \sim \tilde{C}_2 \log n, \end{aligned}$$

where

$$(5.26) \quad \begin{aligned} \hat{C}_1 &= \frac{K-1}{2K(1-Kc)}, & \hat{C}_2 &= \frac{(K-1)^2}{2(2K-1)(1-Kc)}, & \hat{C}_3 &= \frac{K-1}{2(1-Kc)}, \\ \bar{C}_1 &= \frac{K-1}{2K}, & \bar{C}_2 &= \frac{(K-1)^2}{2(2K-1)}, & \bar{C}_3 &= \frac{K-1}{2}, \\ \tilde{C}_1 &= \frac{1}{2(1-c)}, & \tilde{C}_2 &= \frac{1}{2}. \end{aligned}$$

The value of  $e$  plays no role for the clustering, but does for the asymptotics.

The proof of Theorem 5.3 is given in Section 13. 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  $\lim_{k \rightarrow \infty} E_k = 0$  the growth rate of  $A_n$  is *strictly smaller* than without seed-bank.

Furthermore, subject to (3.10), 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 (3.11), 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 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 5.2 can in principle be deduced from the asymptotics of  $A_n$  as  $n \rightarrow \infty$  in Theorem 5.3. It would be interesting to work out the details and to identify the limiting processes that control the cluster growth.

## Part II

# PREPARATIONS AND PROOFS

## 6 Proofs: $N < \infty$ , identification of clustering regime

In this section we prove Theorems 3.2–3.3. The integral criterion for  $\rho = \infty$  in (3.8) is explained in Section 6.1. Theorem 3.2 is proved in Section 6.2 and Theorem 3.3 in Section 6.3.

### 6.1 Explanation of clustering criterion for infinite seed-bank

Recall Fig. 6. Suppose that  $g = dg_{\text{FW}}$ , so that we have a dual. We will show that the integral criterion in (3.8) 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 2.9). We argue heuristically as follows. If  $\rho = \infty$ , then the activity times  $\sigma_k$  are much smaller than the sleeping times  $\tau_k$ , and we may assume that  $\tau_k + \sigma_k \asymp \tau_k$ ,  $k \rightarrow \infty$ . Discretising time, we can use the results from [AB16] for the intersection of two independent renewal processes. Then the integral criterion in (3.8) 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  $\asymp \varphi(s)^{-1} s^{-(1-\gamma)}$  [AB16]. 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  $\asymp a_{\varphi(s)^{-1} s^\gamma}^{(N)}(0, 0)$ . Hence the total hazard is  $\asymp \int_1^\infty ds [\varphi(s)^{-1} s^{-(1-\gamma)}]^2 a_{\varphi(s)^{-1} s^\gamma}^{(N)}(0, 0)$ . After the transformation  $t = t(s) = \varphi(s)^{-1} s^\gamma$ , the latter turns into the integral in (3.8), modulo a constant. When carrying out this transformation,

we need that  $s\varphi'(s)/\varphi(s) \rightarrow 0$ , which follows from (3.7), and  $\varphi(t(s))/\varphi(s) \asymp 1$ , which follows from the bound we imposed on  $\psi$  in (3.7) together with the fact that  $\log \varphi(s)/\log s \rightarrow 0$ . This computation is spelled out in Appendix A.

- If  $\gamma = 1$ , then the probability for each of the lineages to be active at time  $s$  decays like  $\hat{\varphi}(s)^{-1}$  [AB16], and so the total time they are active up to time  $s$  is  $\asymp s\hat{\varphi}(s)^{-1}$ . Recall from (3.6) that  $\hat{\varphi}(t) = \mathbb{E}[\tau \wedge t]$  is also slowly varying.) Hence the total hazard is  $\asymp \int_1^\infty ds [\hat{\varphi}(s)^{-1}]^2 a_{\hat{\varphi}(s)^{-1}s}^{(N)}(0, 0)$ . After the transformation  $t = t(s) = \hat{\varphi}(s)^{-1}s$  (for which we can use the same type of computation as in Appendix A), the latter turns into the integral in (3.8), modulo a constant.

## 6.2 Scaling of wake-up time and migration kernel for infinite seed-bank

We can prove Theorem 3.2 by direct computation via assumptions (3.10)–(3.11). We start by computing  $\gamma$ . Afterwards we compute  $\hat{\varphi}(t)$  and  $a_t^{\Omega_N}(0, 0)$ .

**Computation of  $\gamma$ .** Recall (2.39), which reads

$$(6.1) \quad \mathbb{P}(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m} e^{-(e_m/N^m)t}.$$

Since we are interested in the asymptotic behaviour of  $\mathbb{P}(\tau > t)$  as  $t \rightarrow \infty$ , 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 (6.1). Hence we can estimate the latter by an integral and insert the assumptions made in (3.10)–(3.11). 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  $\gamma$  after extracting the  $t$ -dependence:

$$(3.10) \quad \implies \quad \gamma = 1, \quad \varphi(t) \asymp (\log t)^{-\alpha},$$

$$(6.2) \quad (3.11) \quad \implies \quad \gamma = \gamma_{N,K,e} = \frac{\log(N/Ke)}{\log(N/e)}, \quad \varphi(t) \asymp 1.$$

In order to guarantee that  $\rho = \infty$ , we must require that  $\alpha \in (-\infty, 1]$ , respectively,  $K \in [1, \infty)$  (while  $\beta$ , respectively,  $e$  play no role). Subject to (3.10),

$$(6.3) \quad \hat{\varphi}(t) \asymp \begin{cases} (\log t)^{1-\alpha}, & \alpha \in (-\infty, 1), \\ \log \log t, & \alpha = 1, \end{cases}$$

while subject to (3.11),

$$(6.4) \quad \hat{\varphi}(t) \asymp \begin{cases} 1, & K \in (1, \infty), \\ \log t, & K = 1. \end{cases}$$

**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 (2.6) as

$$(6.5) \quad a^{\Omega_N}(0, \eta) = \frac{r_{\|\eta\|}}{N^{\|\eta\|-1}(N-1)}$$

with

$$(6.6) \quad r_{\|\eta\|} = \frac{1}{D(N)} \frac{N-1}{N} \sum_{l \geq \|\eta\|} \frac{c_{l-1}}{N^{l-1}} \frac{1}{N^{l-\|\eta\|}},$$

where  $D(N)$  is a renormalisation constant such that  $\sum_{j \in \mathbb{N}} r_j = 1$ . For transition kernels of the form (6.5), the time- $t$  transition kernel  $a_t^{\Omega_N}(\cdot, \cdot)$  was computed in [FG94] with the help of Fourier analysis, see also [DGW05]. Namely,

$$(6.7) \quad a_t^{\Omega_N}(0, \eta) = \sum_{j \geq k} K_{jk}(N) \frac{\exp[-h_j(N)t]}{N^j}, \quad t \geq 0, \quad \eta \in \Omega_N: d_{\Omega_N}(0, \eta) = k \in \mathbb{N}_0,$$

where

$$(6.8) \quad K_{jk}(N) = \begin{cases} 0, & j = k = 0, \\ -1, & j = k > 0, \\ N - 1, & \text{otherwise,} \end{cases} \quad j, k \in \mathbb{N}_0,$$

and

$$(6.9) \quad h_j(N) = \frac{N}{N-1} r_j(N) + \sum_{i>j} r_i(N), \quad j \in \mathbb{N}.$$

The expressions in (6.6)–(6.9) simplify considerably in the limit as  $N \rightarrow \infty$ , namely, the term with  $i = j$  dominates and

$$(6.10) \quad h_j(N) \sim r_j(N) \sim \frac{c_{j-1}}{D(N)N^{j-1}}, \quad j \in \mathbb{N}, \quad D(N) \sim c_0.$$

We show why this is true for  $h_j(N)$  (the argument for  $r_j(N)$  and  $D(N)$  is similar). Write

$$(6.11) \quad \begin{aligned} h_j(N) &= \frac{N}{N-1} r_j(N) + \sum_{i>j} r_i(N) \\ &= \frac{1}{D(N)} \left( \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} \frac{1}{N^{l-j}} + \frac{N-1}{N} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} \sum_{i<j \leq l} \frac{1}{N^{l-i}} \right) \\ &= \frac{1}{D(N)} \frac{c_{j-1}}{N^{j-1}} \left( 1 + \left[ 1 + O\left(\frac{1}{N}\right) \right] \left( \frac{c_{j-1}}{N^{j-1}} \right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} \right). \end{aligned}$$

Hence it suffices to show that

$$(6.12) \quad \limsup_{N \rightarrow \infty} \left( \frac{c_{j-1}}{N^{j-1}} \right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} = 0, \quad j \in \mathbb{N}.$$

To do so, note that, since  $\limsup_{k \rightarrow \infty} \frac{1}{k} \log c_k < \log N$  by (2.7), for  $N$  large enough we have

$$(6.13) \quad \sup_{k \in \mathbb{N}_0} c_k^{1/k} < N.$$

Let  $\bar{N} = \inf\{N \in \mathbb{N} : \sup_{k \in \mathbb{N}_0} c_k^{1/k} < N\}$ . Then

$$(6.14) \quad \begin{aligned} \limsup_{N \rightarrow \infty} \left( \frac{c_{j-1}}{N^{j-1}} \right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} &\leq \limsup_{N \rightarrow \infty} \frac{1}{c_{j-1}} \sum_{l>j} \frac{\bar{N}^{l-1}}{N^{l-1}} N^{j-1} \\ &= \frac{\bar{N}^{j-1}}{c_{j-1}} \limsup_{N \rightarrow \infty} \frac{\bar{N}}{1 - \frac{\bar{N}}{N}} = 0, \quad j \in \mathbb{N}, \end{aligned}$$

which settles (6.12).

To understand what (6.9) gives for finite  $N$ , note that for asymptotically polynomial coefficients (recall (3.10))

$$(6.15) \quad \begin{aligned} \left( \frac{c_{j-1}}{N^{j-1}} \right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} &= [1 + o(1)] \frac{N^{j-1}}{F(j-1)^{-\phi}} \sum_{l>j} \frac{F(l-1)^{-\phi}}{N^{l-1}} \\ &= [1 + o(1)] \sum_{l>j} \frac{(l-1)^{-\phi}}{(j-1)^{-\phi}} N^{-(l-j)} = [1 + o(1)] \sum_{k \geq 1} \left( 1 + \frac{k}{j-1} \right)^{-\phi} N^{-k}, \quad j \in \mathbb{N}. \end{aligned}$$

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 (3.11)),

$$(6.16) \quad \left( \frac{c_{j-1}}{N^{j-1}} \right)^{-1} \sum_{l>j} \frac{c_{l-1}}{N^{l-1}} = \sum_{k \geq 1} \left( \frac{c}{N} \right)^{-k} = \frac{c}{N-c}.$$

Hence, for both choices of coefficients we have the following:

$$(6.17) \quad \begin{array}{l} \text{For } N \rightarrow \infty \text{ the quantities } h_j(N), r_j(N) \text{ are bounded from above} \\ \text{and below by positive finite constants times the right-hand side of} \\ \text{(6.10) uniformly in } j \in \mathbb{N}. \end{array}$$

Picking  $\eta = 0$  ( $k = 0$ ) in (6.7), we obtain

$$(6.18) \quad a_t^{\Omega_N}(0, 0) = \sum_{j \in \mathbb{N}} (N-1) \frac{\exp[-h_j(N)t]}{N^j}.$$

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 (6.18) by an integral. To do so, we change variables by putting  $s = h_j(N)$  and exploit (6.17). Take the logarithm to express  $j$  in terms of  $s$ , compute  $ds/dj$ , and extract the  $t$ -dependence. This gives

$$(6.19) \quad (3.10) \implies a_t^{\Omega_N}(0, 0) \asymp t^{-1} \log^\phi t,$$

$$(3.11) \implies a_t^{\Omega_N}(0, 0) \asymp t^{-1-\delta_{N,c}},$$

where

$$(6.20) \quad \delta_{N,c} = \frac{\log c}{\log(N/c)}.$$

### 6.3 Hierarchical clustering

In this section we prove Theorem 3.3 by substituting the results of Theorem 3.2 into the clustering criterion in (3.8).

Combining (3.9), (6.2)–(6.4) and (6.19)–(6.20), we find the following clustering criterion for *fixed*  $N$  and infinite seed-bank:

- Subject to (3.10), clustering prevails if and only if

$$(6.21) \quad -\phi \leq \alpha \leq 1.$$

- Subject to (3.11), clustering prevails if and only

$$(6.22) \quad \delta_{N,c} \leq -\frac{1 - \gamma_{N,K,e}}{\gamma_{N,K,e}}.$$

In view of (6.2) and (3.14), the condition in (6.22) amounts to

$$(6.23) \quad \log N \times \log(Kc) \leq \log c \times \log(K^2e),$$

where we use that  $c < N$  and  $Ke < N$  (recall (2.7) and (2.12)). The condition in (6.23) holds for all  $N$  when

$$(6.24) \quad Kc = 1 \text{ with } \begin{cases} c = 1, & K^2e \in (0, \infty), \\ c > 1, & K^2e \geq 1, \\ c < 1, & K^2e \leq 1. \end{cases}$$

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

$$(6.25) \quad Kc \leq 1 \leq K,$$

which is the analogue of (6.21).

## 7 Preparation: $N \rightarrow \infty$ , McKean-Vlasov process and mean-field system

To analyse the scaling of our hierarchical system in the hierarchical mean-field limit  $N \rightarrow \infty$ , 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:

1. McKean-Vlasov process (Section 7.1).
2. Mean-field system and McKean-Vlasov limit (Section 7.2).

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 10–12 and the full system with *infinitely many hierarchies* in Section 12. The proofs of the propositions stated in this section will be given in Sections 7.3 and 7.4.

### 7.1 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 7.2–8.1. (In the full system the effective process introduced in (4.9) 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

$$(7.1) \quad z(t) = (x(t), y(t))_{t \geq 0},$$

taking values in  $[0, 1]^2$ , with initial law  $\mathcal{L}[(x(0), y(0))] = \mu$  and with components evolving according to

$$(7.2) \quad \begin{aligned} dx(t) &= c [\mathbb{E}[x(t)] - x(t)] dt + \sqrt{g(x(t))} dw(t) + Ke [y(t) - x(t)] dt, \\ dy(t) &= e [x(t) - y(t)] dt, \end{aligned}$$

where  $\mathbb{E}$  denotes expectation with respect to  $\mu$ . With the help of Itô-calculus we can compute the expectation  $\mathbb{E}[x(t)]$ . Indeed, from (7.2) we get

$$(7.3) \quad \begin{aligned} \frac{d}{dt} \mathbb{E}[x(t)] &= Ke [\mathbb{E}[y(t)] - \mathbb{E}[x(t)]], \\ \frac{d}{dt} \mathbb{E}[y(t)] &= e [\mathbb{E}[x(t)] - \mathbb{E}[y(t)]]. \end{aligned}$$

Define

$$(7.4) \quad \theta_x = \mathbb{E}^\mu[x(0)], \quad \theta_y = \mathbb{E}^\mu[y(0)], \quad \theta = \mathbb{E}^\mu \left[ \frac{x(0) + Ky(0)}{1 + K} \right].$$

Note that (7.3) implies that  $\theta$  is a preserved quantity, i.e.,

$$(7.5) \quad \mathbb{E}^\mu \left[ \frac{x(0) + Ky(0)}{1 + K} \right] = \mathbb{E}^\mu \left[ \frac{x(t) + Ky(t)}{1 + K} \right] = \theta, \quad t \geq 0.$$

Solving (7.3), we find

$$(7.6) \quad \begin{aligned} \mathbb{E}[x(t)] &= \theta + \frac{K}{1 + K} (\theta_x - \theta_y) e^{-(K+1)et}, \\ \mathbb{E}[y(t)] &= \theta - \frac{1}{1 + K} (\theta_x - \theta_y) e^{-(K+1)et}. \end{aligned}$$

In particular, from (4.51) we see that

$$(7.7) \quad \lim_{t \rightarrow \infty} (\mathbb{E}[x(t)], \mathbb{E}[y(t)]) = (\theta, \theta).$$

Hence, in equilibrium we can replace  $\mathbb{E}[x(t)]$  in (7.2) by  $\theta$ . After inserting (7.6) into (7.2), we can use [YW71, 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 (7.2) has a unique strong solution. We will refer to this solution as the *McKean-Vlasov process*.



**Remark 7.1. [Self-consistency]** To prove uniqueness of the solution to (7.2) we can also use [Gar88], where self-consistent mean-field dynamics are treated in detail. The solution has the Feller property.  $\blacksquare$

**Proposition 7.2. [Single-colony McKean-Vlasov process: equilibrium]** *For every initial law  $\mu \in \mathcal{P}([0, 1]^2)$  satisfying*

$$(7.8) \quad \mathbb{E}^\mu \left[ \frac{x(0) + Ky(0)}{1 + K} \right] = \theta, \quad \theta \in [0, 1],$$

*the process in (7.1) converges to a unique equilibrium,*

$$(7.9) \quad \lim_{t \rightarrow \infty} \mathcal{L}[(x(t), y(t))] = \Gamma_\theta,$$

*and*

$$(7.10) \quad \Gamma_\theta \in \mathcal{P}([0, 1]^2),$$

*satisfies*

$$(7.11) \quad \theta = \int_{[0, 1]^2} x \Gamma_\theta(dx, dy) = \int_{[0, 1]^2} y \Gamma_\theta(dx, dy).$$

The proof of Proposition 7.2 is given in Section 7.3. Note that  $\Gamma_\theta = \Gamma_\theta^{g, c, K, e}$  depends on all the parameters appearing in (7.2). In Section 8 we will see that  $\Gamma_\theta$  is continuous as a function of  $\theta$ .

**Remark 7.3. [Non-linear Markov process]** Note that (7.1) 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 (7.2). This is different from the model without seed-bank, where the non-linearity is replaced by a drift towards  $\theta$  that is constant in time. In equilibrium we can replace  $\mathbb{E}[x(t)]$  by  $\theta$  in (7.2), but before equilibrium is reached we cannot, because  $t \mapsto \mathbb{E}[x(t)]$  is not constant, as is clear from (4.51). 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 13.  $\blacksquare$

## 7.2 Mean-field system and McKean-Vlasov limit

In this section we consider a simplified version of the SSDE in (2.20), namely, we restrict to the finite geographic space

$$(7.12) \quad [N] = \{0, 1, \dots, N-1\}, \quad N \in \mathbb{N}.$$

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

$$(7.13) \quad S = \mathfrak{s}^{[N]}, \quad \mathfrak{s} = [0, 1]^2,$$

the system is denoted by

$$(7.14) \quad Z^{[N]}(t) = (X^{[N]}(t), Y^{[N]}(t))_{t \geq 0}, \quad (X^{[N]}(t), Y^{[N]}(t)) = (x_i^{[N]}(t), y_i^{[N]}(t))_{i \in [N]},$$

and its components evolve according to the SSDE

$$(7.15) \quad \begin{aligned} dx_i^{[N]}(t) &= \frac{c}{N} \sum_{j \in [N]} [x_j^{[N]}(t) - x_i^{[N]}(t)] dt + \sqrt{g(x_i^{[N]}(t))} dw_i(t) + Ke [y_i^{[N]}(t) - x_i^{[N]}(t)] dt, \\ dy_i^{[N]}(t) &= e [x_i^{[N]}(t) - y_i^{[N]}(t)] dt, \quad i \in [N], \end{aligned}$$

which is the special case of (2.20) 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.,

$$(7.16) \quad \mathcal{L}[X^{[N]}(0), Y^{[N]}(0)] = \mu^{\otimes [N]}, \quad \mu \in \mathcal{P}([0, 1]^2).$$

By [SS80, Theorem 3.1], the SSDE in (7.15) is the unique weak solution of a well-posed martingale problem. By [SS80, Theorem 3.2], for every deterministic initial state  $(X^{[N]}(0), Y^{[N]}(0))$ , (7.15) has a unique strong solution. We are interested in the limit  $N \rightarrow \infty$ . For the limiting process we define

$$(7.17) \quad (Z(t))_{t \geq 0} = (X(t), Y(t))_{t \geq 0} = ((x_i(t), y_i(t))_{i \in \mathbb{N}_0})_{t \geq 0}$$

with components evolving according to (7.2), i.e.,

$$(7.18) \quad \begin{aligned} dx_i(t) &= c [\mathbb{E}[x_i(t)] - x_i(t)] dt + \sqrt{g(x_i(t))} dw(t) + Ke [y_i(t) - x_i(t)] dt, \\ dy_i(t) &= e [x_i(t) - y_i(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

with  $\mathcal{L}[(X(0), Y(0))] = \mu$  for some exchangeable  $\mu \in \mathcal{P}([0, 1]^2)^{\otimes [\mathbb{N}_0]}$ . Note that (7.18) consists of i.i.d. copies of the single-colony McKean-Vlasov process in (7.1), labelled by  $i \in \mathbb{N}_0$ .

**Proposition 7.4. [Infinite-system McKean-Vlasov limit: convergence]**

Suppose that  $\mathcal{L}[(X^{[N]}(0), Y^{[N]}(0))] = \mu^{[N]}$  is exchangeable and

$$(7.19) \quad \theta = \mathbb{E}^{\mu^{[N]}} \left[ \frac{x(0) + Ky(0)}{1 + K} \right].$$

Then

$$(7.20) \quad \lim_{N \rightarrow \infty} \mathcal{L}[(X^{[N]}(t), Y^{[N]}(t))_{t \geq 0}] = \mathcal{L}[(X(t), Y(t))_{t \geq 0}]$$

with

$$(7.21) \quad \mathcal{L}[(X(0), Y(0))_{t \geq 0}] = \mu, \quad \mu = \lim_{N \rightarrow \infty} \mu^{[N]},$$

where the limit is the McKean-Vlasov process in (7.1)–(7.2).

The proof of Proposition 7.4 is given in Section 7.4. For the system without seed-bank the McKean-Vlasov limit was proved in [Gar88]. The fact that the components decouple is a property referred to as *propagation of chaos*.

### 7.3 Proof of equilibrium and ergodicity

In this section we prove Proposition 7.2.

*Proof.* Note that, by (4.51), we can rewrite (7.2) as

$$(7.22) \quad \begin{aligned} dx(t) &= c \left[ \theta + \frac{K}{1 + K} (\theta_x - \theta_y) e^{-(K+1)et} - x(t) \right] dt + \sqrt{g(x(t))} dw(t) \\ &\quad + Ke [y(t) - x(t)] dt, \\ dy(t) &= e [x(t) - y(t)] dt. \end{aligned}$$

Existence and uniqueness of a strong solution is again standard (see e.g. [YW71, Theorem 1] and recall Remark 7.1). 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 (7.22), with  $\mathcal{L}[(x_1(0), y_1(0))] = \mu_1$  and  $\mathcal{L}[(x_2(0), y_2(0))] = \mu_2$ , where  $\mu_1$  and  $\mu_2$  satisfy

$$(7.23) \quad \mathbb{E}^{\mu_1} \left[ \frac{x_1(0) + Ky_1(0)}{1 + K} \right] = \theta = \mathbb{E}^{\mu_2} \left[ \frac{x_2(0) + Ky_2(0)}{1 + K} \right]$$

for some  $\theta \in [0, 1]$ . Write

$$(7.24) \quad \theta_{x_1} = \mathbb{E}^{\mu_1}[x_1(0)], \quad \theta_{y_1} = \mathbb{E}^{\mu_1}[y_1(0)], \quad \theta_{x_2} = \mathbb{E}^{\mu_2}[x_2(0)], \quad \theta_{y_2} = \mathbb{E}^{\mu_2}[y_2(0)].$$

Couple the two systems by coupling their Brownian motions. Denote the coupled process by

$$(7.25) \quad \begin{aligned} (\bar{z}(t))_{t \geq 0} &= (z_1(t), z_2(t))_{t \geq 0}, \quad z_1(t) = (x_1(t), y_1(t)), \quad z_2(t) = (x_2(t), y_2(t)), \\ \mathcal{L}(\bar{z}(0)) &= \mu_1 \times \mu_2, \end{aligned}$$

which has a unique strong solution. Put

$$(7.26) \quad \Delta(t) = x_1(t) - x_2(t), \quad \delta(t) = y_1(t) - y_2(t).$$

To show that the equilibrium is unique, it is enough to show that

$$(7.27) \quad \lim_{t \rightarrow \infty} \mathbb{E}[|\Delta(t)| + EK|\delta(t)|] = 0.$$

Using a generalised form of Itô's formula, we find

$$(7.28) \quad \begin{aligned} d|\Delta(t)| &= (\text{sign } \Delta(t)) d\Delta(t) + dL_t^0 \\ &= (\text{sign } \Delta(t)) c \left[ \frac{K}{1+K} ((\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})) e^{-(K+1)et} - \Delta(t) \right] dt \\ &\quad + (\text{sign } \Delta(t)) \left( \sqrt{g(x_1(t))} - \sqrt{g(x_2(t))} \right) dw(t) \\ &\quad + (\text{sign } \Delta(t)) Ke [\delta(t) - \Delta(t)] dt, \end{aligned}$$

where we use that the local time  $L_t^0$  (see [RW00, Section IV.43]) of  $\Delta(t)$  at 0 equals 0, since  $g$  is Lipschitz (see [RW00, Proposition V.39.3]). Again using Itô's formula, we also find

$$(7.29) \quad d|\delta(t)| = (\text{sign } \delta(t)) d\delta(t) = (\text{sign } \delta(t)) e [\Delta(t) - \delta(t)] dt.$$

Taking expectations in (7.28)–(7.29), we get

$$(7.30) \quad \begin{aligned} &\frac{d}{dt} \mathbb{E}[|\Delta(t)| + K|\delta(t)|] \\ &= \mathbb{E} \left[ c \left[ (\text{sign } \Delta(t)) \frac{K}{1+K} ((\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})) e^{-(K+1)et} - |\Delta(t)| \right] \right] \\ &\quad + Ke \mathbb{E} \left[ (\text{sign } \Delta(t) - \text{sign } \delta(t)) (\delta(t) - \Delta(t)) \right] \\ &= \mathbb{E} \left[ c (\text{sign } \Delta(t)) \frac{K}{1+K} ((\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})) e^{-(K+1)et} \right] \\ &\quad - c \mathbb{E}[|\Delta(t)|] \\ &\quad - 2Ke \mathbb{E} \left[ 1_{\{\text{sign } \delta(t) \neq \text{sign } \Delta(t)\}} (|\delta(t)| + |\Delta(t)|) \right]. \end{aligned}$$

Define

$$(7.31) \quad h(t) = c \mathbb{E}[|\Delta(t)|] + 2Ke \mathbb{E} \left[ 1_{\{\text{sign } \delta(t) \neq \text{sign } \Delta(t)\}} (|\delta(t)| + |\Delta(t)|) \right].$$

Then  $h(t)$  satisfies

1.  $h(t) > 0$ .
2.  $0 \leq \int_0^\infty dt h(t) \leq 1 + K + c |(\theta_{x_1} - \theta_{x_2}) - (\theta_{y_1} - \theta_{y_2})| \frac{K}{K+1} \frac{1}{e^{(K+1)}} [1 - e^{-(K+1)e}]$ .
3.  $h$  is differentiable with  $h'$  bounded (see [GHO20, Appendix D]).

Hence it follows that  $\lim_{t \rightarrow \infty} h(t) = 0$ , which implies that

$$(7.32) \quad \lim_{t \rightarrow \infty} \mathbb{E}[|\Delta(t)|] = 0.$$

We are left to prove that  $\lim_{t \rightarrow \infty} \mathbb{E}[|\delta(t)|] = 0$ . To do so, we define

$$(7.33) \quad f(t) = \mathbb{E}[|\delta(t)|], \quad G(t) = e \mathbb{E}[(\text{sign } \delta(t))\Delta(t)].$$

Note that  $G$  is bounded and continuous. Taking expectations in (7.29), we find

$$(7.34) \quad \frac{d}{dt} f(t) = -e f(t) + G(t),$$

Solving (7.34) explicitly, we find that

$$(7.35) \quad f(t) = f(r) e^{-e(t-r)} + \int_r^t ds e^{-e(t-s)} G(s), \quad r, t \in \mathbb{R}, t > r \geq 0.$$

By (7.32), 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

$$(7.36) \quad f(t) \leq f(r) e^{-e(t-r)} + \epsilon$$

and, since  $|f| < 1$ , we find, for each  $\epsilon > 0$ ,

$$(7.37) \quad \lim_{t \rightarrow \infty} f(t) < \epsilon.$$

Therefore  $\lim_{t \rightarrow \infty} \mathbb{E}[|\delta(t)|] = 0$ , which completes the proof of uniqueness of the equilibrium for given  $\theta$ .

To prove existence of the equilibrium, let  $(t_n)_{n \in \mathbb{N}}$  be any increasing sequence of times such that  $\lim_{n \rightarrow \infty} t_n = \infty$ . Let  $\mu = \mathcal{L}[(x(0), y(0))]$  be any initial measure of the system in (7.22) with  $\mathbb{E}^\mu \left[ \frac{x(0) + Ky(0)}{1+K} \right] = \theta$ , 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  $\nu = \lim_{k \rightarrow \infty} \mu(t_{n_k})$ . We will show that  $\nu$  is invariant. To that end, recall from Section 7.1 that

$$(7.38) \quad \mathbb{E}^\mu \left[ \frac{x(t) + Ky(t)}{1+K} \right] = \theta, \quad t \geq 0.$$

Hence we can use the coupling in (7.25) to show that the system starting in  $\mu$  and the system starting  $\mu(t)$  converge to the same law as  $t \rightarrow \infty$ , from which it follows that  $\lim_{k \rightarrow \infty} \mu(t + t_{n_k}) = \nu$ . Let  $(S_t)_{t \geq 0}$  denote the semigroup of the system in (7.22). By the Feller property for semigroups,

$$(7.39) \quad S_t \nu = \lim_{k \rightarrow \infty} S_t \mu(t_{n_k}) = \lim_{k \rightarrow \infty} S_{t_{n_k}} (S_t \mu) = \nu,$$

where in the last equality we use the uniqueness of the equilibrium given  $\theta$ . Thus,  $\nu$  is an invariant measure. To exhibit its dependence on  $\theta$  we write  $\nu_\theta$ . Using the same coupling as in (7.25), and starting from  $\mu \times \nu_\theta$  with  $\nu_\theta$  the invariant measure just obtained, we see that for every  $\theta$  the system in (7.2) converges to a unique equilibrium measure  $\nu_\theta$ , and so (7.11) is immediate from (7.38).  $\square$

## 7.4 Proof of McKean-Vlasov limit

In this section we give a sketch of the proof of Proposition 7.4. In Chapters 8-12 we encounter more difficult versions of Proposition 7.4. 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

$$(7.40) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j \in [N]} x_j(0) = \mathbb{E}^{\mu(0)}[x_0] \quad \mathbb{P}\text{-a.s.}$$

By Ioffe's theorem [DGV95, 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 [JM86] we must show propagation of chaos and prove the weak law of large numbers

$$(7.41) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j \in [N]} x_j(t) = \mathbb{E}[x_0(t)].$$

The propagation of chaos and the weak law of large numbers for  $t > 0$  therefore follows from [Gar88, Section 4]. Since the martingale problem is well-posed [Gar88, Section 2], the limiting process exists and is unique.  $\square$

## 8 Proofs: $N \rightarrow \infty$ , mean-field finite-systems scheme

In Sections 8.1 we introduce the so called mean-field finite-systems scheme for the mean-field system introduced in Section 7.2. In Section 8.2 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 9.

### 8.1 Mean-field finite-systems scheme

In this section we describe the limiting dynamics of the finite system in (7.14) 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

$$(8.1) \quad \begin{aligned} dx_i(t) &= c[\theta - x_i(t)] dt + \sqrt{g(x_i(t))} dw_i(t) + Ke[y_i(t) - x_i(t)] dt, \\ dy_i(t) &= e[x_i(t) - y_i(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

where  $\theta$  is defined in (7.11). Note that each component of (8.1) is an autonomous copy of the McKean-Vlasov process in (7.2) in equilibrium.

For the multiscale analysis we will need the following ingredients:

1. The *estimator* for the finite system is defined by

$$(8.2) \quad \bar{\Theta}^{[N]}(t) = \bar{\Theta}^{[N]}(Z^{[N]}(t)) = \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N]}(t) + Ky_i^{[N]}(t)}{1 + K}$$

and its active and dormant counterparts

$$(8.3) \quad \begin{aligned} \bar{\Theta}_x^{[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(t), \\ \bar{\Theta}_y^{[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} y_i^{[N]}(t). \end{aligned}$$

2. The *time scale*  $N$ , on which  $\lim_{N \rightarrow \infty} \mathcal{L}[\bar{\Theta}^{[N]}(L(N)) - \bar{\Theta}^{[N]}(0)] = \delta_0$  for all  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , but not for  $L(N) = N$ . In words,  $N$  is the time scale on which  $\bar{\Theta}^{[N]}(\cdot)$  starts evolving, i.e.,  $(\bar{\Theta}^{[N]}(Ns))_{s>0}$  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”.
3. The *invariant measure*, i.e., the equilibrium measure of a single component in (8.1) written

$$(8.4) \quad \Gamma_\theta,$$

and the *invariant measure* of the infinite system in (8.1), written  $\nu_\theta = \Gamma_\theta^{\otimes \mathbb{N}_0}$ , with  $\theta \in [0, 1]$  controlled by the initial measure (recall (7.4)–(4.51)).

4. The *renormalisation transformation*  $\mathcal{F}: \mathcal{G} \rightarrow \mathcal{G}$ ,

$$(8.5) \quad (\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x) \nu_\theta(dx, dy_0), \quad \theta \in [0, 1],$$

where  $\nu_\theta$  is the equilibrium measure of 8.1. Note that  $\mathcal{F}$  is the same transformation as defined in (4.16), but for the truncated system. Note that we can also write

$$(8.6) \quad (\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x) \Gamma_\theta(dx, dy_0), \quad \theta \in [0, 1],$$

where  $\Gamma_\theta$  is as defined in (8.4).

5. The *macroscopic observable*  $(\bar{\Theta}(s))_{s>0}$  satisfying the SSDE

$$(8.7) \quad d\bar{\Theta}(s) = \frac{1}{1+K} \sqrt{\mathbb{E}^{\Gamma_{\bar{\Theta}(s)}}[g(u)]} dw(s) = \frac{1}{1+K} \sqrt{(\mathcal{F}g)(\bar{\Theta}(s))} dw(s),$$

To obtain the multi-scale limit dynamics for the system in (7.14), we speed up time by a factor  $N$  and define the process

$$(8.8) \quad \left( x_1^{[N]}(s), y_1^{[N]}(s) \right)_{s>0} = \left( \Theta_x^{[N]}(Ns), \Theta_y^{[N]}(Ns) \right)_{s>0},$$

which is the analogue of the 1-block average in (2.22). We use the lower index 1 to indicate that the average is taken over  $[N]$  components. Using (7.15), we see that the dynamics of (8.8) is given by the SSDE

$$(8.9) \quad \begin{aligned} dx_1^{[N]}(s) &= \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i(Ns))} dw(s) + NKe \left[ y_1^{[N]}(s) - x_1^{[N]}(s) \right] ds, \\ dy_1^{[N]}(s) &= Ne \left[ x_1^{[N]}(s) - y_1^{[N]}(s) \right] ds. \end{aligned}$$

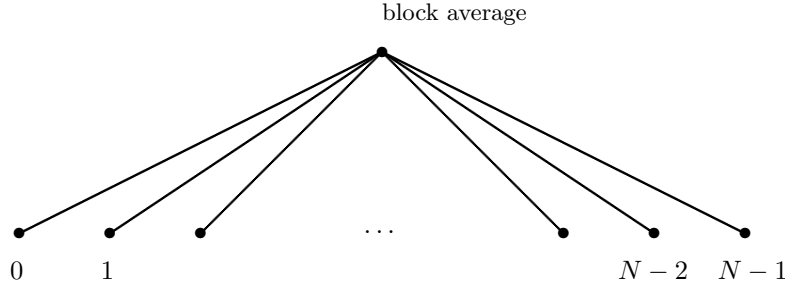


Figure 12: Given the value of the block average, the  $N \gg 1$  constituent components equilibrate on a time scale that is fast with respect to the time scale on which the block average fluctuates. Consequently, the volatility of the block average is the expectation of the volatility of the constituent components under the conditional quasi-equilibrium.

In (8.9), in the limit as  $N \rightarrow \infty$  infinite rates appear in the exchange between the active and the dormant population. However, looking at the process

$$(8.10) \quad \left( \frac{x_1^{[N]}(s) + Ky_1^{[N]}(s)}{1+K} \right)_{s>0} = \left( \bar{\Theta}^{[N]}(Ns) \right)_{s>0}$$

we see that the terms carrying a factor  $N$  in front cancel out. Consequently, for the process in (8.10) we can use ideas from [DG93a] to prove tightness as  $N \rightarrow \infty$  in the *classical topology* of continuum path processes. We will show in Section 9.3 that

$$(8.11) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \left[ x_1^{[N]}(s) - y_1^{[N]}(s) \right] \right)_{s \geq 0} \right] = \mathcal{L} [(0)_{s \geq 0}]$$

in the *Meyer-Zheng topology*.

Combining (8.10) and (8.11), we obtain the multiple space-time scaling behaviour of the system in (7.14).

**Proposition 8.1. [Mean-field: finite-systems scheme]** *Suppose that the SSDE in (7.15) has initial measure  $\mathcal{L}[Z^{[N]}(0)] = \mu^{\otimes [N]}$  for some  $\mu \in \mathcal{P}([0, 1]^2)$ . Let*

$$(8.12) \quad \theta = \mathbb{E}^\mu \left[ \frac{x + Ky_0}{1 + K} \right].$$

(a) *For the averages in (8.8),*

$$(8.13) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( x_1^{[N]}(s), y_{0,1}^{[N]}(s) \right)_{s>0} \right] = \mathcal{L} \left[ \left( x_1^{\text{N}_0}(s), y_{0,1}^{\text{N}_0}(s) \right)_{s>0} \right]$$

*in the Meyer-Zheng topology,*

*where the limit process is the unique solution of the SSDE*

$$(8.14) \quad \begin{aligned} dx_1^{\text{N}_0}(s) &= \frac{1}{1+K} \sqrt{(\mathcal{F}g)(x_1^{\text{N}_0}(s))} dw(s), \\ y_{0,1}^{\text{N}_0}(s) &= x_1^{\text{N}_0}(s), \end{aligned}$$

*with initial state*

$$(8.15) \quad \left( x_1^{\text{N}_0}(0), y_{0,1}^{\text{N}_0}(0) \right) = (\theta, \theta).$$

(b) *For the weighted sum of the averages in (8.10),*

$$(8.16) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \bar{\Theta}^{[N]}(Ns) \right)_{s>0} \right] = \mathcal{L} \left[ \left( \bar{\Theta}(s) \right)_{s>0} \right],$$

*where the limit is the macroscopic observable in (8.7) with initial state*

$$(8.17) \quad \bar{\Theta}(0) = \theta.$$

(c) *Define*

$$(8.18) \quad \nu_\theta(s) = \int_{[0,1]} Q_s(\theta, d\theta') \nu_{\theta'} \in \mathcal{P}([0, 1]^2),$$

*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  $s \in (0, \infty)$ ,*

$$(8.19) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( X^{[N]}(Ns+t), Y^{[N]}(Ns+t) \right)_{t>0} \right] = \mathcal{L} \left[ \left( Z^{\nu_\theta(s)}(t) \right)_{t>0} \right]$$

*where, conditional on  $\bar{\Theta}(s) = \theta$ ,  $(z^{\nu_\theta(s)}(t))_{t \geq 0}$  is the random process in (8.1) 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 8.1 is given in Section 8.2.

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 (8.10) the diffusion function  $g$  is replaced by a *renormalised diffusion function*  $\mathcal{F}g$ , defined by (8.5) (recall Fig. 12). In section 8.2 we will show that  $\mathcal{F}\mathcal{G} \subset \mathcal{G}$ , i.e.,  $\mathcal{F}$  preserves the class of diffusion functions defined in (2.15).
- 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 (8.9)). 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 (7.15) cancel out against each other.
- Comparing the system in (7.14) with the system of interacting Fisher-Wright diffusions in the mean-field limit studied in [DG93b], we see from (8.14) 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 (8.10) follows an autonomous SDE, with convergence in the classical topology, i.e., in  $C_b([0, \infty), [0, 1])$ . The Brownian motion in (7.1) 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  $N \rightarrow \infty$ , 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 8.2. [Basic multi-scale]** Note that Proposition 8.1 already reveals several phenomena that we encountered in Theorems 4.2 and 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 8.3. [Interchange of limits]** The notation  $x_1^{N_0}, y_{0,1}^{N_0}$  indicates that the limit arises from taking averages over  $[N]$  and letting  $N \rightarrow \infty$ . Note that, for i.i.d. initial states,

$$(8.20) \quad x_1^{N_0}(0) = \lim_{N \rightarrow \infty} x_1^{[N]}(0) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} x_i(0) = \theta \quad \mathbb{P}\text{-a.s.}$$

On the other hand, picking any sequence of times  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , we get

$$(8.21) \quad x_1^{N_0}(0+) = \lim_{N \rightarrow \infty} x_1^{[N]}(\frac{L(N)}{N}t) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} x_i(L(N)t) = \theta \quad \mathbb{P}\text{-a.s.}$$

The mismatch between (8.20) and (8.21) indicates that we must be careful with interchanging the limits  $N \rightarrow \infty$  and  $s \downarrow 0$ . This is why (8.13), which lives on the fast time scale, is restricted to  $s > 0$ . ■

## 8.2 Abstract scheme behind finite-systems scheme

To prove Proposition 8.1, we follow the abstract scheme outlined in [DGV95, p. 2314–2315] and based on [DG93b], [DG93a]. 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 9.

**Step 1. Equilibrium of the single components.** This step fixes the one-dimensional distributions of the single components when  $t, N \rightarrow \infty$  in a combined way, and is the equivalent of [DG93b, Proposition 1]. Recall that  $\bar{\Theta}^{[N]}$  is defined in (8.2).

**Proposition 8.4. [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  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , and suppose that*

$$(8.22) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \bar{\Theta}^{[N_k]}(N_k s) \right] = P_s, \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \sup_{0 \leq t \leq L(N_k)} \left| \bar{\Theta}^{[N_k]}(N_k s) - \bar{\Theta}^{[N_k]}(N_k s - t) \right| \right] = \delta_0, \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left( X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s) \right) = \nu(s). \end{aligned}$$



Then  $\nu(s)$  is of the form

$$(8.23) \quad \nu(s) = \int_{[0,1]} P_s(d\theta) \nu_\theta,$$

where  $\nu_\theta$  is the equilibrium measure of the process defined in (8.1).

Proposition 8.4 follows from the following seven lemmas, which are the analogues of the five lemmas used in [DG93b, p. 477–478] for the system without seed-bank.

The first lemma establishes convergence of the infinite system in (8.1) to its equilibrium.

**Lemma 8.5. [Convergence for the infinite system]** *Let  $\mu$  be an exchangeable probability measure on  $([0, 1]^2)^{\mathbb{N}_0}$ . Then for the system  $(Z(t))_{t \geq 0}$  given by (8.1) with  $\mathcal{L}(Z(0)) = \mu$ ,*

$$(8.24) \quad \lim_{t \rightarrow \infty} \mathcal{L}[Z(t)] = \nu_\theta,$$

where  $\nu_\theta$  is of the form

$$(8.25) \quad \nu_\theta = \Gamma_\theta^{\otimes \mathbb{N}_0},$$

with  $\Gamma_\theta$  the equilibrium of the single-colony process defined in (8.4). Moreover,  $\nu_\theta$  is ergodic.

The second lemma establishes the continuity of the equilibrium with respect to its center of drift  $\theta$ .

**Lemma 8.6. [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  $[0, 1] \rightarrow \mathcal{P}([0, 1]^{\mathbb{N}_0})$  given by*

$$(8.26) \quad \theta \mapsto \nu_\theta$$

is continuous. Furthermore, if  $h$  is a Lipschitz function on  $[0, 1]$ , then also  $\mathcal{F}h$  defined by

$$(8.27) \quad (\mathcal{F}h)(\theta) = \mathbb{E}^{\nu_\theta} [h(\cdot)] = \int_{([0,1]^2)^{\mathbb{N}_0}} \nu_\theta(dz) h(x_0)$$

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.

**Lemma 8.7. [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 (8.3), and define*

$$(8.28) \quad \Delta_{\bar{\Theta}}^{[N]}(t) = \Theta_x^{[N]}(t) - \bar{\Theta}_y^{[N]}(t).$$

Then

$$(8.29) \quad \mathbb{E} \left[ \left| \Delta_{\bar{\Theta}}^{[N]}(t) \right| \right] \leq \sqrt{\mathbb{E} \left[ \left( \Delta_{\bar{\Theta}}^{[N]}(0) \right)^2 \right]} e^{-(Ke+e)t} + \sqrt{\frac{\|g\|}{N(Ke+e)}}.$$

**Remark 8.8. [Key estimate for Meyer-Zheng convergence]** The estimate in (8.29) in Lemma 8.7 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 (8.29) shows that  $\mathbb{E} \left[ \left| \Delta_{\bar{\Theta}}^{[N]}(Ns) \right| \right]$  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  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , and define the distribution  $\mu_N$  by continuing the configuration of  $(X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N)))$  periodically to  $([0, 1]^2)^{\mathbb{N}_0}$ . Define

$$(8.30) \quad \bar{\Theta}^{[N]} = \bar{\Theta}^{[N]}(Ns - L(N)).$$

Note that

$$(8.31) \quad \begin{aligned} \bar{\Theta}^{[N]} &= \frac{\frac{1}{N} \sum_{j \in [N]} x_j^{[N]}(Ns - L(N)) + \frac{K}{N} \sum_{j \in [N]} y_j^{[N]}(Ns - L(N))}{1 + K} \\ &= \frac{1}{N} \sum_{j \in [N]} \frac{x_j^{\mu_N}(0) + Ky_j^{\mu_N}(0)}{1 + K}. \end{aligned}$$

Thus,  $\bar{\Theta}^{[N]}$  is a random variable whose law depends on  $\mathcal{L}[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))]$   $= \mu_N$ . The infinite system with initial law  $\mu_N$  is denoted by

$$(8.32) \quad (X^{\mu_N}(t), Y^{\mu_N}(t))_{t \geq 0} = (x_i^{\mu_N}(t), y_i^{\mu_N}(t))_{i \in \mathbb{N}_0, t \geq 0}$$

and evolves according to

$$(8.33) \quad \begin{aligned} dx_i^{\mu_N}(t) &= c[\bar{\Theta}^{[N]} - x_i^{\mu_N}(t)] dt + \sqrt{g(x_i^{\mu_N}(t))} dw_i(t) + Ke[y_i^{\mu_N}(t) - x_i^{\mu_N}(t)] dt, \\ dy_i^{\mu_N}(t) &= e[x_i^{\mu_N}(t) - y_i^{\mu_N}(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

where  $\{w_i\}_{i \in \mathbb{N}_0}$  is a collection of independent Brownian motions.

**Lemma 8.9. [Comparison of finite and infinite systems]** Fix  $s > 0$  and assume that, for any  $L(N)$  satisfying  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ ,

$$(8.34) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} |\bar{\Theta}^{[N]}(Ns) - \bar{\Theta}^{[N]}(Ns - t)| = 0 \text{ in probability.}$$

Let

$$(8.35) \quad (X^{\mu_N}(t), Y^{\mu_N}(t))_{t \geq 0}$$

be the infinite system defined in (8.33) starting in the distribution  $\mu_N$ , where  $\mu_N$  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  $t \geq 0$ ,

$$(8.36) \quad \begin{aligned} \lim_{N \rightarrow \infty} |\mathbb{E}[f(X^{\mu_N}(t), Y^{\mu_N}(t)) - f(X^{[N]}(Ns - L(N) + t), Y^{[N]}(Ns - L(N) + t))]| &= 0 \\ \forall f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}. \end{aligned}$$

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  $N \rightarrow \infty$ , we need the following regularity property for the estimator  $\bar{\Theta}^{[N]}$ . This is stated in our fifth lemma.

**Lemma 8.10. [Stability of the estimator for the conserved quantity]** Define  $\mu_N$  as in Lemma 8.9. Let  $(x_i, y_i)_{i \in [N]}$  be distributed according to the exchangeable probability measure  $\mu_N$  on  $([0, 1]^2)^{\mathbb{N}_0}$  restricted to  $([0, 1]^2)^{[N]}$ . Suppose that  $\lim_{N \rightarrow \infty} \mu_N = \mu$  for some exchangeable probability measure  $\mu$  on  $([0, 1]^2)^{\mathbb{N}_0}$ . Define a random variable  $\phi$  on  $(\mu, ([0, 1]^2)^{\mathbb{N}_0})$  by putting

$$(8.37) \quad \phi = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i + Ky_i}{1 + K},$$

and a random variable  $\phi_N$  on  $(\mu_N, ([0, 1]^2)^{\mathbb{N}_0})$  by putting

$$(8.38) \quad \phi_N = \frac{1}{N} \sum_{i \in [N]} \frac{x_i + Ky_i}{1 + K}.$$

Then

$$(8.39) \quad \lim_{N \rightarrow \infty} \mathcal{L}[\phi_N] = \mathcal{L}[\phi].$$

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  $N \rightarrow \infty$ .

**Lemma 8.11. [Uniformity of the ergodic theorem for the infinite system]** *Let  $\mu_N$  be defined as in Lemma 8.9. 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  $\mu = \lim_{k \rightarrow \infty} \mu_{N_k}$ . Define*

$$(8.40) \quad \bar{\Theta} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} \frac{x_i^\mu + K y_i^\mu}{1 + K} \quad \text{in } L_2(\mu),$$

and let  $(X^\mu(t), Y^\mu(t))_{t \geq 0}$  be the infinite system evolving according to

$$(8.41) \quad \begin{aligned} dx_i^\mu(t) &= c [\bar{\Theta} - x_i^\mu(t)] dt + \sqrt{g(x_i^\mu(t))} dw_i(t) + K e [y_i^\mu(t) - x_i^\mu(t)] dt, \\ dy_i^\mu(t) &= e [x_i^\mu(t) - y_i^\mu(t)] dt, \quad i \in \mathbb{N}_0. \end{aligned}$$

Then

1. For all  $t \geq 0$ ,

$$(8.42) \quad \begin{aligned} \lim_{k \rightarrow \infty} |\mathbb{E}[f(X^{\mu_{N_k}}(t), Y^{\mu_{N_k}}(t))] - \mathbb{E}[f(X^\mu(t), Y^\mu(t))]| &= 0, \\ \forall f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}. \end{aligned}$$

2. There exists a sequence  $(\bar{L}(N))_{N \in \mathbb{N}}$  satisfying  $\lim_{N \rightarrow \infty} \bar{L}(N) = \infty$  and  $\lim_{N \rightarrow \infty} \bar{L}(N)/N = 0$  such that

$$(8.43) \quad \begin{aligned} \lim_{k \rightarrow \infty} |\mathbb{E}[f(X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))) \\ - f(X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k)))]| \\ + |\mathbb{E}[f(X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k)))] - \mathbb{E}[f(X^\mu(\bar{L}(N_k)), Y^\mu(\bar{L}(N_k)))]| &= 0 \\ \forall f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}. \end{aligned}$$

**Remark 8.12. [Existence of  $\bar{\Theta}$ ]** Note that the limit in (8.40) is well-defined by the ergodic theorem in  $L_2$ , since  $\mu$  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 8.13. [Coupling of finite systems]** *Let  $(X^{[N],1}, Y^{[N],1})$  be a finite system evolving according to (7.15) 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 (7.15) 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  $(\mu^{[N],2}, ([0, 1]^2)^{\mathbb{N}_0})$  and  $\bar{\Theta}_1$  and  $\bar{\Theta}_2$  on  $(\mu, ([0, 1]^2)^{\mathbb{N}_0})$  by*

$$(8.44) \quad \begin{aligned} \bar{\Theta}^{[N],1} &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N],1} + K y_i^{[N],1}}{1 + K}, & \bar{\Theta}^{[N],2} &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N],2} + K y_i^{[N],2}}{1 + K}, \\ \bar{\Theta}_1 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i^1 + K y_i^1}{1 + K}, & \bar{\Theta}_2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i^2 + K y_i^2}{1 + K}, \end{aligned}$$

and let  $(\bar{\Theta}^{[N],1}(t))_{t \geq 0}$  and  $(\bar{\Theta}^{[N],2}(t))_{t \geq 0}$  be defined according to (8.2) for  $(X_1^{[N]}, Y_1^{[N]})$ , respectively,  $(X_2^{[N]}, Y_2^{[N]})$ . Assume that

$$(8.45) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} |\bar{\Theta}^{[N],k}(0) - \bar{\Theta}^{[N],k}(t)| = 0 \quad \text{in probability, } k \in \{1, 2\},$$

and suppose that  $\tilde{\mu}(\{\bar{\Theta}_1 = \bar{\Theta}_2\}) = 1$ . Then, for any sequence  $t(N) \rightarrow \infty$ ,

$$(8.46) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ |x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K |y_i^{[N],1}(t(N)) - y_i^{[N],2}(t(N))| \right] = 0.$$

**Step 2. Convergence of the estimator.** This step is the equivalent of [DG93b, 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 8.14. [Convergence of average sum process]**

$$(8.47) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \bar{\Theta}^{[N]}(Ns) \right)_{s>0} \right] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}],$$

where  $(\bar{\Theta}(s))_{s>0}$  evolves according to

$$(8.48) \quad d\bar{\Theta}(s) = \frac{1}{(1+K)} \sqrt{(\mathcal{F}g)(\bar{\Theta}(s))} dw(s).$$

Proposition 8.14 follows from the following three lemmas, which are the equivalent of the three lemmas used in [DG93b, p. 488–493] for the system without seed-bank.

**Lemma 8.15. [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*

$$(8.49) \quad \left\langle \bar{\Theta}^{[N]}(Ns) \right\rangle_{s>0} = \frac{1}{(1+K)^2} \int_0^s dr \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Nr)).$$

- (2) *Let  $L(N)$  be such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . Then*

$$(8.50) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} |\bar{\Theta}^{[N]}(Ns) - \bar{\Theta}^{[N]}(Ns-t)| = 0 \text{ in probability.}$$

- (3)  *$(\mathcal{L}[(\bar{\Theta}^{[N]}(Ns))_{s>0}])_{N \in \mathbb{N}}$  is tight as a sequence of probability measures on  $\mathcal{C}([0, \infty), [0, 1])$ .*

**Lemma 8.16. [Martingale property of limit process]** *Let  $(N_k)_{k \in \mathbb{N}}$  be any subsequence such that*

$$(8.51) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \bar{\Theta}^{[N_k]}(N_k s) \right)_{s>0} \right] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}].$$

*Then  $(\bar{\Theta}(s))_{s>0}$  is a square-integrable martingale with continuous paths, and*

$$(8.52) \quad \left( \bar{\Theta}^2(s) - \int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)] \right)_{s>0}$$

*is a martingale.*

**Lemma 8.17. [Uniqueness]** *The following martingale problem has a unique solution:*

$$(8.53) \quad \left( \bar{\Theta}_s \right)_{s>0} \text{ is a continuous martingale with values in } [0, 1],$$

$$\left( \bar{\Theta}^2(s) - \frac{1}{(1+K)^2} \int_0^s dr \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)] \right)_{s>0} \text{ is a martingale.}$$

*The solution of (8.53) 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 4.1. We have to prove the following proposition.

**Proposition 8.18. [Convergence in Meyer-Zheng topology]** *If*

$$(8.54) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \bar{\Theta}(Ns) \right)_{s>0} \right] = \mathcal{L} \left[ \left( \bar{\Theta}(s) \right)_{s>0} \right],$$

*then*

$$(8.55) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( x_1^{[N]}(t), y_1^{[N]}(t) \right)_{t \geq 0} \right] = \mathcal{L} \left[ \left( x_1^{\mathbb{N}_0}(t), y_1^{\mathbb{N}_0}(t) \right)_{t \geq 0} \right]$$

*in the Meyer-Zheng topology,*

*where  $(x_1^{\mathbb{N}_0}(t), y_1^{\mathbb{N}_0}(t))_{t \geq 0}$  evolves according to (8.14).*

To prove Proposition 8.18 we will use Lemma 8.7 in combination with the following three general lemmas about the Meyer-Zheng topology, which are proven in Appendix B.3.

**Lemma 8.19. [Path 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  $t \geq 0$ ,*

$$(8.56) \quad \lim_{n \rightarrow \infty} \mathbb{E}[d(Z_n(t), Z(t))] = 0,$$

then,

$$(8.57) \quad \lim_{n \rightarrow \infty} (Z_n(t))_{t \geq 0} = (Z(t))_{t \geq 0} \text{ in probability in the Meyer-Zheng topology.}$$

**Lemma 8.20. [Convergence of the joint law of paths in the Meyer-Zheng topology]** *Let  $(X_n)_{n \in \mathbb{N}}, (Y_n)_{n \in \mathbb{N}}, X$  be stochastic processes on a metric space  $(E, d)$  and let  $c \in E$  be a constant. If  $\lim_{n \rightarrow \infty} \mathcal{L}[X_n] = \mathcal{L}[X]$  in the Meyer-Zheng topology and for all  $t \geq 0$ ,  $\lim_{n \rightarrow \infty} \mathbb{E}[d(Y_n(t), c)] = 0$ , then  $\lim_{n \rightarrow \infty} \mathcal{L}[(X_n, Y_n)] = \mathcal{L}[(X, c)]$  in the Meyer-Zheng topology.*

**Lemma 8.21. [Continuous mapping theorem in the Meyer-Zheng topology]** *Let  $f: E \rightarrow E$  be a continuous function and  $x \in M_E[0, \infty)$ .*

(a) *The function*

$$(8.58) \quad h: \Psi \rightarrow \Psi, \quad \psi_x \rightarrow \psi_{f(x)},$$

*is continuous.*

(b) *If the stochastic processes  $(X_n)_{n \in \mathbb{N}}, X$  on state space  $(E, d)$  satisfy*

$$(8.59) \quad \lim_{n \rightarrow \infty} \mathcal{L}[X_n] = \mathcal{L}[X] \text{ in the Meyer-Zheng topology,}$$

then

$$(8.60) \quad \lim_{n \rightarrow \infty} \mathcal{L}[f(X_n)] = \mathcal{L}[f(X)] \text{ in the Meyer-Zheng topology.}$$

Note that Lemma 8.21 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 8.1.

Having completed the abstract scheme of steps 1–4, we set out to prove the constituent propositions and lemmas.

## 9 Proofs: $N \rightarrow \infty$ , mean-field, proof of abstract scheme

In Sections 9.1–9.4 we prove the propositions and the lemmas stated in Steps 1–4 in Section 8.2.

### 9.1 Proof of step 1. Equilibrium of the single components

We start by proving Proposition 8.4 with the help of the seven lemmas stated in Step 1 of Section 8.2. Afterwards we prove each of the lemmas.

#### • Proof of Proposition 8.4

*Proof.* We use an argument similar to the one used in [DG93b, Section 2 (i)]. Let  $(L(N))_{N \in \mathbb{N}}$  be any sequence satisfying  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . Let  $\mu_N$  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 8.4, we can pass to a further subsequence and obtain

$$(9.1) \quad \lim_{k \rightarrow \infty} \mu_{N_k} = \mu.$$

Since we assumed that  $\mathcal{L}[X^{[N]}(0), Y^{[N]}(0)]$  is exchangeable and the dynamics preserves exchangeability, the measures  $\mu_{N_k}$  are exchangeable and also the limiting law  $\mu$  is exchangeable. Define  $\phi$  as in (8.37) in Lemma 8.10. Then we can condition on  $\phi$  and write

$$(9.2) \quad \mu = \int_{[0,1]} \mu_\rho \, d\Lambda(\rho),$$

where  $\Lambda(\cdot) = \mathcal{L}[\phi]$ . By assumption we know that

$$(9.3) \quad \lim_{k \rightarrow \infty} \mathcal{L}[\bar{\Theta}^{[N_k]}(N_k s)] = P_s$$

and

$$(9.4) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \sup_{0 \leq t \leq L(N_k)} \left| \bar{\Theta}^{[N_k]}(N_k s) - \bar{\Theta}^{[N_k]}(N_k s - t) \right| \right] = \delta_0.$$

Hence

$$(9.5) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \bar{\Theta}^{[N_k]}(N_k s - L(N_k)) \right] = P_s.$$

Recall that

$$(9.6) \quad \Lambda = \mathcal{L}[\phi] = \mathcal{L} \left[ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i + K y_i}{1 + K} \right] \quad \text{on } (\mu, ([0, 1]^2)^{\mathbb{N}_0}).$$

By Lemma 8.10, if  $\phi_{N_k} = \frac{1}{N_k} \sum_{i \in [N_k]} \frac{x_i + K y_i}{1 + K}$  on  $(\mu_{N_k}, ([0, 1]^2)^{\mathbb{N}_0})$ , then  $\lim_{k \rightarrow \infty} \mathcal{L}[\phi_{N_k}] = \mathcal{L}[\phi]$ . Taking the subsequence  $(\mu_{N_k})_{k \in \mathbb{N}}$ , we get  $\Lambda(\cdot) = P_s(\cdot)$ , and hence

$$(9.7) \quad \mu = \int_{[0,1]} \mu_\rho \, dP_s(\rho).$$

Let  $\bar{L}(N)$  be the sequence constructed in Lemma 8.11[b]. We can require that  $\bar{L}(N) \leq L(N)$  for all  $N \in \mathbb{N}$ . Write

$$(9.8) \quad \begin{aligned} & \mathcal{L}[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))] \\ &= \mathcal{L}[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{N_k}(N_k s - L(N_k) + \bar{L}(N_k))] \\ & - \mathcal{L}[X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k))], \\ & + \mathcal{L}[X^{\mu_{N_k}}(\bar{L}(N_k)), Y^{\mu_{N_k}}(\bar{L}(N_k))] - \mathcal{L}[X^\mu(\bar{L}(N_k)), Y^\mu(\bar{L}(N_k))] \\ & + \mathcal{L}[X^\mu(\bar{L}(N_k)), Y^\mu(\bar{L}(N_k))]. \end{aligned}$$

By Lemma 8.11, the first and the second term tend to zero as  $k \rightarrow \infty$ . Hence

$$(9.9) \quad \mathcal{L}[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))]$$

tends to  $\mathcal{L}[X^\mu(L(N_k)), Y^\mu(L(N_k))]$  as  $k \rightarrow \infty$ . By (9.7),

$$(9.10) \quad \mathcal{L}[X^\mu(\bar{L}(N_k)), Y^\mu(\bar{L}(N_k))] = \int_{[0,1]} \mathcal{L}[X^{\mu_\rho}(\bar{L}(N_k)), Y^{\mu_\rho}(\bar{L}(N_k))] \, dP_s(\rho).$$

Since  $\lim_{k \rightarrow \infty} \bar{L}(N_k) = \infty$ , by Lemma 8.5 we have

$$(9.11) \quad \lim_{k \rightarrow \infty} \mathcal{L}[X^{\mu_\rho}(\bar{L}(N_k)), Y^{\mu_\rho}(\bar{L}(N_k))] = \nu_\rho.$$

Therefore, by (9.8) and Lemma 8.6,

$$(9.12) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathcal{L}[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))] \\ &= \int_{[0,1]} \nu_\rho \, dP_s(\rho). \end{aligned}$$

To show that

$$(9.13) \quad \lim_{k \rightarrow \infty} \mathcal{L}[X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s)] = \int_{[0,1]} \nu_\rho \, dP_s(\rho).$$

we invoke Lemma 8.13. Let  $(X^{[N],1}, Y^{[N],1})$  be the finite system starting from

$$(9.14) \quad \mathcal{L}[X^{[N]}(N_s - L(N)), Y^{[N]}(N_s - L(N))],$$

and let  $(\bar{L}(N))_{N \in \mathbb{N}}$  be the sequence such that (9.12) holds. Let  $(X^{[N],2}, Y^{[N],2})$  be the finite system starting from

$$(9.15) \quad \mathcal{L}[X^{[N]}(N_s - \bar{L}(N)), Y^{[N]}(N_s - \bar{L}(N))].$$

Choose for the sequence  $t(N)$  in Lemma 8.13 the sequence  $\bar{L}(N)$ . Let  $\mu^{[N],1}$  be defined by periodic continuation of  $(X^{[N]}(N_s - L(N)), Y^{[N]}(N_s - L(N)))$ , and  $\mu^{[N],2}$  by periodic continuation of  $(X^{[N]}(N_s - \bar{L}(N)), Y^{[N]}(N_s - \bar{L}(N)))$ . Defining  $\bar{\Theta}_1$  and  $\bar{\Theta}_2$  according to (8.44), where for  $\mu^{[N],2}$  we replace  $L(N)$  by  $\bar{L}(N)$ , we get

$$(9.16) \quad \lim_{k \rightarrow \infty} |\bar{\Theta}_1^{N_k} - \bar{\Theta}_2^{N_k}| = \lim_{k \rightarrow \infty} |\bar{\Theta}^{N_k}(N_k s - L(N_k)) - \bar{\Theta}^{N_k}(N_k s - \bar{L}(N_k))| = 0 \text{ in probability}$$

by the assumptions in (8.22). Hence, if  $\mu$  is any weak limit point of the sequence  $(\mu^{[N_k],1} \times \mu^{[N_k],2})_{k \in \mathbb{N}}$ , then

$$(9.17) \quad \mu(\bar{\Theta}_1 = \bar{\Theta}_2) = 1.$$

By passing to a further subsequence, we can now apply Lemma 8.13, to obtain

$$(9.18) \quad \lim_{k \rightarrow \infty} \mathbb{E} \left[ |x_{i,1}^{N_k}(\bar{L}(N_k)) - x_{i,2}^{N_k}(\bar{L}(N_k))| + K |y_{i,1}^{N_k}(\bar{L}(N_k)) - y_{i,2}^{N_k}(\bar{L}(N_k))| \right] = 0.$$

Note that

$$(9.19) \quad \begin{aligned} & \mathcal{L}[X_1(\bar{L}(N_k)), Y_1(\bar{L}(N_k))] \\ &= \mathcal{L}[X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))], \\ & \mathcal{L}[X_2(\bar{L}(N_k)), Y_2(\bar{L}(N_k))] = \mathcal{L}[X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s)]. \end{aligned}$$

Moreover, we know from (9.12) that

$$(9.20) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ X^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)), Y^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k)) \right] = \int_{[0,1]} \nu_\rho P_s(d\rho).$$

Combining (9.18)–(9.20), we find that

$$(9.21) \quad \mathcal{L}[X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s)] = \int_{[0,1]} \nu_\rho P_s(d\rho).$$

□

In the remainder of this section we prove Lemmas 8.5–8.11 and 8.13.

### • Proof of Lemma 8.5

*Proof.* Since the components of the infinite system in (8.1) evolve independently, it is enough to show that each component converges to  $\Gamma_\theta$ . This convergence follows from the proof of Proposition 7.2 (see Section 7.3). Hence the infinite system defined by (7.18) converges to  $\nu_\theta = \Gamma_\theta^{\otimes \mathbb{N}_0}$ . Ergodicity of  $\nu_\theta$  with respect to translations follows from Kolmogorov's zero-one law. □

• **Proof of Lemma 8.7**

*Proof.* Using the definition of  $\Theta_x^{[N]}(t)$ ,  $\Theta_y^{[N]}(t)$  in (8.3) and the SSDE in (7.15), we find the following evolution for the averages:

$$(9.22) \quad \begin{aligned} d\Theta_x^{[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i(t))} dw_i(t) + Ke [\Theta_y^{[N]}(t) - \Theta_x^{[N]}(t)] dt, \\ d\Theta_y^{[N]}(t) &= e [\Theta_x^{[N]}(t) - \Theta_y^{[N]}(t)] dt. \end{aligned}$$

Consequently,

$$(9.23) \quad \begin{aligned} d(\Delta_{\Theta}^{[N]}(t))^2 &= 2\Delta_{\Theta}^{[N]}(t) d\Delta_{\Theta}^{[N]}(t) + 2d\langle \Delta_{\Theta}^{[N]} \rangle(t) \\ &= \Delta_{\Theta}^{[N]}(t) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i(t))} dw_i(t) - (Ke + e) (\Delta_{\Theta}^{[N]}(t))^2 dt \\ &\quad + 2\frac{1}{N^2} \sum_{i \in [N]} g(x_i(t)) dt, \end{aligned}$$

and hence

$$(9.24) \quad \frac{d}{dt} \mathbb{E} [(\Delta_{\Theta}^{[N]}(t))^2] = -2(Ke + e) \mathbb{E} [(\Delta_{\Theta}^{[N]}(t))^2] + \frac{2}{N^2} \sum_{i \in [N]} g(x_i(t))$$

and

$$(9.25) \quad \mathbb{E} [(\Delta_{\Theta}^{[N]}(t))^2] = \mathbb{E} [(\Delta_{\Theta}^{[N]}(0))^2] e^{-2(Ke+e)t} + \int_0^t dr e^{-2(Ke+e)(t-r)} \frac{2}{N^2} \sum_{i \in [N]} g(x_i(r)).$$

Therefore we get the bound

$$(9.26) \quad \mathbb{E} [|\Delta_{\Theta}^{[N]}(t)|] \leq \sqrt{\mathbb{E} [(\Delta_{\Theta}^{[N]}(0))^2]} e^{-(Ke+e)t} + \sqrt{\frac{2\|g\|}{N(Ke+e)}}.$$

□

• **Proof of Lemma 8.9**

*Proof.* To compare the systems in (7.15) and (8.33), we couple them via their Brownian motions. Therefore for all  $i \in [N]$  we assume that the evolution in (7.15) and (8.33) is driven by the same Brownian motion,  $\tilde{w}_i = w_i$ . If  $i \notin [N]$ , then we set  $w_i = w_j$  for  $j = i \bmod N$ . 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

$$(9.27) \quad \begin{aligned} \mathcal{L}[\tilde{z}(0)] &= \mathcal{L}[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))] \times \mu_N \\ &= \mathcal{L}[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))]^2. \end{aligned}$$

Define

$$(9.28) \quad \Delta_i^N(t) = \tilde{x}_i^{[N]}(t) - \tilde{x}_i^{\mu N}(t), \quad \delta_i^N(t) = \tilde{y}_i^{[N]}(t) - \tilde{y}_i^{\mu N}(t).$$

To prove that the coupling is successful, we show that, for all  $t \geq 0$ ,

$$(9.29) \quad \lim_{N \rightarrow \infty} \mathbb{E} [|\Delta_i^N(t)| + K|\delta_i^N(t)|] = 0 \quad \forall i \in \mathbb{N}_0.$$

From now on we will only consider sites  $i \in [0, N]$  for which both infinite systems have the same Brownian motion.



From (7.15) and (8.33) it follows that

$$\begin{aligned}
& d [|\Delta_i^N(t)| + K|\delta_i^N(t)|] \\
&= (\text{sign } \Delta_i^N(t)) d\Delta_i^N(t) + dL_t^0 + K \text{sign } \delta_i^N(t) d\delta_i^N(t) \\
&= -c (\text{sign } \Delta_i^N(t)) \Delta_i^N(t) dt + c (\text{sign } \Delta_i^N(t)) [\bar{\Theta}^{[N]}(t) - \bar{\Theta}^{[N]}] dt \\
(9.30) \quad &+ c (\text{sign } \Delta_i^N(t)) [\Theta_x^{[N]}(t) - \bar{\Theta}^{[N]}(t)] dt \\
&+ (\text{sign } \Delta_i^N(t)) \left( \sqrt{g(x_i^{[N]}(t))} - \sqrt{g(x_i^{\mu_N}(t))} \right) dw_i(t) \\
&+ (\text{sign } \Delta_i^N(t)) Ke [\delta_i^N(t) - \Delta_i^N(t)] dt \\
&+ (\text{sign } \delta_i^N(t)) Ke [\Delta_i^N(t) - \delta_i^N(t)] dt,
\end{aligned}$$

where we use that the local time  $L_t^0$  is zero, since  $g$  is Lipschitz (see [RW00, Proposition V.39.3]).

Taking expectations in (9.30), we find

$$\begin{aligned}
(9.31) \quad \frac{d}{dt} \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] &= -c \mathbb{E}[|\Delta_i^N(t)|] \\
&+ c \mathbb{E}[(\text{sign } \Delta_i^N(t)) [\bar{\Theta}^{[N]}(t) - \bar{\Theta}^{[N]}]] \\
&+ c \mathbb{E}[(\text{sign } \Delta_i^N(t)) [\Theta_x^{[N]}(t) - \bar{\Theta}^{[N]}(t)]] \\
&+ Ke \mathbb{E}[(\text{sign } \Delta_i^N(t) - \text{sign } \delta_i^N(t)) [\delta_i^N(t) - \Delta_i^N(t)]].
\end{aligned}$$

Note that we can rewrite (9.31) as

$$\begin{aligned}
(9.32) \quad \frac{d}{dt} \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] &= -c \mathbb{E}[|\Delta_i^N(t)|] \\
&- 2Ke \mathbb{E} \left[ \mathbb{1}_{\text{sign } \Delta_i^N(t) \neq \text{sign } \delta_i^N(t)} [|\delta_i^N(t)| + |\Delta_i^N(t)|] \right] \\
&+ c \mathbb{E}[(\text{sign } \Delta_i^N(t)) [\bar{\Theta}^{[N]}(t) - \bar{\Theta}^{[N]}]] \\
&+ c \mathbb{E}[(\text{sign } \Delta_i^N(t)) [\Theta_x^{[N]}(t) - \bar{\Theta}^{[N]}(t)]].
\end{aligned}$$

It therefore follows that

$$\begin{aligned}
(9.33) \quad \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] &= \mathbb{E}[|\Delta_i^N(0)| + K|\delta_i^N(0)|] \\
&- c \int_0^t dr \mathbb{E}[|\Delta_i^N(r)|] \\
&- 2Ke \int_0^t dr \mathbb{E} \left[ \mathbb{1}_{\text{sign } \Delta_i^N(r) \neq \text{sign } \delta_i^N(r)} [|\delta_i^N(r)| + |\Delta_i^N(r)|] \right] \\
&+ \int_0^t dr c \mathbb{E}[(\text{sign } \Delta_i^N(r)) [\bar{\Theta}^{[N]}(r) - \bar{\Theta}^{[N]}]] \\
&+ \int_0^t dr c \mathbb{E}[(\text{sign } \Delta_i^N(r)) [\Theta_x^{[N]}(r) - \bar{\Theta}^{[N]}(r)]].
\end{aligned}$$

Note that, by the choice of initial distribution for the coupling, we have

$$(9.34) \quad \mathbb{E}[|\Delta_i^N(0)| + K|\delta_i^N(0)|] = 0.$$

Therefore we get

$$\begin{aligned}
(9.35) \quad 0 &\leq \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] \\
&\leq -c \int_0^t dr \mathbb{E}[|\Delta_i^N(r)|] \\
&\quad - 2Ke \int_0^t dr \mathbb{E} \left[ \mathbf{1}_{\text{sign } \Delta_i^N(r) \neq \text{sign } \delta_i^N(r)} [|\delta_i^N(r)| + |\Delta_i^N(r)|] \right] \\
&\quad + \int_0^t dr c \mathbb{E} [|\bar{\Theta}^{[N]}(r) - \bar{\Theta}^{[N]}|] \\
&\quad + \int_0^t dr c \mathbb{E} [|\Theta_x^{[N]}(r) - \bar{\Theta}^{[N]}(r)|] \\
&\leq t \left( \sup_{0 \leq r \leq t} c \mathbb{E} [|\bar{\Theta}^{[N]}(r) - \bar{\Theta}^{[N]}|] + c \mathbb{E} [|\Theta_x^{[N]}(r) - \bar{\Theta}^{[N]}(r)|] \right).
\end{aligned}$$

Hence, by the assumption in (8.34) and Lemma 8.7 (recall (9.27)), we see that, for all  $t > 0$ ,

$$(9.36) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] = 0.$$

Therefore, for every Lipschitz function  $f \in \mathcal{C}([0, 1], \mathbb{R})$  of  $x_i(t)$ ,

$$(9.37) \quad \lim_{n \rightarrow \infty} \left| \mathbb{E}[f(x_i^{[N]}(t)) - f(x_i^{\mu_N}(t))] \right| \leq \lim_{n \rightarrow \infty} \text{Lip} f \mathbb{E}[|\Delta_i^N(L(N))|] = 0,$$

and the same holds for Lipschitz functions of  $y_i$ . Using that the Lipschitz functions are dense in  $\mathcal{C}([0, 1], \mathbb{R})$ , we obtain that the result actually holds for all  $f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}$  depending on finitely many components. This in turn implies that the result holds for all  $f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}$ .  $\square$

### • Proof of Lemma 8.10

*Proof.* Define

$$(9.38) \quad D^N(Z) = \frac{1}{N} \sum_{j \in [N]} \frac{x_j + Ky_j}{1+K}, \quad D(Z) = \lim_{N \rightarrow \infty} D^N(Z) \text{ in } L_2(\mu).$$

Since  $\mu$  is translation invariant with  $\int_{[0,1]^2} \frac{x_0 + Ky_0}{1+K} d\mu < 1$ , the  $L_2(\mu)$ -limit  $D(Z)$  exists by the ergodic theorem. Since, by assumption,  $\mu_N \rightarrow \mu$  as  $N \rightarrow \infty$  for all fixed  $M \in \mathbb{N}_0$ , we have

$$(9.39) \quad \lim_{N \rightarrow \infty} \mathcal{L}_{\mu_N}[D^M(Z)] = \mathcal{L}_\mu[D^M(Z)].$$

Therefore, in order to prove Lemma 8.10, we are left to show

$$(9.40) \quad \lim_{M \rightarrow \infty} \sup_{N \geq M} \|D^M(Z) - D^N(Z)\|_{L_2(\mu_N)} = 0.$$

This can be done by using Fourier transforms and spectral densities, and to do so we follow the same strategy as in [DG93b, Lemma 2.5].

Define

$$(9.41) \quad \bar{\theta}^N = \mathbb{E}^{\mu_N} \left[ \frac{x_0 + Ky_0}{1+K} \right].$$

Since  $\mu_N$  is translation invariant on  $\mathbb{N}_0$ , by Herglotz's theorem there exists a unique measure  $\lambda_N$  such that, for all  $j, k \in \mathbb{N}_0$ ,

$$(9.42) \quad \mathbb{E}^{\mu_N} \left[ \left( \frac{x_j + Ky_j}{1+K} - \bar{\theta}^N \right) \left( \frac{x_k + Ky_k}{1+K} - \bar{\theta}^N \right) \right] = \int_{(-\pi, \pi]} \lambda_N(du) e^{i(j-k)u}.$$

For  $N \in \mathbb{N}_0$ , define

$$(9.43) \quad D^N(u) = \frac{1}{N} \sum_{j \in [N]} e^{ij u}.$$

By (9.42), it follows that

$$(9.44) \quad \|D^M(Z) - D^N(Z)\|_{L_2(\mu_N)} = \|D^M(u) - D^N(u)\|_{L_2(\lambda_N)}.$$

Polynomials of the type  $D^N(u)$  are called trigonometric polynomials and satisfy:

1.  $\lim_{N \rightarrow \infty} D^N(u) = 1_{\{0\}}(u)$ .
2. For  $\delta > 0$  and  $M < \infty$  there exists an  $\epsilon(M, \delta)$  such that, for all  $N \geq M$ ,

$$(9.45) \quad |D^N(u) - 1_{\{0\}}(u)| \leq 1_{(-\delta, \delta) \setminus \{0\}} + \epsilon(M, \delta) \text{ with } \epsilon(M, \delta) \rightarrow 0 \text{ as } M \rightarrow \infty.$$

Hence it follows that

$$(9.46) \quad \|D^M(u) - D^N(u)\|_{L_2(\lambda_N)}^2 \leq 2\lambda_N((-\delta, \delta) \setminus \{0\}) + 2\epsilon(M, \delta).$$

Now let  $M \rightarrow \infty$ , to obtain

$$(9.47) \quad \sup_{N \geq M} \|D^M(u) - D^N(u)\|_{L_2(\lambda_N)} \leq 2\lambda_N((-\delta, \delta) \setminus \{0\}).$$

Subsequently let  $\delta \rightarrow 0$ , so that  $(-\delta, \delta) \setminus \{0\} \rightarrow \emptyset$  and

$$(9.48) \quad \lim_{M \rightarrow \infty} \sup_{N \geq M} \|D^M(u) - D^N(u)\|_{L_2(\lambda_N)}^2 = 0.$$

□

### • Proof of Lemma 8.11

*Proof.* We first prove Lemma 8.11[1]. Afterwards we construct  $(\bar{L}(N))_{N \in \mathbb{N}}$  to prove Lemma 8.11[2].

Since  $\lim_{k \rightarrow \infty} \mu_{N_k} = \mu$ , Lemma 8.10 implies that  $\lim_{k \rightarrow \infty} \mathcal{L}[\bar{\Theta}^{[N_k]}] = \mathcal{L}[\bar{\Theta}]$ . For ease of notation we drop the subsequence notation in the remainder of this proof. By Skorohod's theorem we can construct the random variables  $(z_i^{\mu_N})_{N \in \mathbb{N}}$  and  $z_i^\mu$  on one probability space such that  $\lim_{N \rightarrow \infty} z_i^{\mu_N} = z_i^\mu$  a.s. Then, as in the proof of Lemma 8.10, we obtain

$$(9.49) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\bar{\Theta}^{[N]} - \bar{\Theta}|] = 0.$$

To prove the claim we couple the two infinite systems, namely,

$$(9.50) \quad \begin{aligned} dx_i^{\mu_N}(t) &= c [\bar{\Theta}^{[N]} - x_i^{\mu_N}(t)] dt + \sqrt{g(x_i^{\mu_N}(t))} dw_i(t) + Ke [y_i^{\mu_N}(t) - x_i^{\mu_N}(t)] dt, \\ dy_i^{\mu_N}(t) &= e [x_i^{\mu_N}(t) - y_i^{\mu_N}(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

and

$$(9.51) \quad \begin{aligned} dx_i^\mu(t) &= c [\bar{\Theta} - x_i^\mu(t)] dt + \sqrt{g(x_i^\mu(t))} dw_i(t) + Ke [y_i^\mu(t) - x_i^\mu(t)] dt, \\ dy_i^\mu(t) &= e [x_i^\mu(t) - y_i^\mu(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

are coupled by using the same Brownian motions in (9.50) and (9.51). 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 Skorohod's theorem, we have that

$$(9.52) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^{\mu_N}(0)| + K|\delta_i^{\mu_N}(0)|] = 0.$$

To prove that, for all  $t \geq 0$ ,

$$(9.53) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|] = 0,$$

we proceed as in the proof of Lemma 8.9. By Itô-calculus, we find that

$$\begin{aligned}
(9.54) \quad \mathbb{E}[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|] &= \mathbb{E}[|\Delta_i^{\mu_N}(0)| + K|\delta_i^{\mu_N}(0)|] \\
&\quad - c \int_0^t dr \mathbb{E}[|\Delta_i^{\mu_N}(r)|] \\
&\quad - 2Ke \int_0^t dr \mathbb{E}\left[1_{\text{sign } \Delta_i^{\mu_N}(r) \neq \text{sign } \delta_i^{\mu_N}(r)} [|\delta_i^{\mu_N}(r)| + |\Delta_i^{\mu_N}(r)|]\right] \\
&\quad + \int_0^t dr c \mathbb{E}\left[|\bar{\Theta}^{[N]} - \bar{\Theta}|\right] \\
&\leq \mathbb{E}[|\Delta_i^{\mu_N}(0)| + K|\delta_i^{\mu_N}(0)|] + tc \mathbb{E}[|\bar{\Theta}^{[N]} - \bar{\Theta}|].
\end{aligned}$$

From (9.54) it follows that, for every  $t \geq 0$ ,

$$(9.55) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|] = 0.$$

We next construct the sequence  $(\bar{L}(N))_{N \in \mathbb{N}}$ . From (9.36) and (9.55), we have

$$(9.56) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^N(t)| + K|\delta_i^N(t)|] + \mathbb{E}[|\Delta_i^{\mu_N}(t)| + K|\delta_i^{\mu_N}(t)|] = 0.$$

Let  $(t_k)_{k \in \mathbb{N}}$  be an increasing sequence such that  $\lim_{k \rightarrow \infty} t_k = \infty$  and  $\lim_{k \rightarrow \infty} t_k/k = 0$ . By (9.56), we can for each  $k$  find an  $N_k \in \mathbb{N}$  such that, for all  $N \geq N_k$ ,

$$(9.57) \quad \mathbb{E}[|\Delta_i^N(t_k)| + K|\delta_i^N(t_k)|] + \mathbb{E}[|\Delta_i^{\mu_N}(t_k)| + K|\delta_i^{\mu_N}(t_k)|] < \frac{1}{k}.$$

Requiring that  $N_{k+1} > N_k$ , we obtain a strictly increasing sequence  $(N_k)_{k \in \mathbb{N}}$  that partitions  $\mathbb{N}$ . Set

$$(9.58) \quad \bar{L}(N) = \sum_{k \in \mathbb{N}} t_k 1_{\{N_k, \dots, N_{k+1}-1\}}(N).$$

We show that  $\bar{L}(N)$  satisfies the required properties:

- $\lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^N(\bar{L}(N))| + K|\delta_i^N(\bar{L}(N))|] + \mathbb{E}[|\Delta_i^{\mu_N}(\bar{L}(N))| + K|\delta_i^{\mu_N}(\bar{L}(N))|] = 0$ . To proof this, we fix  $\epsilon > 0$  and let  $K$  be such that  $\frac{1}{K} < \epsilon$ . Then, for all  $N \geq N_K$ ,

$$\begin{aligned}
(9.59) \quad &\mathbb{E}[|\Delta_i^N(\bar{L}(N))| + K|\delta_i^N(\bar{L}(N))|] + \mathbb{E}[|\Delta_i^{\mu_N}(\bar{L}(N))| + K|\delta_i^{\mu_N}(\bar{L}(N))|] \\
&= \sum_{k \in \mathbb{N}} \mathbb{E}[|\Delta_i^N(t_k)| + K|\delta_i^N(t_k)|] + \mathbb{E}[|\Delta_i^{\mu_N}(t_k)| + K|\delta_i^{\mu_N}(t_k)|] 1_{\{N_k, \dots, N_{k+1}-1\}}(N) \\
&< \frac{1}{K} < \epsilon.
\end{aligned}$$

We conclude that

$$(9.60) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^N(\bar{L}(N))| + K|\delta_i^N(\bar{L}(N))|] + \mathbb{E}[|\Delta_i^{\mu_N}(\bar{L}(N))| + K|\delta_i^{\mu_N}(\bar{L}(N))|] = 0.$$

- $\lim_{N \rightarrow \infty} \bar{L}(N) = \infty$ . By (9.58), 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  $t_k \rightarrow \infty$ . We conclude that  $\lim_{N \rightarrow \infty} \bar{L}(N) = \infty$ .
- $\lim_{N \rightarrow \infty} \bar{L}(N)/N = 0$ . Recall that  $\lim_{k \rightarrow \infty} t_k/k = 0$  and  $N_k \geq k$  by construction. Hence  $\lim_{N \rightarrow \infty} \bar{L}(N)/N \leq \lim_{N \rightarrow \infty} \sum_{k \in \mathbb{N}} (t_k/k) 1_{\{N_k, \dots, N_{k+1}-1\}}(N) = 0$ .

Choosing  $(t_k)_{k \in \mathbb{N}} = (L(N))_{N \in \mathbb{N}}$ , we complete the proof of Lemma 8.11.  $\square$

• **Proof of Lemma 8.6**

*Proof.* The goal is to prove that  $\nu_\theta$  is continuous in  $\theta$ . To do so, let  $(\theta_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 1]$  (note that  $\theta_n$  is not a random variable) such that  $\lim_{n \rightarrow \infty} \theta_n = \theta$ . Couple the two infinite systems

$$(9.61) \quad \begin{aligned} dx_i^n(t) &= c[\theta_n - x_i^n(t)] dt + \sqrt{g(x_i^n(t))} dw_i(t) + Ke[y_i^n(t) - x_i^n(t)] dt, \\ dy_i^n(t) &= e[x_i^n(t) - y_i^n(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

and

$$(9.62) \quad \begin{aligned} dx_i(t) &= c[\theta - x_i(t)] dt + \sqrt{g(x_i(t))} dw_i(t) + Ke[y_i(t) - x_i(t)] dt, \\ dy_i(t) &= e[x_i(t) - y_i(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

via their Brownian motions, like in the proof of Lemma 8.11. 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

$$(9.63) \quad \lim_{n \rightarrow \infty} \mathbb{E}[|\Delta_i^n(0)| + K|\delta_i^n(0)|] = 0.$$

By a similar argument as in the proof of Lemma 8.11, we obtain that, for all  $t \geq 0$ ,

$$(9.64) \quad \lim_{n \rightarrow \infty} \mathbb{E}[|\Delta_i^n(t)| + K|\delta_i^n(t)|] = 0.$$

Hence we can construct a sequence  $(L(n))_{n \in \mathbb{N}}$  satisfying  $\lim_{n \rightarrow \infty} L(n) = \infty$  and  $\lim_{n \rightarrow \infty} L(n)/n = 0$  such that

$$(9.65) \quad \lim_{n \rightarrow \infty} \mathbb{E}[|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|] = 0.$$

To prove the continuity of the equilibrium  $\nu_\theta$  in  $\theta$ , we reason as follows. First note that we can couple the system in (9.61) starting from  $\delta_{(\theta_n, \theta_n)}$  with the system in (9.61) starting from  $\nu_{\theta_n}$ . By the uniqueness and convergence to equilibrium (see Lemma 8.5), we see that this coupling is successful. Similarly, we can couple the system in (9.62) starting from  $\delta_{(\theta, \theta)}$  with the system in (9.62) starting from  $\nu_{\theta_n}$ , and see the coupling successful. Finally, we use (9.65) to obtain

$$(9.66) \quad \begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}^{\nu_{\theta_n} \times \nu_\theta} [|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|] \\ & \leq \lim_{n \rightarrow \infty} \mathbb{E}^{\nu_{\theta_n} \times \delta_{(\theta_n, \theta_n)}} [|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|] \\ & \quad + \lim_{n \rightarrow \infty} \mathbb{E}^{\delta_{(\theta, \theta)} \times \delta_{(\theta_n, \theta_n)}} [|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|] \\ & \quad + \lim_{n \rightarrow \infty} \mathbb{E}^{\nu_\theta \times \delta_{(\theta, \theta)}} [|\Delta_i^n(L(n))| + K|\delta_i^n(L(n))|] = 0. \end{aligned}$$

Let  $f$  be a Lipschitz function. Then, by the equilibrium property of  $\nu_{\theta_n}$  and  $\nu_\theta$ ,

$$(9.67) \quad \begin{aligned} & \lim_{n \rightarrow \infty} |\mathbb{E}^{\nu_{\theta_n}} [f(x^n(0))] - \mathbb{E}^{\nu_\theta} [f(x(0))]| \\ & = \lim_{n \rightarrow \infty} |\mathbb{E}^{\nu_{\theta_n}} [f(x^n(L(n)))] - \mathbb{E}^{\nu_\theta} [f(x(L(n)))]| \\ & = \lim_{n \rightarrow \infty} |\mathbb{E}^{\nu_{\theta_n} \times \nu_\theta} [f(x^n(L(n))) - f(x(L(n)))]| \\ & \leq \lim_{n \rightarrow \infty} (\text{Lip} f) \mathbb{E}^{\nu_{\theta_n} \times \nu_\theta} [|x^n(L(n)) - x(L(n))|] = 0. \end{aligned}$$

We can also show this if  $f$  is a Lipschitz function of the  $y$  component. Hence  $\nu_\theta$  is indeed continuous as a function of  $\theta$ .  $\square$

• **Proof of Lemma 8.13**

*Proof.* Note that for all  $N \in \mathbb{N}$  fixed, by Itô-calculus we find from (7.15) that

$$(9.68) \quad \begin{aligned} & \frac{d}{dt} \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K |y_i^{[N],1}(t) - y_i^{[N],2}(t)| \right] \\ &= -\frac{2c}{N} \sum_{j \in [N]} \mathbb{E} \left[ |x_{j,1}^{[N]}(t) - x_{j,2}^{[N]}(t)| \mathbf{1}_{\{\text{sign}(x_{j,1}^{[N]}(t) - x_{j,2}^{[N]}(t)) \neq \text{sign}(x_i^{[N],1}(t) - x_i^{[N],2}(t))\}} \right] \\ & \quad - 2Ke \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| \right. \\ & \quad \left. + |y_i^{[N],1}(t) - y_i^{[N],2}(t)| \mathbf{1}_{\{\text{sign}(x_i^{[N],1}(t) - x_i^{[N],2}(t)) \neq \text{sign}(y_i^{[N],1}(t) - y_i^{[N],2}(t))\}} \right]. \end{aligned}$$

Therefore, for each  $N \in \mathbb{N}$ ,

$$(9.69) \quad t \mapsto \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K |y_i^{[N],1}(t) - y_i^{[N],2}(t)| \right] \text{ is decreasing.}$$

Fix any  $t(N) \rightarrow \infty$ . By the assumption in Lemma 8.13, the proofs of Lemma 8.9 and Lemma 8.11 imply that (9.56) holds for both  $(X^{[N],1}, Y^{[N],1})$  and  $(X^{[N],2}, Y^{[N],2})$ . Using the construction in the proof of Lemma 8.11, we can construct *one* sequence  $(l(N))_{N \in \mathbb{N}}$ , satisfying  $l(N) \leq t(N)$ ,  $\lim_{N \rightarrow \infty} l(N) = \infty$  and  $\lim_{N \rightarrow \infty} l(N)/N = 0$ , such that (9.60) 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

$$(9.70) \quad \begin{aligned} & \mathbb{E} \left[ |x_i^{[N],1}(l(N)) - x_i^{[N],2}(l(N))| + K |y_i^{[N],1}(l(N)) - y_i^{[N],2}(l(N))| \right] \\ & \leq \mathbb{E} \left[ |x_i^{[N],1}(l(N)) - x_i^{\mu,1}(l(N))| + K |y_i^{[N],1}(l(N)) - y_i^{\mu,1}(l(N))| \right] \\ & \quad + \mathbb{E} \left[ |x_i^{\mu,1}(l(N)) - x_i^{\mu,2}(l(N))| + K |y_i^{\mu,1}(l(N)) - y_i^{\mu,2}(l(N))| \right] \\ & \quad + \mathbb{E} \left[ |x_i^{\mu,2}(l(N)) - x_i^{[N],2}(l(N))| + K |y_i^{\mu,2}(l(N)) - y_i^{[N],2}(l(N))| \right]. \end{aligned}$$

Note that in the right-hand side of the inequality the first and the third term tend to zero by (9.60). The second term tends to zero because  $\mu\{\bar{\Theta}_1 = \bar{\Theta}_2\} = 1$ , and hence Lemma 8.5 can be applied. Therefore

$$(9.71) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ |x_i^{[N],1}(l(N)) - x_i^{[N],2}(l(N))| + K |y_i^{[N],1}(l(N)) - y_i^{[N],2}(l(N))| \right] = 0.$$

Using the monotonicity in (9.69), we get

$$(9.72) \quad \begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{E} \left[ |x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K |y_i^{[N],1}(t(N)) - y_i^{[N],2}(t(N))| \right] \\ & \leq \lim_{N \rightarrow \infty} \mathbb{E} \left[ |x_i^{[N],1}(l(N)) - x_i^{[N],2}(l(N))| + K |y_i^{[N],1}(l(N)) - y_i^{[N],2}(l(N))| \right] = 0. \end{aligned}$$

□

Combining the proofs of Proposition 8.4, Lemma 8.11 and Lemma 8.13, we obtain the following corollary. This corollary turns out to be important in Section 9.2 in the proof of Lemma 8.16 to obtain the limiting evolution of the 1-blocks.

**Corollary 9.1.** *Fix  $s > 0$ . Let  $\mu_N$  be the measure obtained by periodic configuration of*

$$(9.73) \quad \mathcal{L}[X^{[N]}(Ns - L(N)), Y^{[N]}(Ns - L(N))],$$

*and let  $\mu$  be a weak limit point of the sequence  $(\mu_N)_{N \in \mathbb{N}}$ . Let*

$$(9.74) \quad \bar{\Theta} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} \frac{x_i^\mu + K y_i^\mu}{1 + K} \text{ in } L^2(\mu).$$

*and let  $(X^{\nu\bar{\Theta}}, Y^{\nu\bar{\Theta}})$  be the infinite system evolving according to (8.41) 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  $t \geq 0$ ,*

$$(9.75) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ |x_i^{[N]}(Ns + t) - x_i^{\nu\bar{\Theta}}(t)| \right] + K \mathbb{E} \left[ |y_i^{[N]}(Ns + t) - y_i^{\nu\bar{\Theta}}(t)| \right] = 0, \quad \forall i \in [N].$$

*Proof.* By Proposition 8.4 we have that  $\lim_{k \rightarrow \infty} \mathcal{L}[X^{[N_k]}(N_k s) + Y^{[N_k]}(N_k s)] = \nu(s) = \nu_{\bar{\Theta}}$ . Let  $\nu_{N_k}$  be defined by periodic continuation of the configuration of  $(X^{[N_k]}(N_k s), Y^{[N_k]}(N_k s))$  and let  $\nu = \lim_{k \rightarrow \infty} \nu_{N_k}$ , 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 8.9 and Lemma 8.11 imply (9.75).  $\square$

## 9.2 Proof of step 2. Convergence of the estimator

In this section we prove the three lemmas stated in Step 2 of Section 8.2. Afterwards we prove Proposition 8.14 with the help of these lemmas.

### • Proof of Lemma 8.15

*Proof.* Recall the definition of  $\bar{\Theta}^{[N]}(t)$  in (8.2). It follows from the SSDE in (7.15) that

$$(9.76) \quad d\bar{\Theta}^{[N]}(t) = \frac{1}{1+K} \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N]}(t))} dw_i(t).$$

Hence we see that  $t \mapsto \bar{\Theta}^{[N]}(t)$  is a continuous martingale. By Itô's formula we have

$$(9.77) \quad \begin{aligned} \mathbb{E}[(\bar{\Theta}^{[N]}(t))^2] &= \mathbb{E}[(\bar{\Theta}^{[N]}(0))^2] + \frac{1}{(1+K)^2} \int_0^t dr \frac{1}{N^2} \sum_{i \in [N]} g(x_i^{[N]}(r)) \\ &\leq 1 + \frac{1}{N} \frac{\|g\|}{(1+K)^2} t. \end{aligned}$$

Since  $g$  is a bounded function, we get that  $t \mapsto \bar{\Theta}^{[N]}(t)$  is square integrable. It follows that,

$$(9.78) \quad \left( (\bar{\Theta}^{[N]}(Ns+t) - \bar{\Theta}^{[N]}(Ns))^2 \right)_{t \geq 0}$$

is a sub-martingale. Therefore, defining the stopping time

$$(9.79) \quad S_\epsilon^N = \inf \left\{ t \geq 0 : (\bar{\Theta}^{[N]}(Ns+t) - \bar{\Theta}^{[N]}(Ns))^2 \geq \epsilon \right\} \wedge L(N),$$

we find, by the continuity of  $t \mapsto \bar{\Theta}^{[N]}(Ns+t)$  and the optional sampling theorem, that

$$(9.80) \quad \mathbb{P}(S_\epsilon^N \in [Ns, Ns + L(N)]) \leq \frac{1}{\epsilon^2} \mathbb{E} \left[ (\bar{\Theta}^{[N]}(Ns + L(N)) - \bar{\Theta}^{[N]}(Ns))^2 \right].$$

Combining (9.77) and (9.80), we find

$$(9.81) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} |\bar{\Theta}^{[N]}(Ns+t) - \bar{\Theta}^{[N]}(Ns)| = 0 \text{ in probability.}$$

To obtain the increasing process, note that by Itô-calculus it follows from (9.76) that

$$(9.82) \quad \langle \bar{\Theta}^{[N]}(Ns) \rangle_{s \geq 0} = \frac{1}{(1+K)^2} \int_0^s dr \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Nr)).$$

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 (9.82) implies that, for all  $N \in \mathbb{N}$  and  $s \geq 0$ ,

$$(9.83) \quad \langle \bar{\Theta}^{[N]}(Ns) \rangle \leq \frac{\|g\|}{(1+K)^2} s$$

and

$$(9.84) \quad \frac{d}{ds} \langle \bar{\Theta}^{[N]}(Ns) \rangle \leq \frac{\|g\|}{(1+K)^2}.$$

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. [Bil99, 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  $[0, T] \rightarrow \mathbb{R}$ . 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. [RY99, Chapter 5]):

$$(9.85) \quad (\bar{\Theta}^{[N]}(Ns) - \bar{\Theta}^{[N]}(0))_{s \geq 0} = w(\langle \bar{\Theta}^{[N]}(Ns) \rangle)_{s \geq 0}.$$

Let  $\chi$  be a standard normal random variable. Then

$$(9.86) \quad \mathbb{E} \left[ \left( w(\langle \bar{\Theta}^{[N]}(Ns) \rangle) - w(\langle \bar{\Theta}^{[N]}(Nr) \rangle) \right)^2 \right] \leq \mathbb{E} \left[ \left( \langle \bar{\Theta}^{[N]}(Ns) \rangle - \langle \bar{\Theta}^{[N]}(Nr) \rangle \right)^2 \right] \mathbb{E} [\chi^4] \\ \leq (s-r)^2 \frac{\|g\|^2}{(1+K)^4} \mathbb{E} [\chi^4].$$

Hence it follows from Kolmogorov's criterion for weak compactness (see e.g. [RY99, Chapter XIII, Theorem 1.8]) that the sequence  $(\mathcal{L}(\langle \bar{\Theta}^{[N]}(Ns) \rangle_{s \geq 0})_{N \in \mathbb{N}}$  is tight in  $\mathcal{P}(C([0, \infty), \mathbb{R}))$ .  $\square$

### • Proof of Lemma 8.16

*Proof.* For ease of notation we will suppress the subsequence notation and assume that

$$(9.87) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ (\bar{\Theta}^{[N]}(Ns))_{s > 0} \right] = \mathcal{L}[(\bar{\Theta}(s))_{s > 0}].$$

The processes  $(\bar{\Theta}^{[N]}(Ns))_{s \geq 0}$  are martingales, see (9.76), measurable w.r.t. the canonical filtration  $(\mathcal{F}_s)_{s \geq 0}$  and so are their weak limit points. Therefore also the weak limit point  $(\bar{\Theta}(s))_{s > 0}$  is a martingale, (see [DG93b, Section 3]). To obtain (8.52), we use the following strategy. Recall from the proof of Lemma 8.15 that the sequence  $\{(\bar{\Theta}^{[N]}(Ns))_{s > 0}\}_{N \in \mathbb{N}}$  is tight. Hence, in order to prove that

$$(9.88) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \langle \bar{\Theta}^{[N]}(Ns) \rangle_{s > 0} \right] = \mathcal{L} \left[ \left( \int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)] \right)_{s > 0} \right],$$

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  $(\int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)])_{s > 0}$ . We will prove a slightly stronger result, namely,

$$(9.89) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \langle \bar{\Theta}^{[N]}(Ns) \rangle - \int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)] \right| \right] = 0.$$

Once we have obtained (9.89) and hence (9.88), by Skorohod'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

$$(9.90) \quad \lim_{N \rightarrow \infty} \langle \bar{\Theta}^{[N]}(Ns) \rangle_{s \geq 0} = \left( \int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)] \right)_{s \geq 0} \quad a.s.$$

Using the continuity of Brownian motion, we get that (recall (9.85))

$$(9.91) \quad \lim_{N \rightarrow \infty} (\bar{\Theta}^{[N]}(Ns))_{s > 0} = \lim_{N \rightarrow \infty} \left[ w(\langle \bar{\Theta}^{[N]}(Ns) \rangle)_{s > 0} + \bar{\Theta}^{[N]}(0) \right] \\ = w \left( \int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}} [g(x_0)] \right)_{s > 0} + \vartheta_0 \quad a.s.$$

Therefore we can choose a version of  $(\bar{\Theta}(s))_{s > 0}$  such that

$$(9.92) \quad \lim_{N \rightarrow \infty} (\bar{\Theta}^{[N]}(Ns))_{s > 0} = \lim_{N \rightarrow \infty} (\bar{\Theta}(s))_{s > 0} \quad a.s.$$



and

$$(9.93) \quad \lim_{N \rightarrow \infty} (\bar{\Theta}^{[N]}(Ns), \langle \bar{\Theta}^{[N]}(Ns) \rangle)_{s>0} = (\bar{\Theta}(s), \langle \bar{\Theta}(s) \rangle)_{s>0} \quad a.s.$$

By the continuous mapping theorem, (8.52) follows. The martingale property follows from the fact that  $(\bar{\Theta}^{[N]}(Ns)^2 - \langle \bar{\Theta}^{[N]}(Ns) \rangle)_{s>0}$  are martingales. Therefore, to finish the proof of Lemma 8.16 we are left to prove (9.89).

To prove (9.89), define the empirical measures on  $[0, 1]$  by

$$(9.94) \quad U^{[N]}(Ns) = \frac{1}{N} \sum_{i \in [N]} \delta_{x_i(Ns)}.$$

Note that we can write

$$(9.95) \quad \begin{aligned} & \mathbb{E} \left[ \left| \langle \bar{\Theta}^{[N]}(Ns) \rangle - \int_0^s dr \frac{1}{(1+K)^2} \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right] \\ &= \frac{1}{(1+K)^2} \mathbb{E} \left[ \left| \int_0^s dr \mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \int_0^s dr \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right] \\ &\leq \frac{1}{(1+K)^2} \int_0^s dr \mathbb{E} \left[ \left| \mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right]. \end{aligned}$$

Hence, to prove (9.89) it is enough to prove that, for all  $r > 0$ ,

$$(9.96) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right] = 0$$

and apply the dominated convergence theorem.

To prove (9.96), we will use the coupling arguments from Section 9.1, as well as ergodicity and invariance under the evolution of  $\nu_{\bar{\Theta}}$ . As before, let  $z^{[N]}(t)$  denote the  $[N]$ -component system  $(x_i^{[N]}(t), y_i^{[N]}(t))_{t \geq 0}$  evolving according to (7.15), 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  $(z^{\mu_N}(t))_{t \geq 0}$  denote the infinite system  $(x_i^{\mu_N}(t), y_i^{\mu_N}(t))_{t \geq 0}$  evolving according to (8.33) with initial law  $\mu_N$  obtained by periodic continuation of the configuration of  $(x_i^{[N]}(Nr - L(N)), y_i^{[N]}(Nr - L(N)))$ , and let  $\mu$  be a weak limit point of the sequence  $\mu_N$ . Note that, for all  $r > 0$ , by Lemma 8.10 we have that  $\lim_{N \rightarrow \infty} \bar{\Theta}^{[N]}(Nr) = \bar{\Theta}(r)$  for  $\bar{\Theta}(r) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i + Ky_i}{1+K}$  in  $L_2(\mu)$ . Let  $\bar{L}(N)$  be the sequence constructed in Corollary 9.1. Then we can write

$$(9.97) \quad \begin{aligned} & \mathbb{E} \left[ \left| \mathbb{E}^{U^{[N]}(Nr)}[g(x_0)] - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right] \\ &\leq \mathbb{E} \left[ \left| \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Nr)) - \frac{1}{N} \sum_{i \in [N]} g(x_i^{\nu_{\bar{\Theta}(r)}}(\bar{L}(N))) \right| \right] \\ &\quad + \mathbb{E} \left[ \left| \frac{1}{N} \sum_{i \in [N]} g(x_i^{\nu_{\bar{\Theta}(r)}}(\bar{L}(N))) - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right] \\ &\leq (\text{Lip } g) \mathbb{E} \left[ \left| x_0^{[N]}(Nr) - x_0^{\nu_{\bar{\Theta}(r)}}(\bar{L}(N)) \right| \right] \\ &\quad + \mathbb{E} \left[ \left| \frac{1}{N} \sum_{i \in [N]} g(x_i^{\nu_{\bar{\Theta}(r)}}(\bar{L}(N))) - \mathbb{E}^{\nu_{\bar{\Theta}(r)}}[g(x_0)] \right| \right], \end{aligned}$$

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  $N \rightarrow \infty$  by Corollary 9.1. Finally, note that by Lemma 8.5  $(x_i)_{i \in \mathbb{N}_0}$  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.  $\square$

• **Proof of Lemma 8.17**

*Proof.* Note that, since  $g$  is Lipschitz, the function

$$(9.98) \quad \theta \mapsto \mathbb{E}^{\nu_\theta}[g],$$

is Lipschitz by Lemma 8.6. Hence, by [YW71, Theorem 1], the SDE

$$(9.99) \quad d\Phi(s) = \frac{1}{(1+K)} \sqrt{\mathbb{E}^{\nu_{\Phi(s)}}[g]} dw(s)$$

has a pathwise unique solution (see [p315–317]KS91) and a unique solution in law (see [YW71, Proposition 1]). This implies that the martingale problem with generator

$$(9.100) \quad \frac{1}{(1+K)^2} \mathbb{E}^{\nu_\Phi}[g] \frac{d}{d\Phi^2}$$

has a unique solution. In particular, choosing  $f(\Phi) = \Phi^2$ , we see that the martingale problem implies that

$$(9.101) \quad \left( \Phi^2(s) - \frac{1}{(1+K)^2} \int_0^s du \mathbb{E}^{\nu_{\Phi(u)}}[g(x_0)] \right)_{s>0}$$

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 (8.53) now follows from the fact that the quadratic variation of a martingale is unique.  $\square$

• **Proof of Proposition 8.14**

*Proof.* Combining Lemma 8.16–8.17, 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 (8.53).  $\square$

### 9.3 Proof of step 3. Convergence of the 1-blocks in the Meyer-Zheng topology

In this section we prove Proposition 8.18 stated in Step 3 of Section 8.2. The Lemmas 8.19, 8.20 and 8.21 are proven in Appendix B.

• **Proof of Proposition 8.18**

*Proof.* By Proposition 8.14 we have that

$$(9.102) \quad \lim_{N \rightarrow \infty} \mathcal{L}[\bar{\Theta}^{[N]}(Ns)_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}]$$

in the normal topology and therefore (see B Lemma B.1)

$$(9.103) \quad \lim_{N \rightarrow \infty} \mathcal{L}[\bar{\Theta}^{[N]}(Ns)_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \text{ in Meyer-Zheng topology.}$$

By Lemma 8.7, for  $s > 0$ ,

$$(9.104) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{[N]}(Ns) - x_1^{[N]}(s) \right| \right] = 0$$

and therefore, by Lemma 8.19,

$$(9.105) \quad \lim_{N \rightarrow \infty} d_P(\psi_{\bar{\Theta}^{[N]}}, \psi_{x_1^{[N]}}) = 0 \text{ in probability.}$$

To apply the above results to our model, we recall the following basic result (see [Bil99, Chapter 1]), which also holds for the Meyer-Zheng topology.

**Lemma 9.2.** *Let  $X_n, Y_n$  be random variables. If*

$$(9.106) \quad \lim_{n \rightarrow \infty} \mathcal{L}[X_n] = \mathcal{L}[X]$$

and

$$(9.107) \quad \lim_{n \rightarrow \infty} d(X_n, Y_n) = 0 \text{ in probability,}$$

then

$$(9.108) \quad \lim_{n \rightarrow \infty} \mathcal{L}[Y_n] = \mathcal{L}[X].$$

Applying Lemma (9.2) to our case, we obtain

$$(9.109) \quad \lim_{N \rightarrow \infty} \mathcal{L}[(x_1^{[N]}(s))_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \text{ in the Meyer-Zheng topology.}$$

The argument for

$$(9.110) \quad \lim_{N \rightarrow \infty} \mathcal{L}[(y_1^{[N]}(s))_{s>0}] = \mathcal{L}[(\bar{\Theta}(s))_{s>0}] \text{ in the Meyer-Zheng topology}$$

follows in the same way. By Lemma 8.20, we obtain

$$(9.111) \quad \lim_{N \rightarrow \infty} \mathcal{L}[(x_1^{[N]}(s), y_1^{[N]}(s) - x_1^{[N]}(s))_{s>0}] = \mathcal{L}[(\bar{\Theta}(s), 0)_{s>0}] \text{ in the Meyer-Zheng topology.}$$

Applying Lemma 8.21 with  $f(x, y) = f(x, y + x)$  and the continuous mapping theorem, we obtain

$$(9.112) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( x_1^{[N]}(s), y_1^{[N]}(s) \right)_{s>0} \right] = \mathcal{L}[(\bar{\Theta}(s), \bar{\Theta}(s))_{s>0}] \text{ in the Meyer-Zheng topology.}$$

□

## 9.4 Proof of step 4. Mean-field finite-systems scheme

### • Proof of Proposition 8.1

*Proof.* Proposition 8.1(b) follows directly from Proposition 8.14. The proof of Proposition 8.1(a) follows from Proposition 8.1(b) and Proposition 8.18.

To prove Proposition 8.1(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 (8.14). 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 8.4, Proposition 8.14 and Lemma 8.15, we get that, for each converging subsequence,

$$(9.113) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ X^{[N_k]}(sN_k + t), Y^{[N_k]}(sN_k + t) \right] = \int Q_s(\theta, d\theta') \nu_{\theta'} = \nu(s),$$

and hence it follows that, for all  $t \geq 0$ ,

$$(9.114) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ X^{[N]}(sN + t), Y^{[N]}(sN + t) \right] = \int Q_s(\theta, d\theta') \nu_{\theta'} = \nu(s).$$

Let  $(X^{\nu(s)}(t), Y^{\nu(s)}(t))_{t \geq 0}$  be the infinite system defined in (7.18), starting from initial measure  $\nu(s)$ . Then it follows from Corollary 9.1 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  $t \geq 0$ ,

$$(9.115) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| z_{(i, R_i)}^{\nu(s)}(t) - z_{(i, R_i)}^{[N]}(sN + t) \right| \right] = 0 \quad \forall (i, R_i) \in \mathbb{Z} \times \{A, D\}.$$

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 the process  $(X^{\nu(s)}(t), Y^{\nu(s)}(t))_{t \geq 0}$ .

Since we want that

$$(9.116) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ (X^{[N]}(sN+t), Y^{[N]}(sN+t))_{t \geq 0} \right] = \mathcal{L} \left[ (X^{\nu(s)}(t), Y^{\nu(s)}(t))_{t \geq 0} \right],$$

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  $(x_i^{[N]}(t))_{t \geq 0}$  and  $(y_i^{[N]}(t))_{t \geq 0}$  that they are tight in path space (see [Bil99, Theorem 7.3]).

To prove that the components are tight, we use a tightness criterion for semi-martingales by Joffe and Metivier, [JM86, Proposition 3.2.3]. To use this criterion, we have to show that for all  $i \in [N]$  the components  $(x_i^{[N]}(t), y_i^{[N]}(t))$  are  $\mathcal{D}$ -semi-martingales as defined in [JM86, Definition 3.1.1]. To do this, let  $\mathcal{C}^* \subset \mathcal{C}_b([0, 1]^2)$  be the set of polynomials on  $[0, 1]^2$ , and define the operator  $G_{\dagger}^{[N]}: (\mathcal{C}^* \times [0, 1]^2 \times [0, \infty), \Omega) \rightarrow \mathbb{R}$  by

$$(9.117) \quad G_{\dagger}^{[N]}(f, (x, y), t, \omega) = \left[ \frac{c}{N} \sum_{j \in [N]} [x_j^{[N]}(t, \omega) - x] + K [y - x] \right] \frac{\partial f}{\partial x} + \frac{1}{2} g(x) \frac{\partial^2 f}{\partial x^2} + e [x - y] \frac{\partial f}{\partial y}.$$

We use the subscript  $\dagger$  to emphasize that  $G_{\dagger}^{[N]}$  is the operator of a  $\mathcal{D}$ -semi-martingale and not a generator. Below we check in 4 steps that the component processes  $(x_i^{[N]}(t), y_i^{[N]}(t))_{t \geq 0}$  are indeed  $\mathcal{D}$ -semi-martingales.

1. The functions

$$(9.118) \quad f_1(x_i, y_i) = x_i, \quad f_2(x_i, y_i) = y_i,$$

are in  $\mathcal{C}^*$ , and so are  $f_1^2, f_1 f_2, f_2^2$ .

2. For every  $((x, y), t, \omega) \in ([0, 1]^2 \times [0, \infty) \times \Omega)$ , the mapping  $f \mapsto G_{\dagger}^{[N]}(f, (x, y), t, \omega)$  is linear on  $\mathcal{C}^*$  and  $G_{\dagger}^{[N]}(f, \cdot, t, \omega) \in \mathcal{C}^*$ .
3. 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  $\sigma$ -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  $(x_j^{[N]}(t))_{t \geq 0}$  are continuous,  $((x, y), t, \omega) \mapsto G_{\dagger}^{[N]}(f, (x, y), t, \omega)$  is  $\mathcal{B}([0, 1]^2) \otimes \mathcal{P}$  measurable for every  $f \in \mathcal{C}^*$ , where  $\mathcal{P}$  is the  $\sigma$ -algebra generated by the sets of the form  $(s, t] \times F$  for  $F \in \mathcal{F}_s$ .
4. Applying Itô's formula to the SSDE in (7.15), we obtain, for every  $f \in \mathcal{C}^*$ ,

$$(9.119) \quad \begin{aligned} f(x_i^{[N]}(t), y_i^{[N]}(t)) &= f(x_i^{[N]}(0), y_i^{[N]}(0)) \\ &+ \int_0^t ds \frac{c}{N} \sum_{j \in [N]} [x_j^{[N]}(s, \omega) - x_i^{[N]}(s)] \frac{\partial f}{\partial x}(x_i^{[N]}(t), y_i^{[N]}(t)) \\ &+ \frac{1}{2} \int_0^t dw_i(s) \sqrt{g(x_i^{[N]}(s))} \frac{\partial f}{\partial x}(x_i^{[N]}(t), y_i^{[N]}(t)) \\ &+ \int_0^t ds K e [y_i^{[N]}(s) - x_i^{[N]}(s)] \frac{\partial f}{\partial x}(x_i^{[N]}(t), y_i^{[N]}(t)) \\ &+ \int_0^t ds e [x_i^{[N]}(s) - y_i^{[N]}(s)] \frac{\partial f}{\partial y}(x_i^{[N]}(t), y_i^{[N]}(t)) \\ &+ \int_0^t ds g(x_i^{[N]}(s)) \frac{\partial^2 f}{\partial x^2}(x_i^{[N]}(t), y_i^{[N]}(t)). \end{aligned}$$

Therefore

$$(9.120) \quad \begin{aligned} M^{[N], f}(t, \omega) &= f(x_i^{[N]}(t, \omega), y_i^{[N]}(t, \omega)) - f(x_i^{[N]}(0, \omega), y_i^{[N]}(0, \omega)) \\ &- \int_0^t ds G_{\dagger}^{[N]}(f(x_i^{[N]}(s, \omega), y_i^{[N]}(s, \omega)), s, \omega) \end{aligned}$$

is a square-integrable martingale on  $(\Omega, (\mathcal{F}_s)_{s \geq 0}, \mathbb{P})$ .

To check that the sequence of component processes  $((x_i^{[N]}(t), y_i^{[N]}(t)))_{N \in \mathbb{N}}$  is tight, we need the local characteristics of the  $\mathcal{D}$ -semi-martingale, which are defined in [JM86, Definition 3.1.2] as (recall (9.118))

$$(9.121) \quad \begin{aligned} b_1^{[N]}((x, y), t, \omega) &= G_{\dagger}^{[N]}(f_1, (x, y), t, \omega), \\ b_2^{[N]}((x, y), t, \omega) &= G_{\dagger}^{[N]}(f_2, (x, y), t, \omega), \\ a_{(1,1)}^{[N]}((x, y), t, \omega) &= G_{\dagger}^{[N]}(f_1 f_1, (x, y), t, \omega) - 2x b_1((x, y), t, \omega), \\ a_{(2,1)}^{[N]}((x, y), t, \omega) &= G_{\dagger}^{[N]}(f_1 f_2, (x, y), t, \omega) - x b_2((x, y), t, \omega) - y b_1((x, y), t, \omega), \\ a_{(1,2)}^{[N]}((x, y), t, \omega) &= a_{(2,1)}((x, y), t, \omega), \\ a_{(2,2)}^{[N]}((x, y), t, \omega) &= G_{\dagger}^{[N]}(f_2 f_2, (x, y), t, \omega) - 2y b_2((x, y), t, \omega). \end{aligned}$$

Hence

$$(9.122) \quad \begin{aligned} b_1^{[N]}((x, y), t, \omega) &= \frac{c}{N} \sum_{j \in [N]} [x_j^{[N]}(t, \omega) - x] + Ke[y - x], \\ b_2^{[N]}((x, y), t, \omega) &= e[x - y], \\ a_{(1,1)}^{[N]}((x, y), t, \omega) &= 2g(x), \\ a_{(1,2)}^{[N]}((x, y), t, \omega) &= a_{(2,1)}((x, y), t, \omega) = 0, \\ a_{(2,2)}^{[N]}((x, y), t, \omega) &= 0. \end{aligned}$$

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 [JM86, H1, H2, H3 in Section 3.2.1] are satisfied.

$H_1$ : For all  $N \in \mathbb{N}$ ,

$$(9.123) \quad \sum_{i \in \{1,2\}} |b_i^{[N]}((x, y), t, \omega)|^2 + \sum_{i,j \in \{1,2\}} |a_{(i,j)}^{[N]}((x, y), t, \omega)|^2 \leq 4(c + Ke + e)^2 + 2\|g\|^2.$$

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  $(A^{[N]}(t))_{t \geq 0}$  in [JM86, 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, [JM86, 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 (9.116) indeed holds.  $\square$

## 10 Preparation: $N \rightarrow \infty$ , two-colour mean-field

In this section we extend the results obtained in Section 8.1 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. In particular, we highlight the role of the second colour. Section 10.1 builds up the setting and states the main scaling result: Proposition 10.2. Section 10.2 provides the proof of this proposition based on a series of lemmas, which are stated and proved first.

### 10.1 Two-colour mean-field finite-systems scheme

**Setup.** In this section we consider a simplified version of our SSDE in (2.20) on the finite geographic space

$$(10.1) \quad [N] = \{0, 1, \dots, N-1\}, \quad N \in \mathbb{N}.$$

The migration kernel  $a^{\alpha N}(\cdot, \cdot)$  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

$$(10.2) \quad S = \mathfrak{s}^{[N]}, \quad \mathfrak{s} = [0, 1] \times [0, 1]^2,$$

and the system, consisting of three components, is denoted by

$$(10.3) \quad \begin{aligned} Z^{[N]}(t) &= (X^{[N]}(t), (Y_0^{[N]}(t), Y_1^{[N]}(t)))_{t \geq 0}, \\ (X^{[N]}(t), (Y_0^{[N]}(t), Y_1^{[N]}(t))) &= (x_i(t), (y_{i,0}(t), y_{i,1}(t)))_{i \in [N]}. \end{aligned}$$

The components of  $(Z^{[N]}(t))_{t \geq 0}$  evolve according to the SSDE

$$(10.4) \quad \begin{aligned} dx_i^{[N]}(t) &= \frac{c_0}{N} \sum_{j \in [N]} [x_j^{[N]}(t) - x_i^{[N]}(t)] dt + \sqrt{g(x_i^{[N]}(t))} dw_i(t) \\ &\quad + K_0 e_0 [y_{i,0}^{[N]}(t) - x_i^{[N]}(t)] dt + \frac{K_1 e_1}{N} [y_{i,1}^{[N]}(t) - x_i^{[N]}(t)] dt, \\ dy_{i,0}^{[N]}(t) &= e_0 [x_i^{[N]}(t) - y_{i,0}^{[N]}(t)] dt, \\ dy_{i,1}^{[N]}(t) &= \frac{e_1}{N} [x_i^{[N]}(t) - y_{i,1}^{[N]}(t)] dt, \quad i \in [N], \end{aligned}$$

which is a special case of (2.20). The initial state is  $\mu(0) = \mu^{\otimes [N]}$  for some  $\mu \in \mathcal{P}([0, 1]^3)$ . The SSDE in (10.4) has a unique weak solution coming from a well-posed martingale problem [SS80, Theorem 3.1]. By [SS80, Theorem 3.2], (10.4) has a unique strong solution for every deterministic initial state  $Z^{[N]}(0)$ . Therefore the solution of (10.4) is Feller and Markov for any initial law. The SSDE in (10.4) can alternatively be written as

$$(10.5) \quad \begin{aligned} dx_i^{[N]}(t) &= c_0 \left[ \frac{1}{N} \sum_{j \in [N]} x_j^{[N]}(t) - x_i^{[N]}(t) \right] dt + \sqrt{g(x_i^{[N]}(t))} dw_i(t) \\ &\quad + K_0 e_0 [y_{i,0}^{[N]}(t) - x_i^{[N]}(t)] dt + \frac{K_1 e_1}{N} [y_{i,1}^{[N]}(t) - x_i^{[N]}(t)] dt, \\ dy_{i,0}^{[N]}(t) &= e_0 [x_i^{[N]}(t) - y_{i,0}^{[N]}(t)] dt, \\ dy_{i,1}^{[N]}(t) &= \frac{e_1}{N} [x_i^{[N]}(t) - y_{i,1}^{[N]}(t)] dt, \quad i \in [N]. \end{aligned}$$

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  $N \rightarrow \infty$ , on time scales  $t$  and  $Ns$ . Heuristically, analysing the SSDE in (10.5), we can foresee the following results, which are made precise in Proposition 10.2.

- **On time scale**  $1 = N^0$  (space-time scale 0), in the limit as  $N \rightarrow \infty$  the colour-1 dormant population  $(Y_1^{[N]}(t))_{t \geq 0}$  in (10.4) 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 (7.1), 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

$$(10.6) \quad z_0^{\text{eff},[N]}(t) = (x_0^{[N]}(t), y_{0,0}^{[N]}(t))_{t \geq 0}$$

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 (4.51).

• **On time scale  $N$**  (space-time scale 1), we look at the averages

$$(10.7) \quad \begin{aligned} (z_1^{[N]}(s))_{s>0} &= \left( x_1^{[N]}(s), (y_{0,1}^{[N]}(s), y_{1,1}^{[N]}(s)) \right)_{s>0} \\ &= \left( \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Ns), \left( \frac{1}{N} \sum_{i \in [N]} y_{i,0}^{[N]}(Ns), \frac{1}{N} \sum_{i \in [N]} y_{i,1}^{[N]}(Ns) \right) \right)_{s>0}. \end{aligned}$$

Again the lower index 1 indicates that the average is the analogue of the 1-block average defined in (2.22). Using (10.4), we see that the dynamics of the system in (10.7) is given by the SSDE

$$(10.8) \quad \begin{aligned} dx_1^{[N]}(s) &= \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Ns))} dw(s) + NK_0 e_0 [y_{0,1}^{[N]}(s) - x_1^{[N]}(s)] ds \\ &\quad + K_1 e_1 [y_{1,1}^{[N]}(s) - x_1^{[N]}(s)] ds, \\ dy_{0,1}^{[N]}(s) &= Ne_0 [x_1^{[N]}(s) - y_{0,1}^{[N]}(s)] ds, \\ dy_{1,1}^{[N]}(s) &= e_1 [x_1^{[N]}(s) - y_{1,1}^{[N]}(s)] ds. \end{aligned}$$

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  $N \rightarrow \infty$  the path becomes rougher and rougher at rarer and rarer times. Using the *Meyer-Zheng topology* we can prove that  $\lim_{N \rightarrow \infty} y_{0,1}^{[N]}(s) = \lim_{N \rightarrow \infty} x_1^{[N]}(s)$  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  $\frac{x_1^{[N]}(s) + K_0 y_{0,1}^{[N]}(s)}{1 + K_0}$  the rates with a factor  $N$  in front cancel out. We will use the quantity  $\frac{x_1^{[N]}(s) + K_0 y_{0,1}^{[N]}(s)}{1 + K_0}$  to obtain results in the classical path-space topology. We call

$$(10.9) \quad (z_1^{[N],\text{eff}}(s))_{s>0} = \left( \frac{x_1^{[N]}(s) + K_0 y_{0,1}^{[N]}(s)}{1 + K_0}, y_{1,1}(s) \right)_{s>0}$$

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 (10.4), we need the infinite-dimensional process

$$(10.10) \quad (Z(t))_{t \geq 0} = ((z_i(t))_{t \geq 0})_{i \in \mathbb{N}_0} = ((x_i(t), (y_{i,0}(t), y_{i,1}(t)))_{t \geq 0})_{i \in \mathbb{N}_0}$$

with state space  $([0, 1]^3)^{\mathbb{N}_0}$  that evolves according to

$$(10.11) \quad \begin{aligned} dx_i(t) &= c_0[\theta - x_i(t)] dt + \sqrt{g(x_i(t))} dw_i(t) + K_0 e_0 [y_{i,0}(t) - x_i(t)] dt, \\ dy_{i,0}(t) &= e_0 [x_i(t) - y_{i,0}(t)] dt, \\ y_{i,1}(t) &= y_{i,1}, \quad i \in \mathbb{N}_0. \end{aligned}$$

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

$$(10.12) \quad (Z^{\text{eff}}(t))_{t \geq 0} = ((z_i^{\text{eff}}(t))_{t \geq 0})_{i \in \mathbb{N}_0} = ((x_i^{\text{eff}}(t), (y_{i,0}^{\text{eff}}(t), y_{i,1}^{\text{eff}}(t)))_{t \geq 0})_{i \in \mathbb{N}_0}$$

with state space  $([0, 1]^2)^{\mathbb{N}_0}$  that evolves according to

$$(10.13) \quad \begin{aligned} dx_i^{\text{eff}}(t) &= c_0[\theta - x_i^{\text{eff}}(t)] dt + \sqrt{g(x_i^{\text{eff}}(t))} dw(t) + K_0 e_0 [y_{i,0}^{\text{eff}}(t) - x_i^{\text{eff}}(t)] dt, \\ dy_{i,0}^{\text{eff}}(t) &= e_0 [x_i^{\text{eff}}(t) - y_{i,0}^{\text{eff}}(t)] dt, \quad i \in \mathbb{N}_0. \end{aligned}$$

Like for the one-colour mean-field finite-systems scheme, we need the following list of ingredients to formally state our multi-scaling properties:

1. For positive times  $t > 0$ , we define the so-called *estimators* for the finite system by:

$$(10.14) \quad \begin{aligned} \bar{\Theta}^{(1),[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N]}(t) + K_0 y_{i,0}^{[N]}(t)}{1 + K_0}, \\ \Theta_x^{(1),[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(t), \\ \Theta_{y_0}^{(1),[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} y_{i,0}^{[N]}(t), \\ \Theta_{y_1}^{(1),[N]}(t) &= \frac{1}{N} \sum_{i \in [N]} y_{i,1}^{[N]}(t). \end{aligned}$$

We abbreviate

$$(10.15) \quad \begin{aligned} \Theta^{(1),[N]}(t) &= \left( \Theta_x^{(1),[N]}(t), \Theta_{y_0}^{(1),[N]}(t), \Theta_{y_1}^{(1),[N]}(t) \right), \\ \Theta^{\text{eff},(1),[N]}(t) &= \left( \bar{\Theta}^{(1),[N]}(t), \Theta_{y_1}^{(1),[N]}(t) \right). \end{aligned}$$

We refer to  $(\Theta^{\text{eff},(1),[N]}(t))_{t \geq 0}$  as the *effective estimator process* and to  $(\Theta^{(1),[N]}(t))_{t \geq 0}$  as the *estimator process*.

2. 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  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , 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 10.1. [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  $k$ .  $\blacksquare$

3. The *invariant measure* (i.e., the equilibrium measure) for the evolution of a single colony in (10.11), written

$$(10.16) \quad \Gamma_{\theta, \theta, y_1},$$

and the *invariant measure* of the infinite system in (10.11), 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  $\nu_\theta$  and the convergence of  $\mathcal{L}[Z(t)_{t \geq 0}]$  towards  $\nu_\theta$  will be shown in the proof of Proposition 10.2.

4. The invariant measure of the effective process in (10.13),

$$(10.17) \quad \Gamma_\theta^{\text{eff}},$$

and the invariant measure for the full process,  $\nu_\theta^{\text{eff}} = (\Gamma_\theta^{\text{eff}})^{\otimes \mathbb{N}_0}$ .

5. The renormalisation transformation  $\mathcal{F}: \mathcal{G} \rightarrow \mathcal{G}$ ,

$$(10.18) \quad (\mathcal{F}g)(\theta) = \int_{([0, 1]^2)^{\mathbb{N}_0}} g(x_0) \nu_\theta^{\text{eff}}(dx_0, dy_{0,0}), \quad \theta \in [0, 1],$$



where  $\Gamma_\theta^{\text{eff}}$  is the equilibrium measure of (10.16). Note that this is the same transformation as defined in (4.16), but for the truncated system. Since  $\nu_\theta^{\text{eff}}$  is a product measure, we can write

$$(10.19) \quad (\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x) \Gamma_\theta^{\text{eff}}(dx, dy_0), \quad \theta \in [0, 1],$$

6. 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

$$(10.20) \quad \begin{aligned} dx_1(s) &= \frac{1}{1+K_0} \left[ \sqrt{(\mathcal{F}g)(x_1(s))} dw(s) + K_1 e_1 [y_{1,1}(s) - x_1(s)] ds \right], \\ y_{0,1}(s) &= x_1(s), \\ dy_{1,1}(s) &= e_1 [x_1(s) - y_{1,1}(s)] ds, \end{aligned}$$

where  $\mathcal{F}g$  is defined in (10.20). The effective process  $(z_1^{\text{eff}}(s))_{s>0} = (x_1^{\text{eff}}(s), y_{1,1}^{\text{eff}}(s))_{s>0}$  on space-time scale 1,

$$(10.21) \quad \begin{aligned} dx_1^{\text{eff}}(s) &= \frac{1}{1+K_0} \left[ \sqrt{(\mathcal{F}g)(x_1^{\text{eff}}(s))} dw(s) + K_1 e_1 [y_{1,1}^{\text{eff}}(s) - x_1^{\text{eff}}(s)] ds \right], \\ dy_{1,1}^{\text{eff}}(s) &= e_1 [x_1^{\text{eff}}(s) - y_{1,1}^{\text{eff}}(s)] ds. \end{aligned}$$

We are now ready to state the scaling limit for the evolution of the averages in (10.7), which we refer to as the *mean-field finite-systems scheme with two colours*.

**Proposition 10.2. [Mean-field: two-colour finite-systems scheme]**

Suppose that  $\mathcal{L}[Z^{[N]}(0)] = \mu^{\otimes [N]}$  for some  $\mu \in \mathcal{P}([0, 1] \times [0, 1]^2)$ . Let

$$(10.22) \quad \vartheta_0 = \mathbb{E}^\mu \left[ \frac{x + K_0 y_0}{1 + K_0} \right], \quad \theta_{y_1} = \mathbb{E}^\mu [y_1].$$

(a) For the effective estimator process defined in (10.15),

$$(10.23) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff},(1),[N]}(Ns) \right)_{s>0} \right] = \mathcal{L} \left[ (z_1^{\text{eff}}(s))_{s>0} \right],$$

where the limit is determined by the unique solution of the SSDE (10.21), with initial state

$$(10.24) \quad z_1^{\text{eff}}(0) = (x_1^{\text{eff}}(0), y_{1,1}^{\text{eff}}(0)) = (\vartheta_0, \theta_{y_1}).$$

(b) Assume for the 1-dormant single components that

$$(10.25) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ Y_1^{[N]}(Ns) \middle| \Theta^{(1),[N]}(Ns) \right] = P_{Y_1(s)}^{z_1(s)}.$$

Define

$$(10.26) \quad \Gamma_{(\vartheta_0, \theta_{y_1})}^{\text{eff}}(s) = \int_{[0,1]^2} S_s((\vartheta_0, \theta_{y_1}), d(u_x, u_y)) \Gamma_{u_x}^{\text{eff}} \in \mathcal{P}([0, 1]^2),$$

where  $S_s((\vartheta_0, \theta_{y_1}), \cdot)$  is the time- $s$  marginal law of the process  $(z_1^{\text{eff}}(s))_{s>0}$  starting from  $(\theta_0, \theta_{y_1}) \in [0, 1]^2$  and  $\Gamma_{u_x}^{\text{eff}}$  is the equilibrium distribution of the system in (10.13) with  $\theta = u_x$  (note that  $\Gamma_{\vartheta_0, \theta_{y_1}}^{\text{eff}}(0) = \Gamma_{\vartheta_0}^{\text{eff}}$ ). Let  $(z^{\text{eff}, \Gamma_{(\vartheta_0, \theta_{y_1})}^{\text{eff}}(s)}(t))_{t \geq 0}$  be the process with initial law  $z^{\text{eff}, \Gamma_{(\vartheta_0, \theta_{y_1})}^{\text{eff}}(s)}(0)$  drawn according to  $\Gamma_{(\vartheta_0, \theta_{y_1})}^{\text{eff}}(s)$  (which is a mixture of random processes in equilibrium) that, conditional on  $x_1^{\text{eff}}(s) = \theta$ , evolves according to (10.13). Then, for every  $s \in (0, \infty)$ ,

$$(10.27) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( z_0^{\text{eff},[N]}(Ns+t) \right)_{t \geq 0} \right] = \mathcal{L} \left[ (z^{\text{eff}, \Gamma_{(\vartheta_0, \theta_{y_1})}^{\text{eff}}(s)}(t))_{t \geq 0} \right].$$

(c) For the averages in (10.7),

$$(10.28) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( z_1^{[N]}(s) \right)_{s>0} \right] = \mathcal{L} \left[ (z_1(s))_{s>0} \right]$$

in the Meyer-Zheng topology,

where the limit process is the unique solution of the SSDE in (10.20) with initial state

$$(10.29) \quad z_1(0) = (x_1(0), y_{0,1}(0), y_{1,1}(0)) = (\vartheta_0, \vartheta_0, \theta_{y_1}).$$

(d) Assume 10.25 and define

$$(10.30) \quad \nu(s) = \int_{[0,1]^3} S_s((\vartheta_0, \vartheta_0, \theta_{y_1}), \cdot) \int_{[0,1]^{\mathbb{N}_0}} P_{Y_1(s)}^{(u_x, u_x, u_{y_1})}(\mathrm{d}\mathbf{y}_1) \nu_{u_x, \mathbf{y}_1},$$

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 (10.20), starting from  $(\vartheta_0, \vartheta_0, \theta_{y_1}) \in [0, 1]^3$ , and  $\nu_{u_x, \mathbf{y}_1}$  is the equilibrium distribution of the system in (10.11) with  $\theta = u_x$  and  $(y_{i,1})_{i \in \mathbb{N}_0} = \mathbf{y}_1$ , (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  $Y_1(s) = \mathbf{y}_1$  evolves according to (10.11) with  $\theta = u_x$  and  $(y_{i,1})_{i \in \mathbb{N}_0} = \mathbf{y}_1$ . Then, for every  $s \in (0, \infty)$ ,

$$(10.31) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( Z^{[N]}(Ns + t) \right)_{t \geq 0} \right] = \mathcal{L} \left[ (z^{\nu(s)}(t))_{t \geq 0} \right].$$

**Remark 10.3 (Law of 1-dormant single components).** Note that

$$(10.32) \quad \left( \mathcal{L} \left[ Y_1^{[N]}(Ns) \mid \left( \bar{\Theta}^{[N]}(Ns), \Theta_{y_1}^{[N]}(Ns) \right) \right] \right)_{N \in \mathbb{N}_0}$$

is a tight sequence of measures. Hence there exist weak limit points. In Section 11 we will see that if there is a higher layer in the hierarchy, then we can show that all weak limit points of (10.32) are the same and we can identify the limit. For Theorems 4.2 and 4.4 we do not need this assumption, since there will always be multiple higher levels.  $\blacksquare$

## 10.2 Proof of the two-colour mean-field finite-systems scheme

The proof of Proposition 10.2, the finite-systems scheme with one level and two colours, follows the strategy used in Section 9 for the proof of Proposition 8.1. 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 (10.15).

$$(10.33) \quad \left( (\Theta^{\text{eff},(1),[N]}(Ns))_{s>0} \right)_{N \in \mathbb{N}}$$

2. Stability property of  $(\Theta^{\text{eff},(1),[N]}(Ns + t))_{t>0}$ , i.e., for  $L(N)$  satisfying  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , and all  $\epsilon > 0$ ,

$$(10.34) \quad \lim_{N \rightarrow \infty} \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N]}(Ns) - \bar{\Theta}^{(1),[N]}(Ns - t) \right| > \epsilon \right] = 0.$$

and

$$(10.35) \quad \lim_{N \rightarrow \infty} \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_1}^{(1),[N]}(Ns) - \Theta_{y_1}^{(1),[N]}(Ns - t) \right| > \epsilon \right] = 0.$$

3. Equilibrium of the infinite system and the one-dimensional distribution of the effective single components  $(Z(Ns + t))_{t>0}$ , analogous to Proposition 8.4.
4. Limiting evolution of the effective processes  $((\Theta^{\text{eff},(1),[N]}(Ns))_{s>0})_{N \in \mathbb{N}}$ .
5. Evolution of the 1-blocks in the Meyer-Zheng topology.
6. Proof of Proposition 10.2.

**Step 1: Tightness of the 1-block estimators.****Lemma 10.4.** [Tightness of the 1-block estimator] *Let*

$$(10.36) \quad (\Theta^{\text{eff.}(1),[N]}(Ns))_{s>0}$$

be defined as in (10.14). Then  $(\mathcal{L}[(\Theta^{\text{eff.}(1),[N]}(Ns))_{s>0}])_{N \in \mathbb{N}}$  is a tight sequence of probability measures on  $\mathcal{C}((0, \infty), [0, 1]^2)$ .

*Proof.* To prove tightness of  $((\Theta^{\text{eff.}(1),[N]}(Ns))_{s>0})_{N \in \mathbb{N}}$ , we will prove for all  $\epsilon > 0$  that the set of measures  $((\Theta^{\text{eff.}(1),[N]}(Ns))_{s \geq \epsilon})_{N \in \mathbb{N}}$  is tight. To do so, fix  $\epsilon > 0$ . We will again use [JM86, Proposition 3.2.3]. From (10.4) we find that the 1-block averages  $(\Theta^{\text{eff.}(1),[N]}(Ns))_{s>0}$  evolve according to

$$(10.37) \quad \begin{aligned} d\bar{\Theta}^{(1),[N]}(Ns) &= \frac{1}{1+K_0} \left[ \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Ns))} dw_i(s) \right. \\ &\quad \left. + K_1 e_1 \left[ \Theta_{y_1}^{(1),[N]} - \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Ns) \right] ds \right], \\ d\Theta_{y_1}^{(1),[N]}(Ns) &= e_1 \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Ns) - \Theta_{y_1}^{(1),[N]}(Ns) \right] ds. \end{aligned}$$

To use [JM86, Proposition 3.2.3], we define  $\mathcal{C}^*$  as the set of polynomials on  $([0, 1]^2)$ . Note that  $(\Theta^{\text{eff.}(1),[N]}(Ns))_{s \geq \epsilon}$  is a semi-martingale. Applying Itô's formula, we get

$$(10.38) \quad \begin{aligned} &f\left(\Theta^{\text{eff.}(1),[N]}(Ns)\right) \\ &= f\left(\Theta^{\text{eff.}(1),[N]}(N\epsilon)\right) \\ &\quad + \int_{\epsilon}^s dw_i(r) \frac{1}{1+K_0} \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Nr))} \frac{\partial f}{\partial x}\left(\Theta^{\text{eff.}(1),[N]}(Nr)\right) \\ &\quad + \int_{\epsilon}^s dr \frac{K_1 e_1}{1+K_0} \left[ \Theta_{y_1}^{(1),[N]}(Nr) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Nr) \right] \frac{\partial f}{\partial x}\left(\Theta^{\text{eff.}(1),[N]}(Nr)\right) \\ &\quad + \int_{\epsilon}^s dr e_1 \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Nr) - \Theta_{y_1}^{(1),[N]}(Nr) \right] \frac{\partial f}{\partial y}\left(\Theta^{\text{eff.}(1),[N]}(Nr)\right) \\ &\quad + \int_{\epsilon}^s dr \frac{1}{2(1+K_0)^2} \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Nr)) \frac{\partial^2 f}{\partial x^2}\left(\Theta^{\text{eff.}(1),[N]}(Nr)\right) \end{aligned}$$

for all  $f \in \mathcal{C}^*$ . Hence, if we define the operator

$$(10.39) \quad \begin{aligned} G_{\dagger}^{(1),[N]}: (\mathcal{C}^*, [0, 1]^2, [\epsilon, \infty), \Omega) &\rightarrow \mathbb{R}, \\ G_{\dagger}^{(1),[N]}(f, (x, y), s, \omega) &= \frac{K_1 e_1}{1+K_0} \left[ y - \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Ns, \omega) \right] \frac{\partial f}{\partial x} \\ &\quad + e_1 \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(Ns, \omega) - y \right] \frac{\partial f}{\partial y} \\ &\quad + \frac{1}{2(1+K_0)^2} \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N]}(Ns, \omega)) \frac{\partial^2 f}{\partial x^2}, \end{aligned}$$

then we see that the process  $(\Theta^{\text{eff.}(1),[N]}(Ns))_{s \geq \epsilon}$  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 [JM86, Proposition 3.2.3] that the sequence  $((\Theta^{\text{eff.}(1),[N]}(Ns))_{s \geq \epsilon})_{N \in \mathbb{N}}$  is tight. Since this is true for all  $\epsilon > 0$ , we conclude that  $(\mathcal{L}[(\Theta^{\text{eff.}(1),[N]}(Ns))_{s>0}])_{N \in \mathbb{N}}$  is tight.  $\square$

**Step 2: Stability of the 1-block estimators.**

**Lemma 10.5. [Stability property of the 1-block estimator]** *Let  $\Theta^{\text{eff},(1),[N]}(t)$  be defined as in (10.14). For any  $L(N)$  satisfying  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ ,*

$$(10.40) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N]}(Ns) - \bar{\Theta}^{(1),[N]}(Ns-t) \right| = 0 \text{ in probability}$$

and

$$(10.41) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_1}^{(1),[N]}(Ns) - \Theta_{y_1}^{(1),[N]}(Ns-t) \right| = 0 \text{ in probability.}$$

*Proof.* Fix  $\epsilon > 0$ . From the SSDE (10.4) we obtain that, for  $N$  large enough,

$$(10.42) \quad \begin{aligned} & \mathbb{P} \left( \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N]}(Ns) - \bar{\Theta}^{(1),[N]}(Ns-t) \right| > \epsilon \right) \\ &= \mathbb{P} \left( \sup_{0 \leq t \leq L(N)} \frac{1}{1+K_0} \left| \int_{Ns-t}^{Ns} dr \frac{K_1 e_1}{N} \left[ \Theta_{y_1}^{(1),[N]}(r) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(r) \right] \right. \right. \\ & \quad \left. \left. + \int_{Ns-t}^{Ns} dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N]}(r))} \right| > \epsilon \right) \\ &\leq \mathbb{P} \left( \left| \frac{L(N) 2K_1 e_1}{N(1+K_0)} \right| + \sup_{0 \leq t \leq L(N)} \left| \frac{1}{1+K_0} \int_{Ns-t}^{Ns} dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N]}(r))} \right| > \epsilon \right) \\ &= \mathbb{P} \left( \sup_{0 \leq t \leq L(N)} \left| \frac{1}{1+K_0} \int_{Ns-t}^{Ns} dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N]}(r))} \right| > \epsilon - \frac{L(N) 2K_1 e_1}{N(1+K_0)} \right) \\ &\leq \mathbb{P} \left( \sup_{0 \leq t \leq L(N)} \left| \frac{1}{1+K_0} \int_{Ns-t}^{Ns} dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N]}(r))} \right| > \frac{\epsilon}{2} \right). \end{aligned}$$

Applying the same optional stopping argument as used in the proof of Lemma 8.15, we find (10.40). For (10.41), note that

$$(10.43) \quad \begin{aligned} & \mathbb{P} \left( \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_1}^{(1),[N]}(Ns) - \Theta_{y_2}^{(1),[N]}(Ns-t) \right| > \epsilon \right) \\ &= \mathbb{P} \left( \sup_{0 \leq t \leq L(N)} \frac{1}{1+K_0} \left| \int_{Ns-t}^{Ns} dr \frac{e_1}{N} \left[ \Theta_{y_1}^{(1),[N]}(r) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N]}(r) \right] \right| > \epsilon \right) \\ &\leq \mathbb{P} \left( \frac{2e_1 L(N)}{(1+K_0)N} > \epsilon \right). \end{aligned}$$

Let  $N \rightarrow \infty$  to obtain (10.41). □

**Step 3: Equilibrium for the infinite system.** To derive the equilibrium of the single components in the infinite system, we derive the following analogue of Proposition 8.4. Recall that the finite system is denoted by  $Z^{[N_k]}$  in (10.3), and recall the list of ingredients in Section 10.1.

**Proposition 10.6. [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  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , and suppose*

that

$$\begin{aligned}
(10.44) \quad & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \Theta^{\text{eff},(1),[N_k]}(N_k s) \right] = P_{\Theta^{\text{eff}}(s)}, \\
& \lim_{k \rightarrow \infty} \mathcal{L} \left[ Y_1^{[N_k]}(N_k s) \middle| \Theta^{\text{eff},(1),[N_k]}(N_k s) \right] = P_{Y_1(s)}^{\Theta^{\text{eff},(1)}(s)}, \\
& \lim_{k \rightarrow \infty} \mathcal{L} \left[ \sup_{0 \leq t \leq L(N_k)} \left| \bar{\Theta}^{[N_k]}(N_k s) - \bar{\Theta}^{[N_k]}(N_k s - t) \right| + \left| \Theta_{y_1}^{[N_k]}(N_k s) - \Theta_{y_1}^{[N_k]}(N_k s - t) \right| \right] \\
& \quad = \delta_0, \\
& \lim_{k \rightarrow \infty} \mathcal{L} [Z^{[N_k]}(N_k s)] = \nu(s).
\end{aligned}$$

Then  $\nu(s)$  is of the form

$$(10.45) \quad \nu(s) = \int_{[0,1]^2} P_{\Theta^{\text{eff}}(s)}(d\theta, d\theta_y) \int_{[0,1]^{\mathbb{N}_0}} P_{Y_1(s)}^{(\theta, \theta_y)}(d\mathbf{y}_1) \nu_{\theta, \mathbf{y}_1},$$

where  $\mathbf{y}_1 = (y_{i,1})_{i \in \mathbb{N}_0}$  is a sequence with elements in  $[0, 1]$ , and  $\nu_{\theta, \mathbf{y}_1}$  is the equilibrium measure of the process in (10.10) evolving according to (10.11) with  $(y_{i,1})_{i \in \mathbb{N}_0}$  given by the sequence  $\mathbf{y}_1$ .

**Preparation for the proof of Proposition 10.6.** The proof of Proposition 10.6 follows the same line of argument as used in the proof of Proposition 8.4. We need lemmas that are similar to Lemmas 8.5-8.11, but this time in the setting of the two-colour hierarchical mean-field finite-systems scheme. Afterwards we prove Proposition 10.6.

**Lemma 10.7. [Convergence for the infinite system]** *Let  $\mu$  be an exchangeable probability measure on  $([0, 1]^3)^{\mathbb{N}_0}$ . Then for the system  $(Z(t))_{t \geq 0}$  given by (10.10) with  $\mathcal{L}[Z(0)] = \mu$ ,*

$$(10.46) \quad \lim_{t \rightarrow \infty} \mathcal{L}[Z(t)] = \nu_{\theta, \mathbf{y}_1},$$

where  $\nu_{\theta, \mathbf{y}_1}$  is of the form

$$(10.47) \quad \nu_{\theta, \mathbf{y}_1} = \prod_{i \in \mathbb{N}_0} \Gamma_{\theta, y_{i,1}}$$

with  $\Gamma_{\theta, y_{i,1}}$  the equilibrium of the  $i$ th single-component process in (10.11).

*Proof.* For each component of the infinite system in (10.10) the 1-dormant single component process  $(y_{i,1}(t))_{t \geq 0}$  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 7.2 (see Section 7.3) 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 10.7 follows.  $\square$

The second lemma establishes the continuity of the equilibrium with respect to  $\theta$ , its center of drift.

**Lemma 10.8. [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  $[0, 1] \times [0, 1]^{\mathbb{N}_0} \rightarrow \mathcal{P}([0, 1]^3)^{\mathbb{N}_0}$  given by*

$$(10.48) \quad (\theta, \mathbf{y}_1) \mapsto \nu_{\theta, \mathbf{y}_1}$$

is continuous. Furthermore, if  $h$  is a Lipschitz function on  $[0, 1]$ , then also  $\mathcal{F}h$  defined by

$$(10.49) \quad (\mathcal{F}h)(\theta) = \mathbb{E}^{\nu_{\theta, \mathbf{y}_1}} [h(\cdot)] = \int_{([0,1]^3)^{\mathbb{N}_0}} \nu_{\theta, \mathbf{y}_1}(dz) h(x_0)$$

is a Lipschitz function on  $[0, 1]$ , whose values are independent of  $\mathbf{y}_1$ .

*Proof.* Lemma 10.8 follows from the proof of Lemma 10.12.  $\square$

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  $N \rightarrow \infty$  and  $t \rightarrow \infty$ .

**Lemma 10.9. [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 (10.14). Then*

$$(10.50) \quad \mathbb{E} \left[ \left| \Theta_x^{(1),[N]}(t) - \Theta_{y_0}^{(1),[N]}(t) \right| \right] \leq \sqrt{\mathbb{E} \left[ \left( \Theta_x^{(1),[N]}(0) - \Theta_{y_0}^{(1),[N]}(0) \right)^2 \right]} e^{-(K_0 e_0 + e_0)t} + \sqrt{\frac{2}{K_0 e_0 + e_0} \left[ \frac{\|g\|}{N} + \frac{4K_1 e_1}{N} \right]}.$$

*Proof.* From (10.4) it follows via Itô-calculus that

$$(10.51) \quad \frac{d}{dt} \mathbb{E} \left[ \left( \Theta_x^{(1),[N]}(t) - \Theta_{y_0}^{(1),[N]}(t) \right)^2 \right] = -2(K_0 e_0 + e_0) \mathbb{E} \left[ \left( \Theta_x^{(1),[N]}(t) - \Theta_{y_0}^{(1),[N]}(t) \right)^2 \right] + h^{[N]}(t),$$

where

$$(10.52) \quad h^{[N]}(t) = \mathbb{E} \left[ \frac{2K_1 e_1}{N} \left( \Theta_x^{(1),[N]}(t) - \Theta_{y_0}^{(1),[N]}(t) \right) \left[ \Theta_{y_1}^{(1),[N]}(t) - \Theta_x^{(1),[N]}(t) \right] \right] + \frac{2}{N^2} \sum_{i \in [N]} \mathbb{E} [g(x_i^{[N]}(r))].$$

Hence

$$(10.53) \quad \mathbb{E} \left[ \left( \Theta_x^{(1),[N]}(t) - \Theta_{y_0}^{(1),[N]}(t) \right)^2 \right] = \mathbb{E} \left[ \left( \Theta_x^{(1),[N]}(0) - \Theta_{y_0}^{(1),[N]}(0) \right)^2 \right] e^{-2(K_0 e_0 + e_0)t} + \int_0^t dr e^{-2(K_0 e_0 + e_0)(t-r)} h^{[N]}(r).$$

Take the square root on both sides and use Jensen's inequality to get (10.50).  $\square$

Like for the mean-field system with one colour, we need to compare the finite system in (10.3) with an infinite system. To derive the analogue of Lemma 8.9, let  $L(N)$  satisfy  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . Define the measure  $\mu_N$  on  $([0, 1]^3)^{\mathbb{N}_0}$  by continuing the configuration of

$$(10.54) \quad Z^{[N]}(Ns - L(N)) = \left( X^{[N]}(Ns - L(N)), \left( Y_0^{[N]}(Ns - L(N)), Y_1^{[N]}(Ns - L(N)) \right) \right)$$

periodically to  $([0, 1]^3)^{\mathbb{N}_0}$ . Let

$$(10.55) \quad \bar{\Theta}^{(1),[N]} = \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N]}(Ns - L(N)) + K_0 y_{i,0}^{[N]}(Ns - L(N))}{1 + K_0}.$$

Let

$$(10.56) \quad (Z^{\mu_N}(t))_{t \geq 0} = (X^{\mu_N}(t), (Y_0^{\mu_N}(t), Y_1^{\mu_N}(t)))_{t \geq 0}$$

be the infinite system evolving according to

$$(10.57) \quad \begin{aligned} dx_i^{\mu_N}(t) &= c_0 [\bar{\Theta}^{(1),[N]} - x_i^{\mu_N}(t)] dt + \sqrt{g(x_i^{\mu_N}(t))} dw_i(t) + K_0 e_0 [y_{i,0}^{\mu_N}(t) - x_i^{\mu_N}(t)] dt, \\ dy_{i,0}^{\mu_N}(t) &= e_0 [x_i^{\mu_N}(t) - y_{i,0}^{\mu_N}(t)] dt, \\ y_{i,1}^{\mu_N}(t) &= y_{i,1}^{\mu_N}(0), \quad i \in \mathbb{N}_0, \end{aligned}$$

starting from initial distribution  $\mu_N$ . Then the following lemma is the equivalent of Lemma 8.9 for the two-colour mean-field system.

**Lemma 10.10. [Comparison of finite and infinite systems]** Fix  $s > 0$ , and let  $L(N)$  satisfy  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N$ . Suppose that

$$(10.58) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N]}(Ns) - \bar{\Theta}^{(1),[N]}(Ns-t) \right| = 0 \quad \text{in probability.}$$

Then, for all  $t \geq 0$ ,

$$(10.59) \quad \lim_{k \rightarrow \infty} \left| \mathbb{E} \left[ f(Z^{\mu_N}(t)) - f(Z^{[N]}(Ns - L(N) + t)) \right] \right| = 0 \quad \forall f \in \mathcal{C}([0,1]^3)^{\mathbb{N}_0}, \mathbb{R}.$$

*Proof.* We proceed as in the proof of Lemma 8.9. We rewrite the SSDE in (10.4) as

$$(10.60) \quad \begin{aligned} dx_i^{[N]}(t) &= c_0 [\Theta^{(1),[N]} - x_i^{[N]}(t)] dt \\ &\quad + c_0 [\bar{\Theta}^{(1),[N]}(t) - \Theta^{(1),[N]}] dt + c_0 [\Theta_x^{(1),[N]}(t) - \bar{\Theta}^{(1),[N]}(t)] dt \\ &\quad + \sqrt{g(x_i^{[N]}(t))} dw_i(t) \\ &\quad + K_0 e_0 [y_{i,0}^{[N]}(t) - x_i^{[N]}(t)] dt + \frac{K_1 e_1}{N} [y_{i,1}^{[N]}(t) - x_i^{[N]}(t)] dt, \\ dy_{i,0}^{[N]}(t) &= e_0 [x_i^{[N]}(t) - y_{i,0}^{[N]}(t)] dt, \\ dy_{i,1}^{[N]}(t) &= \frac{e_1}{N} [x_i^{[N]}(t) - y_{i,1}^{[N]}(t)] dt, \quad i \in [N]. \end{aligned}$$

As before, we consider the finite system in (10.60) as a system on  $([0,1]^3)^{\mathbb{N}_0}$  by periodic continuation, and we couple the finite system in (10.60) and the infinite system in (10.59) via these 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

$$(10.61) \quad \begin{aligned} \Delta_{i,0}^{[N]}(t) &= \tilde{x}_i^{[N]}(t) - \tilde{x}_i^{\mu_N}(t), \\ \delta_{i,0}^{[N]}(t) &= \tilde{y}_{i,0}^{[N]}(t) - \tilde{y}_{i,0}^{\mu_N}(t), \\ \delta_{i,1}^{[N]}(t) &= \tilde{y}_{i,1}^{[N]}(t) - \tilde{y}_{i,1}^{\mu_N}(t). \end{aligned}$$

As in the proof of Lemma 8.9, we have to show that, for all  $t \geq 0$ ,

$$(10.62) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^{[N]}(t)|] = 0, \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\delta_{i,0}^{[N]}(t)|] = 0, \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\delta_{i,1}^{[N]}(t)|] = 0.$$

To prove the third limit in (10.63), note that, by (10.57), (10.60) and the choice of the initial measure in the coupling,

$$(10.63) \quad y_{i,1}^{[N]}(t) = y_{i,1}^{[N]}(0) + \frac{e_1}{N} \int_0^t dr [x_i^{[N]}(r) - y_{i,1}^{[N]}(r)] = y_{i,1}^{\mu_N}(t) + \frac{e_1}{N} \int_0^t dr [x_i^{[N]}(r) - y_{i,1}^{[N]}(r)].$$

Hence

$$(10.64) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\delta_{i,1}^{[N]}(L(N))|] = 0.$$

To prove the first two limits in (10.63), we argue as in the proof of Lemma 8.9, but we need to add extra drift terms towards the first seed-bank. Using Itô-calculus, we obtain

$$(10.65) \quad \begin{aligned} \frac{d}{dt} \mathbb{E}[|\Delta_i^{[N]}(t)| + K|\delta_{i,0}^{[N]}(t)|] &= -c \mathbb{E}[\Delta_i^{[N]}(t)] \\ &\quad - 2K_0 e_0 \mathbb{E} \left[ (|\Delta_i^{[N]}(t)| + |\delta_{i,0}^{[N]}(t)|) \mathbf{1}_{\{\text{sign} \Delta_i^{[N]}(t) \neq \text{sign} \delta_{i,0}^{[N]}(t)\}} \right] \\ &\quad + c \text{sign} \Delta_i^{[N]}(t) [\bar{\Theta}^{(1),[N]}(t) - \bar{\Theta}^{(1),[N]}] \\ &\quad + c \text{sign} \Delta_i^{[N]}(t) [\bar{\Theta}_x^{(1),[N]}(t) - \bar{\Theta}^{(1),[N]}(t)] \\ &\quad + \frac{K_1 e_1}{N} \text{sign} \Delta_i^{[N]}(t) [\delta_{i,1}^{[N]}(t) - \Delta_i^{[N]}(t)]. \end{aligned}$$

This can be rewritten as

$$\begin{aligned}
(10.66) \quad & 0 \leq \mathbb{E}[|\Delta_i^{[N]}(t)| + K_0 |\delta_{i,0}^{[N]}(t)|] \\
& \leq \mathbb{E}[|\Delta_i^{[N]}(0)| + K |\delta_{i,0}^{[N]}(0)|] - c \int_0^t dr \mathbb{E}[\Delta_i^{[N]}(r)] \\
& \quad - 2K_0 e_0 \int_0^t dr \mathbb{E} \left[ (|\Delta_i^{[N]}(r)| + |\delta_{i,0}^{[N]}(r)|) \mathbf{1}_{\{\text{sign} \Delta_i^{[N]}(t) \neq \text{sign} \delta_{i,0}^{[N]}(t)\}} \right] \\
& \quad + c \int_0^t dr |\bar{\Theta}^{(1),[N]}(r) - \Theta^{(1),[N]}| \\
& \quad + c \int_0^t dr |\bar{\Theta}_x^{(1),[N]}(r) - \bar{\Theta}^{(1),[N]}(r)| \\
& \quad + \frac{K_1 e_1}{N} \int_0^t dr \left| \delta_{i,1}^{[N]}(r) - \Delta_i^{[N]}(r) \right|.
\end{aligned}$$

By the construction of the measure  $\mu_N$ , we have

$$(10.67) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^{[N]}(0)| + K_0 |\delta_{i,0}^{[N]}(0)|] = 0.$$

Therefore, for all  $t \geq 0$ ,

$$(10.68) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\Delta_i^{[N]}(t)| + K_0 |\delta_{i,0}^{[N]}(t)|] = 0.$$

Combine this with (10.64) and use that Lipschitz functions are dense in the set of bounded continuous functions. Then, as in the proof of Lemma 8.9, we get the claim in (10.59).  $\square$

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  $N \rightarrow \infty$ , we need the following regularity property for the estimators  $(\bar{\Theta}^{[N]}, \Theta_{y_1}^{[N]})$ .

**Lemma 10.11. [Stability of the estimator for the conserved quantity]** *Define  $\mu_N$  as in Lemma 10.10. Let  $(x_i^N, y_{i,0}^N, y_{i,1}^N)_{i \in [N]}$  be distributed according to the exchangeable probability measure  $\mu_N$  on  $([0, 1]^3)^{N_0}$  restricted to  $([0, 1]^3)^{[N]}$ . Suppose that  $\lim_{N \rightarrow \infty} \mu_N = \mu$  for some exchangeable probability measure  $\mu$  on  $([0, 1]^3)^{N_0}$ . Define the random variable  $\phi$  on  $(\mu, ([0, 1]^3)^{N_0})$  by putting*

$$\begin{aligned}
(10.69) \quad & \phi = (\phi_1, \phi_2), \\
& \phi_1 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i + K y_{i,0}}{1 + K}, \quad \phi_2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} y_{i,1},
\end{aligned}$$

and the random variable  $\phi^{[N]}$  on  $(\mu_N, ([0, 1]^3)^{N_0})$  by putting

$$\begin{aligned}
(10.70) \quad & \phi^{[N]} = (\phi_1^{[N]}, \phi_2^{[N]}) \\
& \phi_1^{[N]} = \frac{1}{N} \sum_{i \in [N]} \frac{x_i^N + K y_{i,0}^N}{1 + K}, \quad \phi_2^{[N]} = \frac{1}{N} \sum_{i \in [N]} y_{i,1}^N.
\end{aligned}$$

Then

$$(10.71) \quad \lim_{N \rightarrow \infty} \mathcal{L}[\phi^{[N]}] = \mathcal{L}[\phi].$$

*Proof.* We can use a similar argument as in the proof of Lemma 8.10. Define

$$(10.72) \quad D^{[N]}(Z) = \left( \frac{1}{N} \sum_{j \in [N]} \frac{x_j + K_0 y_{j,0}}{1 + K_0}, \frac{1}{N} \sum_{j \in [N]} y_{j,1} \right).$$

Then we can proceed as in the proof of Lemma 8.10, using Fourier analysis for both components of  $D^{[N]}(Z)$  separately.  $\square$



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  $N \rightarrow \infty$ .

**Lemma 10.12. [Uniformity of the ergodic theorem for the infinite system]** *Let  $\mu_N$  be defined as in (10.56). 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  $\mu = \lim_{k \rightarrow \infty} \mu_{N_k}$ . Define*

$$(10.73) \quad \Theta = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} \frac{x_i^\mu + K y_{i,0}^\mu}{1 + K} \quad \text{in } L_2(\mu).$$

Let  $Z^\mu(t) = (X^\mu(t), Y_0^\mu(t), Y_1^\mu(t))_{t \geq 0}$  be the infinite system evolving according to

$$(10.74) \quad \begin{aligned} dx_i^\mu(t) &= c[\Theta - x_i^\mu(t)] dt + \sqrt{g(x_i^\mu(t))} dw_i(t) + Ke[y_{i,1}^\mu(t) - x_i^\mu(t)] dt, \\ dy_{i,0}^\mu(t) &= e[x_i^\mu(t) - y_{i,1}^\mu(t)] dt, \\ dy_{i,1}^\mu(t) &= y_{i,1}^\mu(0), \quad i \in \mathbb{N}_0. \end{aligned}$$

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 (10.56). Then

1. For all  $t \geq 0$ ,

$$(10.75) \quad \lim_{k \rightarrow \infty} |\mathbb{E}[f(Z^{\mu_{N_k}}(t))] - \mathbb{E}[f(Z^\mu(t))]| = 0 \quad \forall f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}.$$

2. There exists a sequence  $\bar{L}(N)$  satisfying  $\lim_{N \rightarrow \infty} \bar{L}(N) = \infty$  and  $\lim_{N \rightarrow \infty} \bar{L}(N)/N = 0$  such that

$$(10.76) \quad \begin{aligned} \lim_{k \rightarrow \infty} &|\mathbb{E}[f(Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))) - f(Z^{\mu_{N_k}}(\bar{L}(N_k)))] \\ &+ |\mathbb{E}[f(Z^{\mu_{N_k}}(\bar{L}(N_k)))] - \mathbb{E}[f(Z^\mu(\bar{L}(N_k)))]| = 0 \quad \forall f \in \mathcal{C}([0, 1]^2)^{\mathbb{N}_0}, \mathbb{R}. \end{aligned}$$

*Proof.* As in the proof of Lemma 8.11, 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

$$(10.77) \quad \lim_{N \rightarrow \infty} y_{i,1}^{\mu_N}(0) = y_{i,1}^\mu(0) \text{ a.s.}$$

and

$$(10.78) \quad \lim_{N \rightarrow \infty} \mathbb{E}[|\bar{\Theta}^{[N]} - \Theta|] = 0.$$

Via a similar coupling as in Lemma (10.10), it follows via Itô-calculus that (10.75) holds. Combining (10.64), (10.68), (10.77) and (10.78), we obtain, via a similar construction as in the proof of Lemma 8.11, a sequence  $\bar{L}(N)$  such that

$$(10.79) \quad \begin{aligned} \lim_{N \rightarrow \infty} &\mathbb{E}[|\Delta_i^N(\bar{L}(N))| + K_0 |\delta_{i,0}^N(\bar{L}(N))| + K_1 |\delta_{i,1}^N(\bar{L}(N))|] \\ &+ \mathbb{E}[|\Delta_i^{\mu_N}(\bar{L}(N))| + K_0 |\delta_{i,0}^{\mu_N}(\bar{L}(N))| + K_1 |\delta_{i,1}^{\mu_N}(\bar{L}(N))|] = 0. \end{aligned}$$

As in the proof of Lemma 8.11, we can again use Lipschitz functions to conclude (10.76).  $\square$

**Lemma 10.13. [Coupling of finite systems]** *Let*

$$(10.80) \quad Z^{[N],1} = (X^{[N],1}, Y_0^{[N],1}, Y_1^{[N],1})$$

*be the finite system evolving according to (10.4) 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*

$$(10.81) \quad Z^{[N],2} = (X^{[N],2}, Y_0^{[N],2}, Y_1^{[N],2})$$

be the finite system evolving according to (10.4) 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  $(([0, 1]^3)^{\mathbb{N}_0} \times ([0, 1]^3)^{\mathbb{N}_0}, \tilde{\mu})$  by

$$(10.82) \quad \begin{aligned} \bar{\Theta}^{[N],1} &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N],1} + K_0 y_{i,0}^{[N],1}}{1 + K_0}, & \bar{\Theta}^{[N],2} &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N],2} + K_0 y_{i,0}^{[N],2}}{1 + K_0}, \\ \bar{\Theta}^1 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i^1 + K_0 y_{i,0}^1}{1 + K_0}, & \bar{\Theta}^2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i^2 + K_0 y_{i,0}^2}{1 + K_0}, \end{aligned}$$

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 (10.14) for  $Z^{[N],1}$ , respectively,  $Z^{[N],2}$ . Suppose that

$$(10.83) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left( \left| \bar{\Theta}^{[N],k}(0) - \bar{\Theta}^{[N],k}(t) \right| \right) = 0 \text{ in probability, } k \in \{1, 2\},$$

and suppose that  $\tilde{\mu}(\{\bar{\Theta}_1 = \bar{\Theta}_2, Y_1^1 = Y_1^2\}) = 1$ . Then, for any  $t(N) \rightarrow \infty$ ,

$$(10.84) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} [ & |x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K_0 |y_{i,0}^{[N],1}(t(N)) - y_{i,0}^{[N],2}(t(N))| \\ & + K_1 |y_{i,1}^{[N],1}(t(N)) - y_{i,1}^{[N],2}(t(N))| ] = 0. \end{aligned}$$

*Proof.* Via standard Itô-calculus we obtain from (10.4) that

$$(10.85) \quad \begin{aligned} & \frac{d}{dt} \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K_0 |y_{i,0}^{[N],1}(t) - y_{i,0}^{[N],2}(t)| + K_1 |y_{i,1}^{[N],1}(t) - y_{i,1}^{[N],2}(t)| \right] \\ &= -\frac{2c}{N} \sum_{j \in [N]} \mathbb{E} \left[ |x_j^{[N],1}(t) - x_j^{[N],2}(t)| \mathbf{1}_{\{\text{sign}(x_j^{[N],1}(t) - x_j^{[N],2}(t)) \neq \text{sign}(x_i^{[N],1}(t) - x_i^{[N],2}(t))\}} \right] \\ & \quad - 2K_0 e_0 \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K |y_{i,0}^{[N],1}(t) - y_{i,0}^{[N],2}(t)| \right. \\ & \quad \left. \times \mathbf{1}_{\{\text{sign}(x_i^{[N],1}(t) - x_i^{[N],2}(t)) \neq \text{sign}(y_{i,0}^{[N],1}(t) - y_{i,0}^{[N],2}(t))\}} \right] \\ & \quad - 2\frac{K_1 e_1}{N} \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K_1 |y_{i,0}^{[N],1}(t) - y_{i,0}^{[N],2}(t)| \right. \\ & \quad \left. \times \mathbf{1}_{\{\text{sign}(x_i^{[N],1}(t) - x_i^{[N],2}(t)) \neq \text{sign}(y_{i,1}^{[N],1}(t) - y_{i,1}^{[N],2}(t))\}} \right]. \end{aligned}$$

Therefore, for all  $N \in \mathbb{N}$ ,

$$(10.86) \quad t \mapsto \mathbb{E} \left[ |x_i^{[N],1}(t) - x_i^{[N],2}(t)| + K_0 |y_{i,0}^{[N],1}(t) - y_{i,0}^{[N],2}(t)| + K_1 |y_{i,1}^{[N],1}(t) - y_{i,1}^{[N],2}(t)| \right]$$

is a decreasing function. Hence we can use the same strategy as in the proof of Lemma 8.13 to finish the proof.  $\square$

### • Proof of Proposition 10.6

*Proof.* We follow a similar argument as in the proof of Proposition 8.4. Let  $L(N)$  satisfy  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . Let  $\mu_N$  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 10.6, we can pass to a possibly further subsequence and obtain

$$(10.87) \quad \lim_{k \rightarrow \infty} \mu_{N_k} = \mu.$$

Since we assumed that  $\mathcal{L}[Z^{[N]}(0)]$  is exchangeable and the dynamics preserve exchangeability, the measures  $\mu_{N_k}$  are translation invariant and also the limiting law  $\mu$  is translation invariant.

Let  $\phi = (\phi_1, \phi_2)$  be defined as in (10.69) in Lemma 10.11. Then we can condition on  $\phi = (\phi_1, \phi_2)$  and write

$$(10.88) \quad \mu = \int_{[0,1]^2} \mu_\rho \, d\Lambda(\rho),$$

where  $\Lambda(\cdot) = \mathcal{L}[\phi] = \mathcal{L}[(\phi_1, \phi_2)]$  and  $\rho = (\rho_1, \rho_2)$ . By assumption we know that

$$(10.89) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \Theta^{\text{eff},(1),[N_k]}(N_k s) \right] = P_{\Theta^{\text{eff}}(s)}(\cdot)$$

and

$$(10.90) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \sup_{0 \leq t \leq L(N_k)} \left| \bar{\Theta}^{[N_k]}(N_k s) - \bar{\Theta}^{[N_k]}(N_k s - t) \right| + \left| \Theta_{y_1}^{[N_k]}(N_k s) - \Theta_{y_1}^{[N_k]}(N_k s - t) \right| \right] = \delta_0.$$

Hence

$$(10.91) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \Theta^{\text{eff},(1),[N_k]}(N_k s - L(N_k)) \right] = P_{\Theta^{\text{eff}}(s)}(\cdot).$$

Recall that

$$(10.92) \quad \Lambda(\cdot) = \mathcal{L} \left[ \lim_{n \rightarrow \infty} \left( \frac{1}{n} \sum_{i \in [n]} \frac{x_i + K y_{i,0}}{1 + K}, \frac{1}{n} \sum_{i \in [n]} y_{i,1} \right) \right] \quad \text{on } (\mu, ([0, 1]^2)^{\mathbb{N}_0}).$$

By Lemma 8.10, if

$$(10.93) \quad \phi^{N_k} = (\phi_1^{N_k}, \phi_2^{N_k}) = \left( \frac{1}{N_k} \sum_{i \in [N_k]} \frac{x_i + K y_{i,0}}{1 + K}, \frac{1}{N_k} \sum_{i \in [N_k]} y_{i,1}^{[N_k]} \right) \quad \text{on } (\mu_{N_k}, ([0, 1]^3)^{\mathbb{N}_0}),$$

then  $\lim_{k \rightarrow \infty} \mathcal{L}[\phi^{N_k}] = \mathcal{L}[\phi]$ . Taking the subsequence  $(\mu_{N_k})_{k \in \mathbb{N}}$ , we get  $\Lambda(\cdot) = P_{\Theta^{\text{eff}}(s)}(\cdot)$ , and hence

$$(10.94) \quad \mu = \int_{[0,1]} \mu_\rho \, dP_s(\rho).$$

Let  $\bar{L}(N)$  be the sequence constructed in Lemma 10.12[b]. By construction we can require that  $\bar{L}(N) \leq L(N)$  for all  $N \in \mathbb{N}$ . Write

$$(10.95) \quad \begin{aligned} & \mathcal{L} [Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))] \\ &= \mathcal{L} [Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))] - \mathcal{L} [Z^{\mu_{N_k}}(\bar{L}(N_k))], \\ & \quad + \mathcal{L} [Z^{\mu_{N_k}}(\bar{L}(N_k))] - \mathcal{L} [Z^\mu(\bar{L}(N_k))] \\ & \quad + \mathcal{L} [Z^\mu(\bar{L}(N_k))]. \end{aligned}$$

By Lemma 10.12 the first and second differences tend to zero as  $k \rightarrow \infty$ . Hence

$$(10.96) \quad \lim_{k \rightarrow \infty} \mathcal{L} [Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))] = \mathcal{L} [Z^\mu(\bar{L}(N_k))].$$

By (10.88),

$$(10.97) \quad \mathcal{L} [Z^\mu(\bar{L}(N_k))] = \int_{[0,1]^2} \mathcal{L} [Z^{\mu_\rho}(\bar{L}(N_k))] P_{\Theta^{\text{eff}}(s)}(d\rho).$$

For the infinite system  $(Z^{\mu_\rho}(t))_{t \geq 0} = (X^{\mu_\rho}(t), Y_0^{\mu_\rho}(t), Y_1^{\mu_\rho}(t))_{t \geq 0}$  we have

$$(10.98) \quad Y_1^\mu(t) = Y_1^\mu(0) \text{ a.s.}$$

and hence, since  $\lim_{k \rightarrow \infty} \bar{L}(N_k)/N_k = 0$  by (10.44),

$$(10.99) \quad \lim_{k \rightarrow \infty} \mathcal{L}[Y_1^{\mu^\rho}(\bar{L}(N_k))] = \mathcal{L}[Y_1^{\mu^\rho}(0)] \quad \forall \rho \in [0, 1].$$

Therefore

$$(10.100) \quad \lim_{k \rightarrow \infty} \mathcal{L}[Y_1^{\mu^\rho}(\bar{L}(N_k))] = P_{Y_1(s)}^\rho(\cdot)$$

and

$$(10.101) \quad \begin{aligned} & \mathcal{L}[X^{\mu^\rho}(\bar{L}(N_k)), Y_0^{\mu^\rho}(\bar{L}(N_k)), Y_1^{\mu^\rho}(\bar{L}(N_k))] \\ &= \int \mathcal{L}[X_1^{\mu^\rho}(\bar{L}(N_k)), Y_0^{\mu^\rho}(\bar{L}(N_k)), \mathbf{y}_1] dP_{Y_1(s)}^\rho(d\mathbf{y}_1). \end{aligned}$$

Hence, since  $\lim_{k \rightarrow \infty} \bar{L}(N_k) = \infty$ , by Lemma 8.5 we have

$$(10.102) \quad \begin{aligned} \lim_{k \rightarrow \infty} \mathcal{L}[Z^{\mu^\rho}(\bar{L}(N_k))] &= \lim_{k \rightarrow \infty} \mathcal{L}[X_1^{\mu^\rho}(\bar{L}(N_k)), Y_0^{\mu^\rho}(\bar{L}(N_k)), Y_1^{\mu^\rho}(\bar{L}(N_k))] \\ &= \int \nu_{\rho, \mathbf{y}_1} P_{Y_1(s)}^\rho(d\mathbf{y}_1). \end{aligned}$$

Therefore, by (9.8), (10.97) and Lemma 8.6,

$$(10.103) \quad \lim_{k \rightarrow \infty} \mathcal{L}[Z^{[N_k]}(N_k s - L(N_k) + \bar{L}(N_k))] = \int_{[0,1]} P_{\Theta^{\text{eff}}(s)}(d\rho) \int \nu_{\rho, \mathbf{y}_1} P_{Y_1(s)}^\rho(d\mathbf{y}_1).$$

To finish the proof, we proceed as in the proof of Proposition 8.4 and invoke Lemma 10.13. Let  $Z^{[N],1} = (X^{[N],1}, Y_0^{[N],1}, Y_1^{[N],1})$  be the finite system starting from

$$(10.104) \quad \mathcal{L}[Z^{[N]}(Ns - L(N))] = \mathcal{L}[X^{[N]}(Ns - L(N)), Y_0^{[N]}(Ns - L(N)), Y_1^{[N]}(Ns - L(N))].$$

Let  $(\bar{L}(N))_{N \in \mathbb{N}}$  be the sequence constructed in Lemma 10.12. Let  $Z^{[N],2} = (X^{[N],2}, Y_0^{[N],2}, Y_1^{[N],2})$  be the finite system starting from

$$(10.105) \quad \mathcal{L}[X^{[N]}(Ns - \bar{L}(N))] = \mathcal{L}[X^{[N]}(Ns - \bar{L}(N)), Y_0^{[N]}(Ns - \bar{L}(N)), Y_1^{[N]}(Ns - \bar{L}(N))].$$

Choose for  $t(N)$  in Lemma 10.13 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  $\Theta_1$  and  $\Theta_2$  according to (8.44), where under  $\mu^{[N],2}$  we replace  $L(N)$  by  $\bar{L}(N)$ . Then, by the assumptions in (10.44),

$$(10.106) \quad \begin{aligned} \lim_{k \rightarrow \infty} |\Theta^{(1),[N_k],1} - \Theta^{(1),[N_k],2}| &= \lim_{k \rightarrow \infty} |\Theta^{N_k}(N_k s - L(N_k)) - \Theta^{N_k}(N_k s - \bar{L}(N_k))| \\ &= 0 \text{ in probability.} \end{aligned}$$

Using 10.63 we see that also, for all  $i \in [N]$ ,

$$(10.107) \quad \begin{aligned} \lim_{k \rightarrow \infty} |y_{i,1}^{[N_k],1}(0) - y_{i,1}^{[N_k],2}(0)| &= \lim_{k \rightarrow \infty} |y_{i,1}^{[N_k]}(N_k s - L(N_k)) - y_{i,1}^{[N_k]}(N_k s - \bar{L}(N_k))| \\ &= 0 \text{ in probability.} \end{aligned}$$

Therefore, if  $\mu$  is any weak limit point of the sequence  $(\mu^{[N_k],1} \times \mu^{[N_k],2})_{k \in \mathbb{N}}$ , then

$$(10.108) \quad \mu(\{\Theta_1 = \Theta_2, Y_1^1 = Y_1^2\}) = 1.$$

Hence, by possibly passing to a further subsequence, we can now apply Lemma 10.13 to obtain, for all  $i$ ,

$$(10.109) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathbb{E} \left[ |x_i^{[N_k],1}(\bar{L}(N_k)) - x_i^{[N_k],2}(\bar{L}(N_k))| \right. \\ & \quad + K_0 |y_{i,0}^{[N_k],1}(\bar{L}(N_k)) - y_{i,0}^{[N_k],2}(\bar{L}(N_k))| \\ & \quad \left. + K_1 |y_{i,1}^{[N_k],1}(\bar{L}(N_k)) - y_{i,1}^{[N_k],2}(\bar{L}(N_k))| \right] = 0. \end{aligned}$$

Hence

$$(10.110) \quad \lim_{N \rightarrow \infty} \left( \mathcal{L}[Z^{[N],1}(\bar{L}(N_k))] - \mathcal{L}[Z^{[N],2}(\bar{L}(N_k))] \right) = \delta_0$$

and therefore

$$(10.111) \quad \lim_{k \rightarrow \infty} \mathcal{L}(Z^{[N_k]}(N_k s)) = \int_{[0,1]} P_{\Theta^{\text{eff}}(s)}(d\rho) \int \nu_{\rho, \mathbf{y}_1} P_{Y_1(s)}^\rho(d\mathbf{y}_1).$$

This concludes the proof of Proposition 10.6.  $\square$

Like for the one-colour mean-field system, Proposition 10.6 and Lemmas 10.7–10.13 give rise to the following corollary, which will be important to derive the evolution of the 1-blocks on time scale  $Ns$ .

**Corollary 10.14.** *Fix  $s > 0$ . Let  $\mu_N$  be the measure obtained by periodic continuation of*

$$(10.112) \quad Z^{[N]}(Ns - L(N)) = (X^{[N]}(Ns - L(N)), Y_0^{[N]}(Ns - L(N)), Y_1^{[N]}(Ns - L(N))),$$

and let  $\mu$  be a weak limit point of the sequence  $(\mu_N)_{N \in \mathbb{N}}$ . Let

$$(10.113) \quad \Theta = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} \frac{x_i^\mu + K y_i^\mu}{1 + K} \quad \text{in } L_2(\mu),$$

and let  $(Z^{\nu\Theta}(t))_{t \geq 0} = (X^{\nu\Theta}(t), Y_0^{\nu\Theta}(t), Y_1^{\nu\Theta}(t))_{t \geq 0}$  be the infinite system evolving according to (10.74) 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  $(Z^{[N]}(t))_{t \geq 0}$  and  $(Z^{\nu\Theta}(t))_{t \geq 0}$  on one probability space. Then, for all  $t \geq 0$ ,

$$(10.114) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| x_i^{[N]}(Ns + t) - x_i^{\nu\Theta}(t) \right| \right] + K_0 \mathbb{E} \left[ \left| y_{i,0}^{[N]}(Ns + t) - y_{i,0}^{\nu\Theta}(t) \right| \right] \\ + K_1 \mathbb{E} \left[ \left| y_{i,1}^{[N]}(Ns + t) - y_{i,1}^{\nu\Theta}(t) \right| \right] = 0 \quad \forall i \in [N].$$

*Proof.* Proceed as in the proof of Corollary 9.1, but use the setup of the two-colour mean-field system and therefore replace Proposition 8.4, Lemma 8.11 and Lemma 8.13 by, respectively Proposition 10.6, Lemma 10.12 and Lemma 10.13.  $\square$

#### Step 4: Limiting evolution of the 1-blocks.

**Lemma 10.15. [Limiting evolution of the 1-blocks]** *Let  $(z_1^{\text{eff}}(s))_{s > 0}$  be the process defined in (10.21) with initial state*

$$(10.115) \quad z_1^{\text{eff}}(0) = (\vartheta_0, \theta_{y_1}).$$

Then

$$(10.116) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ (\Theta^{\text{eff},(1),[N]}(Ns))_{s > 0} \right] = \mathcal{L} \left[ (z_1^{\text{eff}}(s))_{s > 0} \right].$$

*Proof.* By [YW71], the SSDE in (10.21) has a unique strong solution. Therefore the process  $(z_1^{\text{eff}}(s))_{s > 0}$  is Markov. Its generator  $G$  is given by

$$(10.117) \quad G = \frac{K_1 e_1}{1 + K_0} (y - x) \frac{\partial}{\partial x} + e_1 (x - y) \frac{\partial}{\partial y} + \frac{1}{(1 + K_0)^2} (\mathcal{F}g)(x) \frac{\partial^2}{\partial x^2},$$

and hence  $(z_1^{\text{eff}}(s))_{s \geq 0}$  solves the martingale problem for  $G$ . We will use [JM86, Theorem 3.3.1], to prove that (10.116) holds.

Define

$$(10.118) \quad (\vartheta_0^N, \vartheta_{y_1}^N) = \left( \bar{\Theta}^{(1),[N]}(0), \Theta_{y_1}^{(1),[N]}(0) \right).$$

Since we start from an i.i.d. law, by the law of large numbers we have that

$$(10.119) \quad \lim_{N \rightarrow \infty} \Theta^{\text{eff},(1),[N]}(0) = \lim_{N \rightarrow \infty} (\vartheta_0^N, \vartheta_{y_1}^N) = (\vartheta_0, \theta_{y_1}) \quad a.s.$$

By the SSDE in (10.37) and an optional sampling argument, we have, for all  $N \in \mathbb{N}$ ,

$$(10.120) \quad \lim_{s \downarrow 0} \left( \bar{\Theta}^{(1),[N]}(Ns), \Theta^{(1),[N]}(Ns) \right) = (\vartheta_0^N, \vartheta_{y_1}^N) \quad a.s.$$

Therefore we can continuously extend the process  $(\Theta^{\text{eff},(1),[N]}(Ns))_{s>0}$  to 0 and, in particular,

$$(10.121) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \Theta^{\text{eff},(1),[N]}(0) \right] = \mathcal{L} [z_1^{\text{eff}}(0)].$$

Since we already showed that the processes

$$(10.122) \quad (\Theta^{\text{eff},(1),[N]}(Ns))_{s>0}$$

are  $\mathcal{D}$ -semimartingales, and are trivially bounded, we are left to show that

$$(10.123) \quad \lim_{N \rightarrow \infty} \int_0^s dr \mathbb{E} \left[ \left| G_{\dagger}^{(1),[N]}(f, \Theta^{\text{eff},(1),[N]}(Nr), r, \cdot) - (Gf)(\Theta^{\text{eff},(1),[N]}(Nr)) \right| \right] = 0.$$

Here,  $G_{\dagger}^{(1),[N]}$  is the operator defined in (10.39). 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  $s > 0$ ,

$$(10.124) \quad \lim_{N \rightarrow \infty} \mathbb{E}^{[N]} \left[ \left| G_{\dagger}^{(1),[N]}(f, \Theta^{\text{eff},(1),[N]}(Ns), s, \cdot) - (Gf)(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right] = 0.$$

Note that

$$(10.125) \quad \begin{aligned} & \mathbb{E} \left[ \left| G_{\dagger}^{(1),[N]}(f, \Theta^{\text{eff},(1),[N]}(Ns), s, \cdot) - (Gf)(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right] \\ &= \mathbb{E} \left[ \left| \frac{K_1 e_1}{1 + K_0} \left[ \Theta_{y_1}^{(1),[N]}(Ns) - \frac{1}{N} \sum_{i \in [N]} x_i(Ns, \omega) \right] \frac{\partial f}{\partial x}(\Theta^{\text{eff},(1),[N]}(Ns)) \right. \right. \\ & \quad + e_1 \left[ \frac{1}{N} \sum_{i \in [N]} x_i(Ns, \omega) - \Theta_{y_1}^{(1),[N]}(Ns) \right] \frac{\partial f}{\partial y}(\Theta^{\text{eff},(1),[N]}(Ns)) \\ & \quad + \frac{1}{(1 + K_0)^2} \frac{1}{N} \sum_{i \in [N]} g(x_i(Ns, \omega)) \frac{\partial^2 f}{\partial x^2}(\Theta^{\text{eff},(1),[N]}(Ns)) \\ & \quad - \frac{K_1 e_1}{1 + K_0} \left[ \Theta_{y_1}^{(1),[N]}(Ns) - \bar{\Theta}^{(1),[N]}(Ns) \right] \frac{\partial f}{\partial x}(\Theta^{\text{eff},(1),[N]}(Ns)) \\ & \quad - e_1 \left[ \bar{\Theta}^{(1),[N]}(Ns) - \Theta_{y_1}^{(1),[N]}(Ns) \right] \frac{\partial f}{\partial y}(\Theta^{\text{eff},(1),[N]}(Ns)) \\ & \quad \left. \left. - \frac{1}{(1 + K_0)^2} (\mathcal{F}g)(\bar{\Theta}^{(1),[N]}(Ns)) \frac{\partial^2 f}{\partial x^2}(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right]. \end{aligned}$$

Hence

$$(10.126) \quad \begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| G_{\dagger}^{(1),[N]}(f, \Theta^{\text{eff},(1),[N]}(Ns), s, \cdot) - (Gf)(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right] \\ & \leq \lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{K_1 e_1}{1 + K_0} \left| \bar{\Theta}^{(1),[N]}(Ns) - \frac{1}{N} \sum_{i \in [N]} x_i(Ns, \omega) \right| \left| \frac{\partial f}{\partial x}(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right] \\ & \quad + \lim_{N \rightarrow \infty} \mathbb{E} \left[ e_1 \left| \frac{1}{N} \sum_{i \in [N]} x_i(Ns, \omega) - \bar{\Theta}^{(1),[N]}(Ns) \right| \left| \frac{\partial f}{\partial y}(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right] \\ & \quad + \lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{1}{(1 + K_0)^2} \left| \frac{1}{N} \sum_{i \in [N]} g(x_i(Ns, \omega)) - (\mathcal{F}g)(\bar{\Theta}^{(1),[N]}(Ns)) \right| \left| \frac{\partial^2 f}{\partial x^2}(\Theta^{\text{eff},(1),[N]}(Ns)) \right| \right]. \end{aligned}$$

Note that each of the derivatives is bounded by a constant because we work on  $\mathcal{C}^*$ . The first and the second term tend to zero by Lemma 10.9. For the third term we can use a similar argument as used in (9.97), since we showed Lemmas 10.7–10.13 for the single components in the mean-field system with two colours.  $\square$

**Step 5: Evolution of the averages in the Meyer-Zheng topology.** In this section we prove the following proposition

**Proposition 10.16. [Convergence in the Meyer-Zheng topology]** *Suppose that the effective estimator process defined in (10.15) satisfies*

$$(10.127) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff},(1),[N]}(Ns) \right)_{s>0} \right] = \mathcal{L} \left[ \left( z_1^{\text{eff}}(s) \right)_{s>0} \right].$$

Then for the averages in (10.7),

$$(10.128) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( z_1^{[N]}(s) \right)_{s>0} \right] = \mathcal{L} \left[ \left( z_1(s) \right)_{s>0} \right]$$

*in the Meyer-Zheng topology,*

where the limiting process  $(z_1(s))_{s>0}$  is defined as in (10.20).

To prove Proposition 10.16 we need the following characterisation of continuous functions in the Meyer-Zheng topology

**Lemma 10.17. [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*

$$(10.129) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( X_n(s), Y_n(s) \right)_{s>0} \right] = \mathcal{L} \left[ \left( X(s), Y(s) \right)_{s>0} \right] \text{ in the Meyer-Zheng topology,}$$

then the marginals also converge in the Meyer-Zheng topology, i.e.,

$$(10.130) \quad \begin{aligned} \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( X_n(s) \right)_{s>0} \right] &= \mathcal{L} \left[ \left( X(s) \right)_{s>0} \right] \text{ in the Meyer-Zheng topology,} \\ \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( Y_n(s) \right)_{s>0} \right] &= \mathcal{L} \left[ \left( Y(s) \right)_{s>0} \right] \text{ in the Meyer-Zheng topology.} \end{aligned}$$

The proof of Lemma 10.17 is given in Appendix B.3.

*Proof of Proposition 10.16.* By Lemma 10.9, we have that, for all  $s > 0$ ,

$$(10.131) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{[N]}(Ns) - x_1^{[N]}(s) \right| \right] = 0$$

and

$$(10.132) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{[N]}(Ns) - y_{0,1}^{[N]}(s) \right| \right] = 0.$$

Applying Lemmas 8.19, 8.20 and 8.21, like in the proof of Proposition 8.18, we obtain

$$(10.133) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( x_1^{[N]}(s), y_{0,1}^{[N]}(s), \bar{\Theta}^{[N]}(Ns), \Theta_{y_{1,1}}^{[N]}(Ns) \right)_{s>0} \right] \\ = \mathcal{L} \left[ \left( x_1^{\text{eff}}(s), x_1^{\text{eff}}(s), x_1^{\text{eff}}(s), y_1^{\text{eff}}(s) \right)_{s>0} \right] \text{ in the Meyer-Zheng topology.} \end{aligned}$$

Applying Lemma 10.17, we get the claim.  $\square$

**Step 6: Proof of the two-colour mean-field finite-systems scheme.**

*Proof.* The proof of Proposition 10.2(a) follows directly from Lemma 10.15. The proof of Proposition 10.2(b) is a consequence of Proposition 10.2(d). The proof of Proposition 10.2(c) follows from Proposition 10.2(a) by applying Proposition 10.16. The proof of Proposition (10.2)(d) follows by the same argument as used in the proof Proposition 8.1(c) in Section 9.4. 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 10.11 instead of the infinite system defined in (8.1). We now use the two-dimensional transition kernel in (10.30), 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 (8.18).  $\square$

**11 Preparation:  $N \rightarrow \infty$ , two-level three-colour**

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 11.1 we give the set-up of the two-level three-colour mean-field model. In Section 11.2 we give a scheme to prove the analysis of the two-level three-colour mean-field model. Finally, in Section 11.3 we prove the steps of the scheme given in Section 11.2.

**11.1 Two-level three-colour mean-field finite-systems scheme**

We consider a restricted version of the SSDE in (2.20) on the finite geographic space

$$(11.1) \quad [N^2] = \{0, 1, \dots, N^2 - 1\}, \quad N \in \mathbb{N}.$$

This space should be interpreted as grouping the  $N$ -blocks consisting of  $N$  colonies together, i.e.

$$(11.2) \quad [N^2] = \bigcup_{l=0}^{N-1} \{Nl, Nl+1, \dots, Nl+N-1\}.$$

With this interpretation we can use the metric  $d_{[N^2]}$  that is induced by the metric  $d_{\Omega_N}$  on the hierarchical group  $\Omega_N$  (recall (2.4)). 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

$$(11.3) \quad a^{[N^2]}(i, j) = 1_{\{d_{[N^2]}(i, j) \leq 1\}} \frac{c_0}{N} + \frac{c_1}{N^3},$$

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

$$(11.4) \quad S = \mathfrak{s}^{[N^2]}, \quad \mathfrak{s} = [0, 1] \times [0, 1]^3,$$

and the restricted system is denoted by

$$(11.5) \quad \begin{aligned} (Z^{[N^2]}(t))_{t \geq 0} &= \left( X^{[N^2]}(t), \left( Y_0^{[N^2]}(t), Y_1^{[N^2]}(t), Y_2^{[N^2]}(t) \right) \right)_{t \geq 0}, \\ \left( X^{[N^2]}(t), \left( Y_0^{[N^2]}(t), Y_1^{[N^2]}(t), Y_2^{[N^2]}(t) \right) \right) &= \left( x_i^{[N^2]}(t), \left( y_{i,0}^{[N^2]}(t), y_{i,1}^{[N^2]}(t), y_{i,2}^{[N^2]}(t) \right) \right)_{i \in [N^2]}. \end{aligned}$$



The components of the restricted system  $(Z^{[N^2]}(t))_{t \geq 0}$  evolve according to the SSDE

$$\begin{aligned}
dx_i^{[N^2]}(t) &= \frac{c_0}{N} \sum_{j \in [N^2]} 1_{\{d_{[N^2]}(i,j) \leq 1\}} [x_j^{[N^2]}(t) - x_i^{[N^2]}(t)] dt \\
&\quad + \frac{c_1}{N^3} \sum_{j \in [N^2]} [x_j^{[N^2]}(t) - x_i^{[N^2]}(t)] dt + \sqrt{g(x_i^{[N^2]}(t))} dw_i(t) \\
&\quad + K_0 e_0 [y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt \\
(11.6) \quad &\quad + \frac{K_1 e_1}{N} [y_{i,1}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt \\
&\quad + \frac{K_2 e_2}{N^2} [y_{i,2}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt, \\
dy_{i,0}^{[N^2]}(t) &= e_0 [x_i^{[N^2]}(t) - y_{i,0}^{[N^2]}(t)] dt, \\
dy_{i,1}^{[N^2]}(t) &= \frac{e_1}{N} [x_i^{[N^2]}(t) - y_{i,1}^{[N^2]}(t)] dt, \\
dy_{i,2}^{[N^2]}(t) &= \frac{e_2}{N^2} [x_i^{[N^2]}(t) - y_{i,2}^{[N^2]}(t)] dt, \quad i \in [N^2],
\end{aligned}$$

which is a special case of (2.20). By [SS80, Theorem 3.1], the SSDE in (11.6) is the unique solution. It is important to note that we can write the SSDE also

$$\begin{aligned}
dx_i^{[N^2]}(t) &= c_0 \left[ \frac{1}{N} \sum_{j \in [N]_i} x_j^{[N^2]}(t) - x_i^{[N^2]}(t) \right] dt \\
&\quad + \frac{c_1}{N} \left[ \frac{1}{N^2} \sum_{j \in [N^2]} x_j^{[N^2]}(t) - x_i^{[N^2]}(t) \right] dt + \sqrt{g(x_i^{[N^2]}(t))} dw_i(t) \\
(11.7) \quad &\quad + K_0 e_0 [y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt + \frac{K_1 e_1}{N} [y_{i,1}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt \\
&\quad + \frac{K_2 e_2}{N^2} [y_{i,2}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt, \\
dy_{i,0}^{[N^2]}(t) &= e_0 [x_i^{[N^2]}(t) - y_{i,0}^{[N^2]}(t)] dt, \\
dy_{i,1}^{[N^2]}(t) &= \frac{e_1}{N} [x_i^{[N^2]}(t) - y_{i,1}^{[N^2]}(t)] dt, \\
dy_{i,2}^{[N^2]}(t) &= \frac{e_2}{N^2} [x_i^{[N^2]}(t) - y_{i,2}^{[N^2]}(t)] dt, \quad i \in [N^2],
\end{aligned}$$

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 (11.7) 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 (11.7).

• **On time scale 1 =  $N^0$**  (i.e., space-time scale 0) in the limit as  $N \rightarrow \infty$ , the colour-1 dormant population and the colour-2 dormant population do not yet move. Hence

$$(11.8) \quad (y_{1,0}^{[N^2]}(t_0), y_{2,0}^{[N^2]}(t_0))_{t_0 \geq 0},$$

converges as  $N \rightarrow \infty$  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 (7.1), 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

$$(11.9) \quad (z_0^{[N^2]}(t_0))_{t_0 \geq 0} = (x_0^{[N^2]}(t_0), y_{0,0}^{[N^2]}(t_0))_{t_0 \geq 0},$$

will be called *the single colony effective process*.

• **On time scale  $N^1$**  (i.e., space-time scale 1), we look at the averages

$$(11.10) \quad \begin{aligned} (z_1^{[N^2]}(t_1))_{t_1 > 0} &= \left( x_1^{[N^2]}(t_1), \left( y_{0,1}^{[N^2]}(t_1), y_{1,1}^{[N^2]}(t_1), y_{2,1}^{[N^2]}(t_1) \right) \right)_{t_1 > 0} \\ &= \left( \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt_1), \left( \frac{1}{N} \sum_{i \in [N]} y_{i,0}^{[N^2]}(Nt_1), \frac{1}{N} \sum_{i \in [N]} y_{i,1}^{[N^2]}(Nt_1), \frac{1}{N} \sum_{i \in [N]} y_{i,2}^{[N^2]}(Nt_1) \right) \right)_{t_1 > 0}. \end{aligned}$$

(Recall Remark 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 (2.22). Using (11.6), we see that the dynamics of the system in (11.10) is given by the SSDE

$$(11.11) \quad \begin{aligned} dx_1^{[N^2]}(t_1) &= c_1 \left[ \frac{1}{N^2} \sum_{j \in [N^2]} x_j(Nt_1) - x_1(t_1) \right] dt_1 + \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i(Nt_1))} dw(t_1) \\ &\quad + NK_0 e_0 \left[ y_{0,1}^{[N^2]}(t_1) - x_1^{[N^2]}(t_1) \right] dt_1 \\ &\quad + K_1 e_1 \left[ y_{1,1}^{[N^2]}(t_1) - x_1^{[N^2]}(t_1) \right] dt_1 \\ &\quad + \frac{K_2 e_2}{N} \left[ y_{2,1}^{[N^2]}(t_1) - x_1^{[N^2]}(t_1) \right] dt_1, \\ dy_{0,1}^{[N^2]}(t_1) &= Ne_0 \left[ x_1^{[N^2]}(t_1) - y_{0,1}^{[N^2]}(t_1) \right] dt_1, \\ dy_{1,1}^{[N^2]}(t_1) &= e_1 \left[ x_1^{[N^2]}(t_1) - y_{1,1}^{[N^2]}(t_1) \right] dt_1, \\ dy_{2,1}^{[N^2]}(t_1) &= \frac{e_2}{N} \left[ x_1^{[N^2]}(t_1) - y_{1,1}^{[N^2]}(t_1) \right] dt_1. \end{aligned}$$

In the limit  $N \rightarrow \infty$  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  $N \rightarrow \infty$ , an infinitely strong interaction with the active population. Therefore we expect that, in the limit as  $N \rightarrow \infty$ , its path becomes rougher and rougher at rarer and rarer times. We will need to use the *Meyer-Zheng topology* to prove that

$$(11.12) \quad \lim_{N \rightarrow \infty} y_{0,1}^{[N^2]}(t_1) = \lim_{N \rightarrow \infty} x_1^{[N^2]}(t_1) \text{ for most } t_1.$$

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  $N \rightarrow \infty$ , 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 (2.22),

$$(11.13) \quad \begin{aligned} & \left( x_2^{[N^2]}(t_2), (y_{0,2}^{[N^2]}(t_2), y_{1,2}^{[N^2]}(t_2), y_{2,2}^{[N^2]}(t_2)) \right)_{t_2 > 0} \\ &= \left( \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(N^2 t_2), \right. \\ & \quad \left. \left( \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,0}^{[N^2]}(N^2 t_2), \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,1}^{[N^2]}(N^2 t_2), \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}(N^2 t_2) \right) \right)_{t_2 > 0}, \end{aligned}$$

which evolves according to the SSDE

$$(11.14) \quad \begin{aligned} dx_2^{[N^2]}(t_2) &= \sqrt{\frac{1}{N^2} \sum_{i \in [N^2]} g(x_i(N^2 t_2))} dw(t_2) \\ & \quad + N^2 K_0 e_0 \left[ y_{0,2}^{[N^2]}(t_2) - x_2^{[N^2]}(t_2) \right] dt_2 \\ & \quad + N K_1 e_1 \left[ y_{1,2}^{[N^2]}(t_2) - x_2^{[N^2]}(t_2) \right] dt_2 \\ & \quad + K_2 e_2 \left[ y_{2,2}^{[N^2]}(t_2) - x_2^{[N^2]}(t_2) \right] dt_2, \\ dy_{0,2}^{[N^2]}(t_2) &= N^2 e_0 \left[ x_2^{[N^2]}(t_2) - y_{0,2}^{[N^2]}(t_2) \right] dt_2, \\ dy_{1,2}^{[N^2]}(t_2) &= N e_1 \left[ x_2^{[N^2]}(t_2) - y_{1,2}^{[N^2]}(t_2) \right] dt_2, \\ dy_{2,2}^{[N^2]}(t_2) &= e_2 \left[ x_2^{[N^2]}(t_2) - y_{2,2}^{[N^2]}(t_2) \right] dt_2. \end{aligned}$$

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  $N \rightarrow \infty$ , 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  $N \rightarrow \infty$ , 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

$$(11.15) \quad \frac{x_2^{[N^2]}(t_2) + K_0 y_{0,2}^{[N^2]}(t_2) + K_1 y_{1,2}^{[N^2]}(t_2)}{1 + K_0 + K_1},$$

for which we find

$$(11.16) \quad \begin{aligned} & d \left[ \frac{x_2^{[N^2]}(t_2) + K_0 y_{0,2}^{[N^2]}(t_2) + K_1 y_{1,2}^{[N^2]}(t_2)}{1 + K_0 + K_1} \right] \\ &= \frac{1}{1 + K_0 + K_1} \sqrt{\frac{1}{N^2} \sum_{i \in [N^2]} g(x_i(N^2 t_2))} dw(t_2) + K_2 e_2 \left[ y_{2,2}^{[N^2]}(t_2) - x_2^{[N^2]}(t_2) \right] dt_2, \end{aligned}$$

we see that the infinite rates cancel out. We will call

$$(11.17) \quad \left( \frac{x_2^{[N^2]}(t_2) + K_0 y_{0,2}^{[N^2]}(t_2) + K_1 y_{1,2}^{[N^2]}(t_2)}{1 + K_0 + K_1}, y_{2,2}^{[N^2]}(t_2) \right)_{t_2 > 0}$$

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

$$(11.18) \quad \begin{aligned} dx_0(t) &= c_0 [\theta - x_0(t)] dt + \sqrt{g(x_0(t))} dw(t) \\ &\quad + K_0 e_0 [y_{0,0}(t) - x_0(t)] dt, \\ dy_{0,0}(t) &= e_0 [x_0(t) - y_{0,0}(t)] dt, \\ y_{1,0}(t) &= y_{1,0}, \\ y_{2,0}(t) &= y_{2,0}, \end{aligned}$$

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

$$(11.19) \quad \begin{aligned} dx_0^{\text{eff}}(t) &= c_0 [\theta - x_0^{\text{eff}}(t)] dt + \sqrt{g(x_0^{\text{eff}}(t))} dw(t) + K_0 e_0 [y_{0,0}^{\text{eff}}(t) - x_0^{\text{eff}}(t)] dt, \\ dy_{0,0}^{\text{eff}}(t) &= e_0 [x_0^{\text{eff}}(t) - y_{0,0}^{\text{eff}}(t)] dt, \quad i \in \mathbb{N}_0, \end{aligned}$$

where  $\theta \in [0, 1]$ . By [YW71], (11.18) and (11.19) 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:

1. For positive times  $t > 0$ , we define the following *1-block estimators* for the finite system:

$$(11.20) \quad \begin{aligned} \bar{\Theta}^{(1),[N^2]}(t) &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N^2]}(t) + K_0 y_{i,0}^{[N^2]}(t)}{1 + K_0}, \\ \Theta_x^{(1),[N^2]}(t) &= \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(t), \\ \Theta_{y_0}^{(1),[N^2]}(t) &= \frac{1}{N} \sum_{i \in [N]} y_{i,0}^{[N^2]}(t), \\ \Theta_{y_1}^{(1),[N^2]}(t) &= \frac{1}{N} \sum_{i \in [N]} y_{i,1}^{[N^2]}(t), \\ \Theta_{y_2}^{(1),[N^2]}(t) &= \frac{1}{N} \sum_{i \in [N]} y_{i,2}^{[N^2]}(t). \end{aligned}$$

We abbreviate

$$(11.21) \quad \begin{aligned} \Theta^{(1),[N^2]}(t) &= \left( \Theta_x^{(1),[N^2]}(t), \left( \Theta_{y_0}^{(1),[N^2]}(t), \Theta_{y_1}^{(1),[N^2]}(t), \Theta_{y_2}^{(1),[N^2]}(t) \right) \right), \\ \Theta^{\text{aux},(1),[N^2]}(t) &= \left( \bar{\Theta}^{(1),[N^2]}(t), \Theta_{y_1}^{(1),[N^2]}(t), \Theta_{y_2}^{(1),[N^2]}(t) \right), \\ \Theta^{\text{eff},(1),[N^2]}(t) &= \left( \bar{\Theta}^{(1),[N^2]}(t), \Theta_{y_1}^{(1),[N^2]}(t) \right). \end{aligned}$$

We call  $(\Theta^{(1),[N^2]}(t))_{t > 0}$  the *1-block estimator process*,  $(\Theta^{\text{aux},(1),[N^2]}(t))_{t > 0}$  the *auxiliary 1-block estimator process* and  $(\Theta^{\text{eff},(1),[N^2]}(t))_{t > 0}$  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:

$$\begin{aligned}
\bar{\Theta}^{(2),[N^2]}(t) &= \frac{1}{N^2} \sum_{i \in [N^2]} \frac{x_i^{[N^2]}(t) + K_0 y_{i,0}^{[N^2]}(t) + K_1 y_{i,1}^{[N^2]}(t)}{1 + K_0 + K_1}, \\
\Theta_x^{(2),[N^2]}(t) &= \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(t), \\
(11.22) \quad \Theta_{y_0}^{(2),[N^2]}(t) &= \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,0}^{[N^2]}(t), \\
\Theta_{y_1}^{(2),[N^2]}(t) &= \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,1}^{[N^2]}(t), \\
\Theta_{y_2}^{(2),[N^2]}(t) &= \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}(t).
\end{aligned}$$

We abbreviate

$$\begin{aligned}
(11.23) \quad \Theta^{(2),[N^2]}(t) &= \left( \Theta_x^{(2),[N^2]}(t), \left( \Theta_{y_0}^{(2),[N^2]}(t), \Theta_{y_1}^{(2),[N^2]}(t), \Theta_{y_2}^{(2),[N^2]}(t) \right) \right), \\
\Theta^{\text{eff},(2),[N^2]}(t) &= \left( \bar{\Theta}^{(2),[N^2]}(t), \Theta_{y_1}^{(2),[N^2]}(t) \right).
\end{aligned}$$

We call  $(\Theta^{\text{eff},(2),[N^2]}(t))_{t>0}$  the *effective 2-block estimator process* and  $(\Theta^{(2),[N^2]}(t))_{t>0}$  as the *2-block estimator process*.

2. 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  $L(N) \rightarrow \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , 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  $L(N) \rightarrow \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , 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”.

3. The *invariant measure* for the evolution of a single colony in (11.18), written

$$(11.24) \quad \Gamma_{\theta, \mathbf{y}_0}^{(0)}, \quad \mathbf{y}_0 = (\theta, y_{1,0}, y_{2,0}),$$

and the invariant measure of the level-0 effective process evolving according to (11.19), written

$$(11.25) \quad \Gamma_{\theta}^{\text{eff},(0)}.$$

4. The renormalisation transformation  $\mathcal{F}: \mathcal{G} \rightarrow \mathcal{G}$ ,

$$(11.26) \quad (\mathcal{F}g)(\theta) = \int_{[0,1]^2} g(x) \Gamma_{\theta}^{\text{eff},(0)}(dx, dy_0), \quad \theta \in [0, 1],$$

where  $\Gamma_{\theta}^{\text{eff},(0)}$  is the equilibrium measure in (11.25). Note that this is the same transformation as defined in (4.16), 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.

5. The limiting 1-block process is given by  $(z_1(t))_{t>0} = (x_1(t), (y_{0,1}(t), y_{1,1}(t), y_{2,1}(t)))_{t>0}$  and evolves according to

$$(11.27) \quad \begin{aligned} dx_1(t) &= \frac{1}{1+K_0} \left[ c_1[\theta - x_1(t)] dt + \sqrt{(\mathcal{F}^{(1)}g)(x_1(t))} dw(t) \right. \\ &\quad \left. + K_1 e_1 [y_{1,1}(t) - x_1(t)] dt \right], \\ y_{0,1}(t) &= x_1(t), \\ dy_{1,1}(t) &= e_1 [x_1(t) - y_{1,1}(t)] dt, \\ y_{2,1}(t) &= y_{2,1}, \end{aligned}$$

where  $\theta \in [0, 1]$ , and  $y_{2,1} \in [0, 1]$ , and  $\mathcal{F}^{(1)}$  is the renormalisation transformation defined in (11.26). The limiting 1-block process for the auxiliary estimator process is given by  $(z_1^{\text{aux}}(t))_{t>0} = (x_1^{\text{aux}}(t), y_{1,1}^{\text{aux}}(t), y_{2,1}^{\text{aux}}(t))_{t>0}$  and evolves according to

$$(11.28) \quad \begin{aligned} dx_1^{\text{aux}}(t) &= \frac{1}{1+K_0} \left[ c_1[\theta - x_1^{\text{aux}}(t)] dt + \sqrt{(\mathcal{F}^{(1)}g)(x_1^{\text{aux}}(t))} dw(t) \right. \\ &\quad \left. + K_1 e_1 [y_{1,1}^{\text{aux}}(t) - x_1^{\text{aux}}(t)] dt \right], \\ dy_{1,1}^{\text{aux}}(t) &= e_1 [x_1^{\text{aux}}(t) - y_{1,1}^{\text{aux}}(t)] dt, \\ y_{2,1}^{\text{aux}}(t) &= y_{2,1}, \end{aligned}$$

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  $(z_1^{\text{eff}}(t))_{t>0} = (x_1^{\text{eff}}(t), y_{1,1}^{\text{eff}}(t))_{t>0}$  and evolves according to

$$(11.29) \quad \begin{aligned} dx_1^{\text{eff}}(t) &= \frac{1}{1+K_0} \left[ c_1[\theta - x_1^{\text{eff}}(t)] dt + \sqrt{(\mathcal{F}^{(1)}g)(x_1^{\text{eff}}(t))} dw(t) \right. \\ &\quad \left. + K_1 e_1 [y_{1,1}^{\text{eff}}(t) - x_1^{\text{eff}}(t)] dt \right], \\ dy_{1,1}^{\text{eff}}(t) &= e_1 [x_1^{\text{eff}}(t) - y_{1,1}^{\text{eff}}(t)] dt, \end{aligned}$$

for  $\theta \in [0, 1]$ . By [YW71], (11.27), (11.28) and (11.29) have a unique strong solution.

6. The *invariant measure* of the infinite system in (11.27), written

$$(11.30) \quad \Gamma_{\theta, y_1}^{(1)}, \quad y_1 = (\theta, \theta, y_{2,1}),$$

and the invariant measures of the level-1 limiting estimator process evolving according to (11.28) and the level-1 effective process evolving according to (11.29),

$$(11.31) \quad \Gamma_{\theta}^{\text{aux},(1)}, \Gamma_{\theta}^{\text{eff},(1)}.$$

7. The first iteration of the renormalisation transformation,

$$(11.32) \quad (\mathcal{F}^{(2)}g)(\theta) = \int_{[0,1]^2} (\mathcal{F}g)(x) \Gamma_{\theta}^{\text{eff},(1)}(dx, dy_1), \quad \theta \in [0, 1].$$

Hence

$$(11.33) \quad (\mathcal{F}^{(2)}g)(\theta) = \int_{[0,1]^2} \Gamma_{\theta}^{\text{eff},(1)}(du, dv) \int_{[0,1]^2} g(x) \Gamma_u^{\text{eff},(0)}(dx, dy).$$

8. The limiting 2-block process  $(z_2(t))_{t>0} = (x_2(t), (y_{0,2}(t), y_{1,2}(t), y_{2,2}(t)))_{t>0}$  evolves according to

$$(11.34) \quad \begin{aligned} dx_2(t) &= \frac{1}{1 + K_0 + K_1} \left[ \sqrt{(\mathcal{F}^{(2)}g)(x_2(t))} dw(t) + K_2 e_2 [y_{2,2}(t) - x_2(t)] dt \right], \\ y_{0,2}(t) &= x_2(t), \\ dy_{1,2}(t) &= x_2(t), \\ dy_{2,2}(t) &= e_2 [x_2(t) - y_{2,2}(t)] dt, \end{aligned}$$

where  $\mathcal{F}^{(2)}g$  is defined as in (11.32). 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

$$(11.35) \quad \begin{aligned} dx_2^{\text{eff}}(t) &= \frac{1}{1 + K_0 + K_1} \left[ \sqrt{(\mathcal{F}^{(2)}g)(x_2^{\text{eff}}(t))} dw(t) + K_2 e_2 [y_{2,2}^{\text{eff}}(t) - x_2^{\text{eff}}(t)] dt \right], \\ dy_{2,2}^{\text{eff}}(t) &= e_2 [x_2^{\text{eff}}(t) - y_{2,2}^{\text{eff}}(t)] dt. \end{aligned}$$

We are now ready to state the scaling limit for the evolution of the averages in (10.7).

**Proposition 11.1. [Mean-field: two-level three-colour finite-systems scheme]** *Suppose that  $\mu(0) = \mu^{\otimes [N^2]}$  for some  $\mu \in \mathcal{P}([0, 1] \times [0, 1]^2)$ . Let*

$$(11.36) \quad \begin{aligned} \vartheta_0 &= \mathbb{E}^\mu \left[ \frac{x + K_0 y_0}{1 + K_0} \right], \quad \vartheta_1 = \mathbb{E}^\mu \left[ \frac{x + K_0 y_0 + K_1 y_1}{1 + K_0 + K_1} \right], \\ \theta_{y_1} &= \mathbb{E}^\mu [y_1], \quad \theta_{y_2} = \mathbb{E}^\mu [y_2]. \end{aligned}$$

and recall the limiting process  $(z_2(t))_{t>0}$  in (11.34) and the limiting process  $(z_1(t))_{t>0}$  in (11.27). Assume for the 2-dormant 1-blocks that

$$(11.37) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ Y_{2,1}^{[N^2]}(Nt_2) \middle| \Theta^{(2), [N^2]}(N^2 t_2) \right] = P^{z_2(t_2)},$$

and for the 2-dormant 0-blocks (= single colonies) that

$$(11.38) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ Y_{2,0}^{[N^2]}(Nt_2 + Nt_1) \middle| \Theta^{\text{eff}, (1), [N^2]}(N^2 t_2 + Nt_1) \right] = P^{z_1(t_1)}.$$

Then the following hold:

- (a) For the effective 2-block estimator process defined in (11.23),

$$(11.39) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (2), [N^2]}(N^2 t_2) \right)_{t_2 > 0} \right] = \mathcal{L} \left[ (z_2^{\text{eff}}(t_2))_{t_2 > 0} \right],$$

where the limit is determined by the unique solution of the SSDE (11.35) with initial state

$$(11.40) \quad z_2^{\text{eff}}(0) = (x_2^{\text{eff}}(0), y_2^{\text{eff}}(0)) = (\vartheta_1, \theta_{y_2}).$$

- (b) For the effective 1-block estimator process defined in (11.21),

$$(11.41) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (1), [N^2]}(N^2 t_2 + Nt_1) \right)_{t_1 > 0} \right] = \mathcal{L} \left[ (z_1^{\text{eff}}(t_1))_{t_1 > 0} \right],$$

where, conditional on  $x_2^{\text{eff}}(t_2) = u$ , the limit process is the unique solution of the SSDE in (11.29) with  $\theta$  replaced by  $u$  and with initial measure  $\Gamma_u^{\text{eff}, (1)}$ .

- (c) For the single colony effective process defined in (11.9),

$$(11.42) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( z_0^{\text{eff}, [N^2]}(N^2 t_2 + Nt_1 + t_0) \right)_{t_0 \geq 0} \right] = \mathcal{L} \left[ (z_0^{\text{eff}}(t_0))_{t_0 \geq 0} \right],$$

where, conditional on  $x_1^{\text{eff}}(t_1) = v$ , the limit process is the unique solution of the SSDE in (11.19) with  $\theta$  replaced by  $v$  and with initial measure  $\Gamma_v^{\text{eff}, (0)}$ .

(d) For the 2-block estimator process defined in (11.23),

$$(11.43) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(2), [N^2]}(N^2 t_2) \right)_{t_2 > 0} \right] = \mathcal{L} \left[ (z_2(t_2))_{t_2 > 0} \right]$$

in the Meyer-Zheng topology,

where the limit process is the unique solution of the SSDE in (11.34) with initial state

$$(11.44) \quad z_2(0) = (\vartheta_1, (\vartheta_1, \vartheta_1, \theta_{y_2})).$$

(e) Fix  $t_2 > 0$ . Assume (11.37). Define

$$(11.45) \quad \Gamma^{(1)}(t_2) = \int_{[0,1]^4} S_{t_2}^{[2]}((\vartheta_1, (\vartheta_1, \vartheta_1, \theta_{y_2})), d(u_x, u_x, u_x, u_{y_{2,2}})) \\ \int_{[0,1]} P^{(u_x, u_x, u_x, u_{y_{2,2}})}(dy_{2,1}) \Gamma_{(u_x, (u_x, u_x, y_{2,1}))}^{(1)} \in \mathcal{P}([0,1]^4),$$

where  $\Gamma_{(u_x, (u_x, u_x, y_{2,1}))}^{(1)}$  is the equilibrium measure in (11.30) 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 (11.43) 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 (11.27) 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 (11.21),

$$(11.46) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(1), [N^2]}(N^2 t_2 + N t_1) \right)_{t_1 > 0} \right] = \mathcal{L} \left[ (z^{\Gamma^{(1)}(t_2)}(t_1))_{t_1 > 0} \right]$$

in the Meyer-Zheng topology.

(f) Let  $z_1(t_1)$  be the limiting process obtained in (e). Assume (11.38). Define, for  $t_2 \in (0, \infty)$ ,

$$(11.47) \quad \Gamma^{(0)}(t_2) = \int_{[0,1]^4} \Gamma^{(1)}(t_2)(dz_1) \int_{[0,1]} P^{z_1}(dy_{2,0}) \Gamma_{(x_1, (x_1, y_{1,1}, y_{2,0}))}^{(0)},$$

where  $\Gamma^{(1)}(t_2)$  is as defined in (11.45). Let  $(z^{\Gamma^{(0)}(t_2)}(t_0))_{t_0 \geq 0}$  be the random process in (11.18) with  $z^{\Gamma^{(0)}(t_2)}(0)$  drawn according to  $\Gamma^{(0)}(t_2)$  which is a mixture of random processes in equilibrium. Then

$$(11.48) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( z_0^{[N^2]}(N^2 t_2 + N t_1 + t_0) \right)_{t_0 \geq 0} \right] = \mathcal{L} \left[ (z^{\Gamma^{(0)}(t_2)}(t_0))_{t_0 \geq 0} \right].$$

**Remark 11.2.** Note that Proposition 11.1(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 11.3.** Note that in Propostion 11.1(f)  $\Gamma_{(x_1, (x_1, y_{1,1}, y_{2,0}))}^{(0)}$  is the equilibrium measure of (11.18) (see also (11.24)), 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 equilibrium on the space-time scale where it is effective, in this case space-time-scale 1. Since we start at times  $N^2 t_2$ , 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  $N^2 t_2$  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 11.3.8 we will treat this effect in detail. ■



## 11.2 Scheme for the two-level three-colour mean-field analysis.

In this section we give a scheme to prove Proposition 11.1. The proof of the steps in the scheme will be written in Section 11.3. To analyse the two-level hierarchical mean-field system we use the results obtained in Sections 8.2, 9 and 10.2.

The scheme for the two-level three-colour hierarchical mean-field system comes in 11 steps. Recall the estimators defined in (11.20) and (11.22).

1. Tightness of the effective 2-block estimator processes

$$(11.49) \quad \left( \left( \Theta^{\text{eff},(2),[N^2]}(N^2 t_2) \right)_{t_2 > 0} \right)_{N \in \mathbb{N}} = \left( \left( \bar{\Theta}^{(2),[N^2]}(N^2 t_2), \Theta_{y_2}^{(2),[N^2]}(N^2 t_2) \right)_{t_2 > 0} \right)_{N \in \mathbb{N}}.$$

2. Stability property of the 2-block estimators, i.e., for  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ ,

$$(11.50) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \bar{\Theta}^{(2),[N^2]}(N^2 t_2 - Nt) \right| = 0 \text{ in probability}$$

and

$$(11.51) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_2}^{(2),[N^2]}(N^2 t_2) - \Theta_{y_2}^{(2),[N^2]}(N^2 t_2 - Nt) \right| = 0 \text{ in probability.}$$

3. Tightness of the effective 1-block estimator process (recall (11.21)),

$$(11.52) \quad \left( \left( \Theta^{\text{aux},(1),[N^2]}(N^2 t_2 + Nt_1) \right)_{t_1 > 0} \right)_{N \in \mathbb{N}}.$$

4. Stability property of  $(\Theta^{\text{aux},(1),[N^2]}(N^2 t_2 + Nt_1))_{t_1 > 0}$ , i.e., for  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , for all  $\epsilon > 0$ ,

$$(11.53) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N^2]}(N^2 t_2 + Nt_1) - \bar{\Theta}^{(1),[N^2]}(N^2 t_2 + Nt_1 - t) \right| = 0$$

in probability,

$$(11.54) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_1}^{(1),[N^2]}(N^2 t_2 + Nt_1) - \Theta_{y_1}^{(1),[N^2]}(N^2 t_2 + Nt_1 - t) \right| = 0$$

in probability,

$$(11.55) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_2}^{(1),[N^2]}(N^2 t_2 + Nt_1) - \Theta_{y_2}^{(1),[N^2]}(N^2 t_2 + Nt_1 - t) \right| = 0$$

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

$$(11.56) \quad \left( \left( \Theta_i^{\text{aux},(1),[N^2]}(N^2 t_2 + Nt_1) \right)_{t_1 > 0, i \in [N]} \right)_{N \in \mathbb{N}}$$

is tight. From the tightness in steps 1 and 3 we can construct a subsequence  $(N_k)_{k \in \mathbb{N}}$  along which

$$(11.57) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff},(2),[N_k^2]}(N_k^2 t_2) \right)_{t_2 > 0} \right],$$

$$\lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right)_{t_1 > 0, i \in [N_k]} \right]$$

both exists. Define the measure

$$(11.58) \quad \nu^{(0)}(t_2) = \prod_{i \in \mathbb{N}_0} \Gamma_i^{(0)}(t_2).$$

Show that along the same subsequence the single components converge to the infinite system, i.e.,

$$(11.59) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( Z^{[N_k^2]}(N_k^2 t_2 + N_k t_1 + t_0) \right)_{t_0 \geq 0} \right] = \mathcal{L} \left[ (Z^{\nu^{(0)}(t_2)}(t_0))_{t_0 \geq 0} \right].$$

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 (11.18), where  $\theta$  is now a random variable that inherits its law from

$$(11.60) \quad \lim_{k \rightarrow \infty} \mathcal{L}[(\Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{i \in [N_k^2]}],$$

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

$$(11.61) \quad \lim_{k \rightarrow \infty} \mathcal{L}[(\Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{i \in [N_k^2]}].$$

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

$$(11.62) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right)_{t_1 > 0, i \in [N_k]} \right].$$

7. Identify the limit  $\lim_{k \rightarrow \infty} \mathcal{L}[(\Theta^{\text{eff},(2),[N_k^2]}(N_k^2 t_2))_{t_2 > 0}]$  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  $N \rightarrow \infty$ , 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

$$(11.63) \quad \left( \Theta_x^{(1),[N^2]}(N^2 t_2 + N t_1), \Theta_{y_0}^{(1),[N^2]}(N^2 t_2 + N t_1), \Theta_{y_1}^{(1),[N^2]}(N^2 t_2 + N t_1), \Theta_{y_2}^{(1),[N^2]}(N^2 t_2 + N t_1) \right)_{t_1 > 0}$$

and

$$(11.64) \quad \left( \Theta_x^{(2),[N^2]}(N^2 t_2), \Theta_{y_0}^{(2),[N^2]}(N^2 t_2), \Theta_{y_1}^{(2),[N^2]}(N^2 t_2), \Theta_{y_2}^{(2),[N^2]}(N^2 t_2) \right)_{t_2 > 0}.$$

11. Combine the above steps to complete the proof of Proposition 11.1.

### 11.3 Proof of two-level three-colour mean-field finite-systems scheme

In this section we prove the steps in the scheme given in Section 11.2.

### 11.3.1 Tightness of the 2-block estimators

In this section we prove step 1 of the scheme.

**Lemma 11.4. [Tightness of the 2-block estimator]** *Let*

$$(11.65) \quad \Theta^{\text{eff},(2),[N^2]}(N^2 t_2) = (\bar{\Theta}^{(2),[N^2]}(N^2 t_2), \Theta_{y_2}^{(2),[N^2]}(N^2 t_2))$$

be defined as in (11.22). Then  $(\mathcal{L}[(\Theta^{\text{eff},(2),[N^2]}(N^2 t_2))_{t_2 > 0}])_{N \in \mathbb{N}}$  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 [JM86, Proposition 3.2.3]. From (11.6) we find that  $(\Theta^{\text{eff},(2),[N^2]}(t))_{t > 0}$  evolves according to

$$(11.66) \quad \begin{aligned} d\bar{\Theta}^{(2),[N^2]}(t) &= \frac{1}{1 + K_0 + K_1} \frac{1}{N^2} \sum_{i \in [N^2]} \sqrt{g(x_i^{[N^2]}(t))} dw_i(t) \\ &+ \frac{1}{1 + K_0 + K_1} \frac{K_2 e_2}{N^2} \left[ \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}(t) - \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(t) \right] dt, \\ d\Theta_{y_2}^{(2),[N^2]}(t) &= \frac{e_2}{N^2} \left[ \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(t) - \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}(t) \right] dt. \end{aligned}$$

Therefore the process  $(\Theta^{\text{eff},(2),[N^2]}(N^2 t_2))_{t_2 > 0}$  evolves according to

$$(11.67) \quad \begin{aligned} d\bar{\Theta}^{(2),[N^2]}(N^2 t_2) &= \frac{1}{1 + K_0 + K_1} \sqrt{\frac{1}{N^2} \sum_{i \in [N^2]} g(x_i^{[N^2]}(N^2 t_2))} dw_i(t_2) \\ &+ \frac{1}{1 + K_0 + K_1} K_2 e_2 \left[ \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}(N^2 t_2) - \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(N^2 t_2) \right] dt_2, \\ d\Theta_{y_2}^{(2),[N^2]}(N^2 t_2) &= e_2 \left[ \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(N^2 t_2) - \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}(N^2 t_2) \right] dt_2. \end{aligned}$$

To use [JM86, Proposition 3.2.3], we define  $\mathcal{C}^*$  as the set of polynomials on  $([0, 1]^2)$ . Since  $(\Theta^{\text{eff},(2),[N^2]}(N^2 t_2))_{t_2 > 0}$  is a semi-martingale, by applying Itô's formula we obtain that  $(\Theta^{\text{eff},(2),[N^2]}(N^2 t_2))_{t_2 > 0}$  is a  $\mathcal{D}$ -semi-martingale with corresponding operator

$$(11.68) \quad \begin{aligned} G_{\dagger}^{(2),[N^2]} : (\mathcal{C}^*, [0, 1]^2, (0, \infty), \Omega) &\rightarrow \mathbb{R}, \\ G_{\dagger}^{(2),[N^2]}(f, (x, y), t, \omega) &= \frac{K_2 e_2}{1 + K_0 + K_1} \left[ y - \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(N^2 t, \omega) \right] \frac{\partial f}{\partial x} \\ &+ e_2 \left[ \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(N^2 t, \omega) - y \right] \frac{\partial f}{\partial y} \\ &+ \frac{1}{2(1 + K_0 + K_1)^2} \frac{1}{N^2} \sum_{i \in [N^2]} g(x_i^{[N^2]}(N^2 t, \omega)) \frac{\partial^2 f}{\partial x^2}. \end{aligned}$$

The conditions  $H_1, H_2, H_3$  in [JM86, Proposition 3.2.3] are satisfied. Hence tightness follows from [JM86, Proposition 3.2.3].  $\square$

### 11.3.2 Stability of the 2-block estimators

**Lemma 11.5. [Stability property of the 2-block estimator]** Let  $(\Theta^{\text{eff},(2),[N^2]}(t))_{t>0}$  be defined as in (11.23). For any  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ ,

$$(11.69) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \bar{\Theta}^{(2),[N^2]}(N^2 t_2 - Nt) \right| = 0 \text{ in probability}$$

and

$$(11.70) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_2}^{(2),[N^2]}(N^2 t_2) - \Theta_{y_2}^{(2),[N^2]}(N^2 t_2 - Nt) \right| = 0 \text{ in probability.}$$

*Proof.* Fix  $\epsilon > 0$ . From the SSDE in (11.66) we obtain that, for  $N$  large enough,

$$(11.71) \quad \begin{aligned} & \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \bar{\Theta}^{(2),[N^2]}(N^2 t_2 - Nt) \right| > \epsilon \right] \\ &= \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \frac{1}{1 + K_0 + K_1} \left| \int_{N^2 t_2 - Nt}^{N^2 t_2} dw_i(r) \frac{1}{N^2} \sum_{i \in [N^2]} \sqrt{g(x_i^{[N^2]}(r))} \right. \right. \\ & \quad \left. \left. + \int_{N^2 t_2 - Nt}^{N^2 t_2} dr \frac{K_2 e_2}{N^2} \left[ \Theta_{y_2}^{(2),[N^2]}(r) - \frac{1}{N^2} \sum_{i \in [N^2]} x_i^{[N^2]}(r) \right] \right| > \epsilon \right] \\ &\leq \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \frac{1}{1 + K_0 + K_1} \left| \int_{N^2 t_2 - Nt}^{N^2 t_2} dw_i(r) \frac{1}{N^2} \sum_{i \in [N^2]} \sqrt{g(x_i^{[N^2]}(r))} \right| \right. \\ & \quad \left. > \epsilon - \frac{K_2 e_2}{1 + K_0 + K_1} \frac{L(N)N}{N^2} \right] \\ &\leq \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \frac{1}{1 + K_0 + K_1} \left| \int_{N^2 t_2 - Nt}^{N^2 t_2} dw_i(r) \frac{1}{N^2} \sum_{i \in [N^2]} \sqrt{g(x_i^{[N^2]}(r))} \right| > \frac{\epsilon}{2} \right]. \end{aligned}$$

By a similar optional stopping time argument as in the proof of Lemma 8.15, the above computation shows that (11.69) holds. Equation (11.70) holds by a similar argument as given in the proof of Lemma 10.5.  $\square$

### 11.3.3 Tightness of the 1-block estimators

**Lemma 11.6. [Tightness of the 1-block estimator]** Let

$$(11.72) \quad \begin{aligned} & \Theta^{\text{aux},(1),[N^2]}(N^2 t_2 + Nt_1) \\ &= (\bar{\Theta}^{(1),[N^2]}(N^2 t_2 + Nt_1), \Theta_{y_1}^{(1),[N^2]}(N^2 t_2 + Nt_1), \Theta_{y_2}^{(1),[N^2]}(N^2 t_2 + Nt_1)) \end{aligned}$$

be defined as in (11.20). Then  $(\mathcal{L}[(\Theta^{\text{aux},(1),[N^2]}(N^2 t_2 + Nt_1))_{t_1 > 0}])_{N \in \mathbb{N}}$  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 [JM86, Proposition 3.2.3]. From (11.11)

we find that the effective process  $(\Theta^{\text{aux},(1),[N^2]}(N^2t_2 + Nt_1))_{t_1>0}$  evolves according to

$$\begin{aligned}
d\bar{\Theta}^{(1),[N^2]}(Nt_1) &= \frac{1}{1+K_0} c_1 \left[ \frac{1}{N^2} \sum_{j \in [N^2]} x_j^{[N^2]}(Nt_1) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt_1) \right] dt_1 \\
&\quad + \frac{1}{1+K_0} \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i^{[N^2]}(Nt_1))} dw_i(t_1) \\
&\quad + \frac{K_1 e_1}{1+K_0} \left[ \Theta_{y_1}^{(1),[N^2]}(Nt_1) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt_1) \right] dt_1 \\
(11.73) \quad &\quad + \frac{K_2 e_2}{N(1+K_0)} \left[ \Theta_{y_2}^{(1),[N^2]}(Nt_1) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt_1) \right] dt_1, \\
d\Theta_{y_1}^{(1),[N^2]}(Nt_1) &= e_1 \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt_1) - \Theta_{y_1}^{(1),[N^2]}(Nt_1) \right] dt_1, \\
d\Theta_{y_2}^{(1),[N^2]}(Nt_1) &= \frac{e_2}{N} \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt_1) - \Theta_{y_2}^{(1),[N^2]}(Nt_1) \right] dt_1.
\end{aligned}$$

To use [JM86, Proposition 3.2.3], we define  $\mathcal{C}^*$  as the set of polynomials on  $([0, 1]^2)$ . Since  $(\Theta^{\text{aux},(1),[N^2]}(N^2t_2 + Nt_1))_{t_1>0}$  is a semi-martingale, by applying Itô's formula we obtain that  $(\Theta^{\text{aux},(1),[N^2]}(N^2t_2 + Nt_1))_{t_1>0}$  is a  $\mathcal{D}$ -semi-martingale with corresponding operator

$$\begin{aligned}
G_{\dagger}^{(1),[N^2]} : (\mathcal{C}^*, [0, 1]^3, (0, \infty), \Omega) &\rightarrow \mathbb{R}, \\
G_{\dagger}^{(1),[N^2]}(f, (x, y_1, y_2), t, \omega) &= \frac{c_1}{1+K_0} \left[ \frac{1}{N^2} \sum_{j \in [N^2]} x_j^{[N^2]}(Nt, \omega) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt, \omega) \right] \frac{\partial f}{\partial x} \\
&\quad + \frac{K_1 e_1}{1+K_0} \left[ y_1 - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt, \omega) \right] \frac{\partial f}{\partial x} \\
(11.74) \quad &\quad + \frac{K_2 e_2}{N(1+K_0)} \left[ y_2(Nt, \omega) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt) \right] \frac{\partial f}{\partial x} \\
&\quad + e_1 \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt, \omega) - y_1 \right] \frac{\partial f}{\partial y_1} \\
&\quad + \frac{e_2}{N} \left[ \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(Nt, \omega) - y_2(Nt, \omega) \right] \frac{\partial f}{\partial y_2} \\
&\quad + \frac{1}{2(1+K_0)^2} \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N^2]}(Nt, \omega)) \frac{\partial^2 f}{\partial x^2}.
\end{aligned}$$

The conditions  $H_1, H_2, H_3$  in [JM86, Proposition 3.2.3] are satisfied as before. Hence we conclude that the sequence  $(\mathcal{L}[(\Theta^{\text{aux},(1),[N^2]}(N^2t_2 + Nt_1))_{t_1>0}])_{N \in \mathbb{N}}$  is tight.  $\square$

### 11.3.4 Stability of the 1-block estimators

**Lemma 11.7. [Stability property of the 1-block estimator]** *Let  $\Theta^{\text{aux},(1),[N^2]}(t)$  be defined as in (10.14). For any  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ ,*

$$(11.75) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N^2]}(N^2t_2 + Nt_1) - \bar{\Theta}^{(1),[N^2]}(N^2t_2 + Nt_1 - t) \right| = 0 \text{ in probability,}$$

$$(11.76) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_1}^{(1), [N^2]}(N^2 t_2 + N t_1) - \Theta_{y_1}^{(1), [N^2]}(N^2 t_2 + N t_1 - t) \right| = 0 \text{ in probability,}$$

$$(11.77) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_2}^{(1), [N^2]}(N^2 t_2 + N t_1) - \Theta_{y_2}^{(1), [N^2]}(N^2 t_2 + N t_1 - t) \right| = 0 \text{ in probability.}$$

*Proof.* Define

$$(11.78) \quad u = N^2 t_2 + N t_1.$$

From the SSDE in (11.7) we obtain that

$$(11.79) \quad \begin{aligned} & \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1), [N^2]}(u) - \bar{\Theta}^{(1), [N^2]}(u - t) \right| > \epsilon \right] \\ &= \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \frac{1}{1 + K_0} \left| \int_{u-t}^u dr \frac{c_1}{N} \left[ \frac{1}{N^2} \sum_{j \in [N^2]} x_j^{[N^2]}(r) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(r) \right] \right. \right. \\ & \quad \left. \left. + \int_{u-t}^u dr \frac{K_1 e_1}{N} \left[ \Theta_{y_1}^{(1), [N^2]}(r) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(r) \right] \right. \right. \\ & \quad \left. \left. + \frac{K_2 e_2}{N^2} \left[ \frac{1}{N} \sum_{i \in [N]} y_{i,2}^{[N^2]}(r) - \frac{1}{N} \sum_{i \in [N]} x_i^{[N^2]}(r) \right] \right. \right. \\ & \quad \left. \left. + \int_{u-t}^u dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N^2]}(r))} \right| > \epsilon \right] \\ &\leq \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \frac{1}{1 + K_0} \left| \int_{u-t}^u dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N^2]}(r))} \right| \right. \\ & \quad \left. > \epsilon - \frac{L(N) 2(c_1 + K_1 e_1 + \frac{K_2 e_2}{N})}{N(1 + K_0)} \right] \\ &\leq \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \frac{1}{1 + K_0} \left| \int_{u-t}^u dw_i(r) \frac{1}{N} \sum_{i \in [N]} \sqrt{g(x_i^{[N^2]}(r))} \right| > \frac{\epsilon}{2} \right]. \end{aligned}$$

Via the same optional stopping time argument as in the proof of Lemma 8.15, the above computation shows that (11.75) holds. Note that the extra drift term  $\frac{K_2 e_2}{N}$  does not have any influence. Equations (11.76)–(11.77) hold by a similar argument as given in the proof of Lemma 10.5.  $\square$

### 11.3.5 Limiting evolution for the single components

**Proposition 11.8. [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  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$  such that

$$(11.80) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \Theta^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k t_1) \right] = P_{t_1, t_2}, \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( Y_{1,0}^{[N_k^2]}(N_k^2 t_2 + N_k t_1), Y_{2,0}^{[N_k^2]}(N_k^2 t_2) \right) \middle| \Theta^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k t_1) \right] \\ & \quad = P_1^{\text{eff}}(t_1), \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \sup_{0 \leq t \leq L(N_k)} \left| \bar{\Theta}^{[N_k^2]}(N_k^2 t_2 + N_k t_1) - \bar{\Theta}^{[N_k^2]}(N_k^2 t_2 + N_k t_1 - t) \right| \right. \\ & \quad \left. + \left| \Theta_{y_1}^{[N_k^2]}(N_k^2 t_2 + N_k t_1) - \Theta_{y_1}^{[N_k^2]}(N_k^2 t_2 + N_k t_1 - t) \right| \right. \\ & \quad \left. + \left| \Theta_{y_2}^{[N_k^2]}(N_k^2 t_2 + N_k t_1) - \Theta_{y_2}^{[N_k^2]}(N_k^2 t_2 + N_k t_1 - t) \right| \right] = \delta_0, \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left( Z^{[N_k^2]}(N_k^2 t_2 + N_k t_1), \right) = \nu(t_1, t_2). \end{aligned}$$

Then  $\nu(t_1, t_2)$  is of the form

$$(11.81) \quad \nu(t_1, t_2) = \int_{[0,1]^2} P_{t_1, t_2}(d\theta^{(1)}, d\theta_y^{(1)}) \int_{[0,1]^{\mathbb{N}_0}} P^{(\theta^{(1)}, \theta_y^{(1)})}(d\mathbf{y}) \nu_{\theta, \mathbf{y}},$$

where

$$(11.82) \quad \nu_{\theta, \mathbf{y}_0} = \prod_{i \in \mathbb{N}_0} \Gamma_{(\theta, \mathbf{y}_0, i)}$$

with  $\Gamma_{(\theta, \mathbf{y}_0, i)}$  the equilibrium measure for the  $i$ 'th single colony defined in (11.24).

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 (11.80) hold. The second line in (11.80) follows from assumptions (11.37) and (11.38). To prove Proposition 11.8 we proceed as in the proof of Proposition 10.6, but with the finite system in (10.4) replaced by the system in (11.6) and the infinite system in (10.11) replaced by the system in (11.18). Note that Lemma 10.7 holds also for the system in (11.18), after adding the non-interacting component  $y_{2,0}$  to the equilibrium. The equivalent of Lemma 10.8 will again follow from the equivalent of Lemma 10.12. We will derive the analogue of Lemmas 10.9 and 10.10 (see Lemma's 11.9 and 11.10 below). Lemma 10.11 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 (11.88)) equals the one for the one-layer two-colour mean-field system, up to a non-interacting component, we obtain an equivalent of Lemma 10.12. Finally, also the equivalent of Lemma 10.13 holds under an additional assumption, see Lemma 11.11. Finally Corollary 11.12 states the equivalent of Corollary 10.14. With the help of the lemma's and the corollary, the proof of Proposition 11.8 follows from the same argument as used in the proof of Proposition 10.6.

**Lemma 11.9. [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 (11.20). Then

$$(11.83) \quad \mathbb{E} \left[ \left| \Theta_x^{(1), [N]}(t) - \Theta_{y_0}^{(1), [N]}(t) \right| \right] \leq \sqrt{\mathbb{E} \left[ \left( \Theta_x^{(1), [N]}(0) - \Theta_{y_0}^{(1), [N]}(0) \right)^2 \right]} e^{-(K_0 e_0 + e_0)t} \\ + \sqrt{\frac{1}{K_0 e_0 + e_0} \left[ \frac{c_1}{N} + \frac{\|g\|}{N} + \frac{K_1 e_1}{N} + \frac{K_2 e_2}{N^2} \right]}.$$

*Proof.* The result follows by Itô-calculus on the SSDE in (11.6) and the same type of argument as used in the proof of Lemma 10.9.  $\square$

Like for the mean-field system with one colour, we need to compare the finite system in (11.6) with an infinite system. To derive the analogue of Lemma 10.10, let  $L(N)$  satisfy  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . 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

$$(11.84) \quad u = N^2 t_2 + N t_1$$

and let  $\mu_N$  be the measure on  $([0, 1]^3)^{\mathbb{N}_0}$  by continuing the configuration of

$$(11.85) \quad \left( Z^{[N^2]}(u - L(N)) \right) \\ = \left( X^{[N^2]}(u - L(N)), \left( Y_0^{[N^2]}(u - L(N)), Y_1^{[N^2]}(u - L(N)), Y_2^{[N^2]}(u - L(N)) \right) \right)$$

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

$$(11.86) \quad \bar{\Theta}^{(1), [N^2]} = \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N^2]}(u - L(N)) + K_0 y_{i,0}^{[N^2]}(u - L(N))}{1 + K_0}.$$

Let

$$(11.87) \quad (Z^{\mu_N}(t))_{t \geq 0} = (X^{\mu_N}(t), Y_0^{\mu_N}(t), Y_1^{\mu_N}(t), Y_2^{\mu_N}(t))_{t \geq 0}$$

be the infinite system evolving according to

$$(11.88) \quad \begin{aligned} dx_i^{\mu_N}(t) &= c_0 [\bar{\Theta}^{(1),[N^2]} - x_i^{\mu_N}(t)] dt + \sqrt{g(x_i^{\mu_N}(t))} dw_i(t) + K_0 e_0 [y_{i,0}^{\mu_N}(t) - x_i^{\mu_N}(t)] dt, \\ dy_{i,0}^{\mu_N}(t) &= e_0 [x_i^{\mu_N}(t) - y_{i,0}^{\mu_N}(t)] dt, \\ y_{i,1}^{\mu_N}(t) &= y_{i,1}^{\mu_N}(0), \\ y_{i,2}^{\mu_N}(t) &= y_{i,2}^{\mu_N}(0), \quad i \in \mathbb{N}_0, \end{aligned}$$

starting from initial distribution  $\mu_N$ . Then the following Lemma 11.10 is the equivalent of Lemma 10.10 for the three-colour two-layer mean-field system. In particular, the infinite system considered in Lemma 11.10 is similar to the infinite system in Lemma 10.10. The only difference is that there is one more non-interacting component added in (11.88).

**Lemma 11.10. [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  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N$ . Suppose that

$$(11.89) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(1),[N]}(u) - \bar{\Theta}^{(1),[N]}(u-t) \right| = 0 \quad \text{in probability.}$$

Then, for all  $t \geq 0$ ,

$$(11.90) \quad \lim_{k \rightarrow \infty} \left| \mathbb{E} \left[ f(Z^{\mu_N}(t)) - f(Z^{[N^2]}(u - L(N) + t)) \right] \right| = 0 \quad \forall f \in \mathcal{C}([0, 1]^3)^{\mathbb{N}_0}, \mathbb{R}.$$

*Proof.* We proceed as in the proof of Lemma 10.10 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

$$(11.91) \quad \begin{aligned} dx_i^{[N^2]}(t) &= c_0 \left[ \Theta^{(1),[N^2]} - x_i^{[N^2]}(t) \right] dt + c_0 \left[ \bar{\Theta}^{(1),[N^2]}(t) - \Theta^{(1),[N^2]} \right] dt \\ &\quad + c_0 \left[ \Theta_x^{(1),[N^2]}(t) - \bar{\Theta}^{(1),[N^2]}(t) \right] dt + \frac{c_1}{N} \left[ \frac{1}{N^2} \sum_{i \in [N^2]} x_j^{[N^2]}(t) - x_i^{[N^2]}(t) \right] dt \\ &\quad + \sqrt{g(x_i^{[N^2]}(t))} dw_i(t) + K_0 e_0 [y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt \\ &\quad + \frac{K_1 e_1}{N} [y_{i,1}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt + \frac{K_2 e_2}{N^2} [y_{i,0}^{[N^2]}(t) - x_i^{[N^2]}(t)] dt, \\ dy_{i,0}^{[N^2]}(t) &= e_0 [x_i^{[N^2]}(t) - y_{i,0}^{[N^2]}(t)] dt, \\ dy_{i,1}^{[N^2]}(t) &= \frac{e_1}{N} [x_i^{[N^2]}(t) - y_{i,1}^{[N^2]}(t)] dt, \\ dy_{i,2}^{[N^2]}(t) &= \frac{e_2}{N^2} [x_i^{[N^2]}(t) - y_{i,2}^{[N^2]}(t)] dt, \quad i \in [N]. \end{aligned}$$

Using this SSDE we can exactly proceed as in the proof of Lemma 10.10 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 10.10, since its rate of interaction with the active population is even slower than the rate of interaction of the colour-1 seed-bank.  $\square$

Finally, we state the equivalent of Lemma 10.13 for the three-colour two-layer mean-field system.

**Lemma 11.11. [Coupling of finite systems]** Let

$$(11.92) \quad Z^{[N^2],1} = (X^{[N^2],1}, Y_0^{[N^2],1}, Y_1^{[N^2],1}, Y_2^{[N^2],1})$$

be the finite system evolving according to (11.6) 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

$$(11.93) \quad Z^{[N^2],2} = (X^{[N^2],2}, Y_0^{[N^2],2}, Y_1^{[N^2],2}, Y_2^{[N^2],2})$$



be the finite system evolving according to (11.6) 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  $\bar{\mu}$  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

$$(11.94) \quad \begin{aligned} \bar{\Theta}^{[N],1} &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N^2],1} + K_0 y_{i,0}^{[N^2],1}}{1 + K_0}, & \bar{\Theta}^{[N],2} &= \frac{1}{N} \sum_{i \in [N]} \frac{x_i^{[N^2],2} + K_0 y_{i,0}^{[N^2],2}}{1 + K_0}, \\ \bar{\Theta}^1 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i^1 + K_0 y_{i,0}^1}{1 + K_0}, & \bar{\Theta}^2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in [n]} \frac{x_i^2 + K_0 y_{i,0}^2}{1 + K_0}, \end{aligned}$$

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 (10.14) for  $Z^{[N^2],1}$ , respectively,  $Z^{[N^2],2}$ . Suppose that

$$(11.95) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left( \left| \bar{\Theta}^{[N],k}(0) - \bar{\Theta}^{[N],k}(t) \right| \right) = 0 \quad \text{in probability, } k \in \{1, 2\},$$

and suppose that  $\bar{\mu}(\{\bar{\Theta}_1 = \bar{\Theta}_2, Y_1^1 = Y_1^2, Y_2^1 = Y_2^2\}) = 1$ . Then, for any sequence  $(t(N))_{N \in \mathbb{N}}$  with  $\lim_{N \rightarrow \infty} t(N) = \infty$ ,

$$(11.96) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} [ & |x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K_0 |y_{i,0}^{[N],1}(t(N)) - y_{i,0}^{[N],2}(t(N))| \\ & + K_1 |y_{i,1}^{[N],1}(t(N)) - y_{i,1}^{[N],2}(t(N))| + K_2 |y_{i,2}^{[N],1}(t(N)) - y_{i,2}^{[N],2}(t(N))| ] = 0. \end{aligned}$$

*Proof.* Like in the proof of Lemma 10.13, we can show with Itô calculus that the function

$$(11.97) \quad \begin{aligned} t \rightarrow \mathbb{E} [ & |x_i^{[N],1}(t(N)) - x_i^{[N],2}(t(N))| + K_0 |y_{i,0}^{[N],1}(t(N)) - y_{i,0}^{[N],2}(t(N))| \\ & + K_1 |y_{i,1}^{[N],1}(t(N)) - y_{i,1}^{[N],2}(t(N))| + K_2 |y_{i,2}^{[N],1}(t(N)) - y_{i,2}^{[N],2}(t(N))| ] \end{aligned}$$

is monotonically decreasing. Hence we can proceed as in the proof of Lemma 8.13 to show that (11.96) is true.  $\square$

From the above couplings we can derive the following corollary, which is the analogue of Corollary 10.14 for the two-level three-colour mean-field system.

**Corollary 11.12.** Fix  $t_1, t_2 > 0$  and set  $u = N^2 t_2 + N t_1$ . Let  $\mu_N$  be the measure obtained by periodic continuation of

$$(11.98) \quad Z^{[N^2]}(u - L(N)) = (X^{[N^2]}(u - L(N)), Y_0^{[N^2]}(u - L(N)), Y_1^{[N^2]}(u - L(N)), Y_2^{[N^2]}(u - L(N))),$$

and let  $\mu$  be a weak limit point of the sequence  $(\mu_N)_{N \in \mathbb{N}}$ . Let

$$(11.99) \quad \Theta = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i \in [N]} \frac{x_i^\mu + K y_i^\mu}{1 + K} \quad \text{in } L^2(\mu),$$

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 (11.18) with  $\theta = \Theta$  and  $y_{i,1,0}$  and  $y_{i,2,0}$  determined by assumption (11.80) 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  $\lim_{N \rightarrow \infty} \bar{L}(N) = \infty$ ,  $\lim_{N \rightarrow \infty} \frac{\bar{L}(N)}{N} = 0$  and

$$(11.100) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} [ & |x_i^{[N^2]}(Ns) - x_i^{\nu_\Theta}(\bar{L}(N))| ] + K_0 \mathbb{E} [ |y_{i,0}^{[N^2]}(Ns) - y_{i,0}^{\nu_\Theta}(\bar{L}(N))| ] \\ & + K_1 \mathbb{E} [ |y_{i,1}^{[N^2]}(Ns) - y_{i,1}^{\nu_\Theta}(\bar{L}(N))| ] + K_2 \mathbb{E} [ |y_{i,2}^{[N^2]}(Ns) - y_{i,2}^{\nu_\Theta}(\bar{L}(N))| ] = 0, \quad i \in [N]. \end{aligned}$$

Note that Lemmas 11.10, 11.11 and Corollary 11.12 do not only hold for sites  $i$  in the 1-block around 0, but hold for all sites  $i \in [N^2]$ , after we replace  $\Theta_0^{[N]^0}$  by  $\Theta_i^{[N]^i}$ .

### 11.3.6 Limiting evolution of the 1-block estimator process

**Proposition 11.13. [Limiting evolution of the 1-blocks]** Fix  $t_2 > 0$ . Let  $(L(N))_{N \in \mathbb{N}}$  satisfy  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . Let  $(N_k)_{k \in \mathbb{N}}$  be a subsequence such that

$$(11.101) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff},(2),[N_k^2]}(N_k^2 t_2) \right) \right] = P_{t_2}(\cdot), \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ y_{2,1}^{[N_k^2]}(N_k t_2) \middle| \Theta^{(2),[N_k^2]}(N_k^2 t_2) \right] = P^{z_2(t_2)}, \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( Y_{1,0}^{[N_k^2]}(N_k^2 t_2 + N_k t_1), Y_{2,0}^{[N_k^2]}(N_k^2 t_2) \right) \middle| \Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right] = P^{z_1^{\text{eff}}(t_1)}, \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \sup_{0 \leq t \leq L(N_k)} \left| \bar{\Theta}^{(2),[N_k^2]}(N_k^2 t_2) - \bar{\Theta}^{(2),[N_k^2]}(N_k^2 t_2 - N_k t) \right| \right. \\ & \quad \left. + \left| \Theta_{y_2}^{(2),[N_k^2]}(N_k^2 t_2) - \Theta_{y_2}^{(2),[N_k^2]}(N_k^2 t_2 - N_k t) \right| \right] = \delta_0. \end{aligned}$$

Then, for the 1-block around 0,

$$(11.102) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2) \right] = \int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv),$$

where  $\Gamma_{u,y_{2,1}}^{\text{aux},(1)}$  is the equilibrium measure of (11.28) with  $\theta$  replaced by  $u$ , and

$$(11.103) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right)_{t_1 > 0} \right] = \mathcal{L}[(z_1^{\text{aux}}(t_1))_{t_1 > 0}],$$

where  $(z_1^{\text{aux}}(t_1))_{t_1 > 0}$  is the process evolving according to (11.28) with  $\theta$  replaced by the random variable  $\bar{\Theta}^{(2)}(t_2)$  and with initial measure  $\int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv)$ , and  $y_{2,1}$  is a random variable.

Note that by tightness of the 2-blocks and the assumptions in Proposition 11.1, we can always find a subsequence  $(N_k)_{k \in \mathbb{N}}$  such that (11.101) holds and also (11.80) holds. To prepare for the proof of Proposition 11.13, we prove four lemmas: Lemma 11.14 shows that the limiting 1-block system has a unique equilibrium, Lemma 11.16 implies convergence of the active 2-block estimator and the combined 2-block estimator, Lemma 11.17 gives a regularity property for the 2-block estimator, and Lemma 11.18 shows the limiting evolution of the auxiliary 1-block estimator process. Lemma 11.20 proves equation (11.102). After we derive these lemmas we prove Proposition 11.13.

**Lemma 11.14. [1-block equilibrium]** For any initial distribution  $\mu \in ([0,1]^3)$ , the process  $(z_1^{\text{aux}}(t_1))_{t_1 > 0}$  evolving according to (11.28) is well defined and converges to a unique equilibrium measure

$$(11.104) \quad \lim_{t_1 \rightarrow \infty} \mathcal{L}[z_1^{\text{aux}}(t_1)] = \Gamma_{\theta, y_{2,1}}^{\text{aux},(1)}.$$

*Proof.* By [YW71], the SSDE in (11.28) has a unique strong solution. By a similar argument as in the proof of Lemma 10.7, the SSDE in (11.28) converges to a unique equilibrium measure  $\Gamma_{\theta, y_{2,1}}^{\text{aux},(1)}$ .  $\square$

**Remark 11.15. [Equilibrium measure]** Note that Lemma 11.14 still holds when we allow  $\theta$  and  $y_{2,1}$  to be the random variables  $\bar{\Theta}(t_2)$  and  $y_{2,1}$ . Assuming (11.101), we can derive the distributions of  $\bar{\Theta}(t_2)$  and  $y_{2,1}$ , and we can write the equilibrium as  $\int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv)$ . In what follows we abbreviate

$$(11.105) \quad \Gamma_{\bar{\Theta}(t_2), y_{2,1}, i}^{(1)} = \int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}, i}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}, i) P_{t_2}(du, dv).$$

■

**Lemma 11.16.** [2-block averages] *Define*

$$(11.106) \quad \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) = \frac{\Theta_x^{(2),[N^2]}(Nt_1) + K_0 \Theta_{y_0}^{(2),[N^2]}(Nt_1)}{1 + K_0} - \Theta_{y_1}^{(2),[N^2]}(Nt_1).$$

*Then*

$$(11.107) \quad \begin{aligned} & \mathbb{E} \left[ \left| \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right| \right] \\ & \leq \sqrt{\mathbb{E} \left[ \left( \Delta_{\Sigma}^{(2),[N^2]}(0) \right)^2 \right]} e^{-e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) t_1} \\ & + \sqrt{\int_0^{t_1} ds 2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) e^{-2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) (t_1-s)} \mathbb{E} \left[ \left| \bar{\Theta}^{(1),[N^2]}(Ns) - \Theta_x^{(1),[N^2]}(Ns) \right| \right]} \\ & + \sqrt{\frac{1}{e_1} \left[ \frac{K_2 e_2}{N(1+K_0+K_1)} + \frac{\|g\|}{2N(1+K_0+K_1)} \right]}. \end{aligned}$$

*Proof.* For the two-level mean-field system we have the following SSDE for the 2-block averages:

$$(11.108) \quad \begin{aligned} d\Theta_x^{(2),[N^2]}(Nt_1) &= \sqrt{\frac{1}{N^3} \sum_{i \in [N^2]} g(x_i^{[N^2]}(Nt_1))} d\tilde{w}(t_1) \\ &+ NK_0 e_0 \left[ \Theta_{y_0}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1) \right] dt_1 \\ &+ K_1 e_1 \left[ \Theta_{y_1}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1) \right] dt_1 \\ &+ \frac{K_2 e_2}{N} \left[ \Theta_{y_2}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1) \right] dt_1, \\ d\Theta_{y_0}^{(2),[N^2]}(Nt_1) &= Ne_0 \left[ \Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_0}^{(2),[N^2]}(Nt_1) \right] dt_1, \\ d\Theta_{y_1}^{(2),[N^2]}(Nt_1) &= e_1 \left[ \Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_1}^{(2),[N^2]}(Nt_1) \right] dt_1, \\ d\Theta_{y_2}^{(2),[N^2]}(Nt_1) &= \frac{e_2}{N} \left[ \Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_2}^{(2),[N^2]}(Nt_1) \right] dt_1. \end{aligned}$$

Therefore

$$(11.109) \quad \begin{aligned} d \left( \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right)^2 &= 2\Delta_{\Sigma}^{(2),[N^2]}(Nt_1) d\Delta_{\Sigma}^{(2),[N^2]}(Nt_1) + d \left\langle \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right\rangle \\ &= 2\Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \frac{1}{1+K_0} \sqrt{\frac{1}{N^3} \sum_{i \in [N^2]} g(x_i^{[N^2]}(Nt_1))} d\tilde{w}(t_1) \\ &+ 2\Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \frac{K_1 e_1}{(1+K_0)} \left[ \Theta_{y_1}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1) \right] dt_1, \\ &+ 2\Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \frac{K_2 e_2}{N(1+K_0)} \left[ \Theta_{y_2}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1) \right] dt_1, \\ &- 2\Delta_{\Sigma}^{(2),[N^2]}(Nt_1) e_1 \left[ \Theta_x^{(2),[N^2]}(Nt_1) - \Theta_{y_1}^{(2),[N^2]}(Nt_1) \right] dt_1 \\ &+ \frac{1}{(1+K_0)^2} \frac{1}{N^3} \sum_{i \in [N^2]} g(x_i^{[N^2]}(Nt_1)) dt_1. \end{aligned}$$

Hence

$$\begin{aligned}
& \frac{d}{dt} \mathbb{E} \left[ \left( \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right)^2 \right] \\
&= -2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) \mathbb{E} \left[ \left( \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right)^2 \right] \\
&\quad + 2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) \\
(11.110) \quad & \times \mathbb{E} \left[ \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \left( \frac{\Theta_x^{(2),[N^2]}(Nt_1) + K_0 \Theta_{y_0}^{(2),[N^2]}(Nt_1)}{1+K_0} - \Theta_x^{(2),[N^2]}(Nt_1) \right) \right] \\
& + \frac{K_2 e_2}{N(1+K_0)} 2 \mathbb{E} \left[ \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \left[ \Theta_{y_2}^{(2),[N^2]}(Nt_1) - \Theta_x^{(2),[N^2]}(Nt_1) \right] \right] \\
& + \frac{1}{(1+K_0)^2} \mathbb{E} \left[ \frac{1}{N^3} \sum_{i \in [N^2]} g(x_i^{[N^2]}(Nt_1)) \right],
\end{aligned}$$

and therefore

$$\begin{aligned}
(11.111) \quad & \mathbb{E} \left[ \left( \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right)^2 \right] \\
&= \mathbb{E} \left[ \left( \Delta_{\Sigma}^{(2),[N^2]}(0) \right)^2 \right] e^{-2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) t_1} + \int_0^{t_1} ds e^{-2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) (t_1-s)} h^{[N]}(s),
\end{aligned}$$

where

$$\begin{aligned}
(11.112) \quad & h^{[N]}(s) = 2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) \\
& \times \mathbb{E} \left[ \Delta_{\Sigma}^{(2),[N^2]}(Ns) \left( \frac{\Theta_x^{(2),[N^2]}(Ns) + K_0 \Theta_{y_0}^{(2),[N^2]}(Ns)}{1+K_0} - \Theta_x^{(2),[N^2]}(Ns) \right) \right] \\
& + \frac{2K_2 e_2}{N(1+K_0)} \mathbb{E} \left[ \Delta_{\Sigma}^{(2),[N^2]}(Ns) \left[ \Theta_{y_2}^{(2),[N^2]}(Ns) - \Theta_x^{(2),[N^2]}(Ns) \right] \right] \\
& + \frac{1}{(1+K_0)^2} \mathbb{E} \left[ \frac{1}{N^3} \sum_{i \in [N^2]} g(x_i^{[N^2]}(Ns)) \right].
\end{aligned}$$

Therefore

$$\begin{aligned}
(11.113) \quad & \mathbb{E} \left[ \left| \Delta_{\Sigma}^{(2),[N^2]}(Nt_1) \right| \right] \\
& \leq \sqrt{\mathbb{E} \left[ \left( \Delta_{\Sigma}^{(2),[N^2]}(0) \right)^2 \right]} e^{-e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) t_1} \\
& + \sqrt{\int_0^{t_1} ds 2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) e^{-2e_1 \left( \frac{1+K_0+K_1}{1+K_0} \right) (t_1-s)} \mathbb{E} \left[ \left| \bar{\Theta}^{(1),[N^2]}(Ns) - \Theta_x^{(1),[N^2]}(Ns) \right| \right]} \\
& + \sqrt{\frac{1}{e_1} \left[ \frac{K_2 e_2}{N(1+K_0+K_1)} + \frac{\|g\|}{2N(1+K_0+K_1)} \right]}.
\end{aligned}$$

□

Let  $\mu_{N_k}$  be the measure obtained by periodic continuation of the configuration

$$(11.114) \quad Z^{[N^2]}(N_k^2 t_2).$$

Since the state space  $([0, 1] \times [0, 1]^3)^{\mathbb{N}_0}$  is compact, we can pass to a further subsequence, to obtain

$$(11.115) \quad \mu = \lim_{k \rightarrow \infty} \mu_{N_k}.$$

**Lemma 11.17. [Regularity for 2-block estimator]** *Let  $\mu$  and  $\mu_N$  be as defined above. Let  $(x_i, y_{1,i}, y_{2,i})_{i \in \mathbb{N}_0}$  be distributed according to  $\mu$ . Define the random variable*

$$(11.116) \quad \begin{aligned} \phi &= (\phi_1, \phi_2), \\ \phi_1 &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i \in [n^2]} \frac{x_i + K_0 y_{i,0} + K_1 y_{i,1}}{1 + K_0 + K_1}, \quad \phi_2 = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i \in [n^2]} y_{i,2}, \end{aligned}$$

and the random variable  $\phi^{[N]}$  on  $(\mu_N, ([0, 1]^3)^{\mathbb{N}_0})$  by putting

$$(11.117) \quad \begin{aligned} \phi^{[N^2]} &= (\phi_1^{[N^2]}, \phi_2^{[N^2]}), \\ \phi_1^{[N^2]} &= \frac{1}{N^2} \sum_{i \in [N^2]} \frac{x_i^{[N^2]} + K_0 y_{i,0}^{[N^2]} + K_1 y_{i,1}^{[N^2]}}{1 + K_0 + K_1}, \quad \phi_2 = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i \in [N^2]} y_{i,2}^{[N^2]}. \end{aligned}$$

Then

$$(11.118) \quad \lim_{N \rightarrow \infty} \mathcal{L}[\phi^{[N^2]}] = \mathcal{L}[\phi].$$

*Proof.* Use a similar argument as in the proof of Lemma 10.11.  $\square$

We will first determine the limiting evolution of  $(\Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N t_1))_{t_1 > 0}$ . To do so we consider all the  $N_k$  1-blocks in  $[N_k^2]$ . After that we show that

$$(11.119) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2) \right)_{i \in [N_k]} \right] = \prod_{i \in \mathbb{N}_0} \Gamma_{\Theta(t_2), y_{2,1,i}}^{(1)},$$

The limiting 1-block process for the auxiliary estimator process (recall (11.28)) is given by

$$(11.120) \quad \begin{aligned} (\mathbf{z}_1^{\text{aux}}(t))_{t > 0} &= (\mathbf{x}_1^{\text{aux}}(t), \mathbf{y}_{1,1}^{\text{aux}}(t), \mathbf{y}_{2,1}^{\text{aux}}(t))_{t > 0}, \\ \mathbf{z}_1^{\text{aux}}(t) &= (z_{1,i}^{\text{aux}}(t))_{i \in \mathbb{N}_0}, \quad \mathbf{x}_1^{\text{aux}}(t) = (x_{1,i}^{\text{aux}}(t))_{i \in \mathbb{N}_0}, \\ \mathbf{y}_{1,1}^{\text{aux}}(t) &= (y_{1,1,i}^{\text{aux}}(t))_{i \in \mathbb{N}_0} \quad \mathbf{y}_{2,1}^{\text{aux}}(t) = (y_{2,i}^{\text{aux}}(t))_{i \in \mathbb{N}_0} \end{aligned}$$

and its components evolve according to

$$(11.121) \quad \begin{aligned} dx_{1,i}^{\text{aux}}(t) &= \frac{1}{1 + K_0} \left[ c_1 [\bar{\Theta}^{(2)}(t_2) - x_{1,i}^{\text{aux}}(t)] dt + \sqrt{(\mathcal{F}^{(1)}g)(x_{1,i}^{\text{aux}}(t))} dw(t) \right. \\ &\quad \left. + K_1 e_1 [y_{1,1,i}^{\text{aux}}(t) - x_{1,i}^{\text{aux}}(t)] dt \right], \end{aligned}$$

$$\begin{aligned} dy_{1,1,i}^{\text{aux}}(t) &= e_1 [x_{1,i}^{\text{aux}}(t) - y_{1,1,i}^{\text{aux}}(t)] dt, \\ y_{2,1,i}^{\text{aux}}(t) &= y_{2,1,i}, \quad i \in \mathbb{N}_0, \end{aligned}$$

where

$$(11.122) \quad \bar{\Theta}^{(2)}(t_2) = \lim_{N \rightarrow \infty} \sum_{i \in [N^2]} \frac{x_i^{[N^2]} + K_0 y_{i,0}^{[N^2]} + K_1 y_{i,1}^{[N^2]}}{1 + K_0 + K_1} \text{ in } L_2(\mu).$$

Let  $\mu_{N_k}^{(1)}$  be the law obtained by periodic continuation of  $(\Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2))_{i \in [N_k]}$ , and let  $\mu^{(1)} = \lim_{k \rightarrow \infty} \mu_{N_k}^{(1)}$  be any weak limit point of the sequence  $(\mu_{N_k}^{(1)})_{k \in \mathbb{N}}$ .

**Lemma 11.18. [Limiting evolution of auxiliary 1-block estimator]** *Let  $\mathcal{L}[(\mathbf{z}_1^{\text{aux}}(0))] = \mu^{(1)}$ . Then the following hold.*

1. For all  $t_1 > 0$  and  $i \in [N_k]$ ,

$$(11.123) \quad \lim_{k \rightarrow \infty} \mathbb{E} \left[ (1 + K_0) \left( x_{1,i}^{\text{aux}}(t_1) - \bar{\Theta}_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right)^2 \right. \\ \left. + K_1 \left( y_{1,1,i}(t_1) - \Theta_{y_{1,i}}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right)^2 \right. \\ \left. + K_2 \left( y_{2,1,i}^{\text{aux},(1),[N_k^2]}(t_1) - \Theta_{y_{2,i}}^{\text{aux},(1),[N_k]}(N_k^2 t_2 + N_k t_1) \right)^2 \right] = 0.$$

2. For all  $t_2 > 0$ ,

$$(11.124) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ (\Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N t_1))_{t_1 > 0} \right] = \mathcal{L}[(\mathbf{z}_1^{\text{aux}}(t_1))_{t_1 > 0}].$$

*Proof.* Abbreviate

$$(11.125) \quad \begin{aligned} \Delta_i^{(1),[N_k^2]}(N_k t_1) &= x_{1,i}^{\text{aux}}(t_1) - \Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1), \\ \delta_{y_{1,i}}^{(1),[N_k^2]}(N_k t_1) &= y_{1,1,i}(t_1) - \Theta_{y_{1,i}}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1), \\ \delta_{y_{2,i}}^{(1),[N_k^2]}(N_k t_1) &= y_{2,1,i}^{\text{aux}}(t_1) - \Theta_{y_{2,i}}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1). \end{aligned}$$

Extending  $(\Theta^{\text{aux},(1),[N_k^2]}(N_k t_1))_{t_1 > 0}$  as a process on  $\mathbb{N}_0$  by periodic continuation, we can construct  $(\mathbf{z}_1^{\text{aux}}(t_1))_{t_1 > 0}$  and  $(\Theta^{\text{aux},(1)}(N_k t_1))_{t_1 > 0}$  on one probability space such that

$$(11.126) \quad \lim_{k \rightarrow \infty} \Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2) = \mathbf{z}_1^{\text{aux},(1)}(0) \quad a.s.$$

We couple the processes  $(\mathbf{z}_1^{\text{aux}}(t_1))_{t_1 > 0}$  and  $(\Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1))_{t_1 > 0}$  by using the same Brownian motions for both processes. By Itô-calculus we obtain for the coupled process (recall

(11.73))

$$\begin{aligned}
& \mathbb{E} \left[ (1 + K_0) \left( \Delta_i^{(1), [N_k^2]}(N_k t_1) \right)^2 + K_1 \left( \delta_{y_1, i}^{(1), [N_k^2]}(N_k t_1) \right)^2 + K_2 \left( \delta_{y_2, i}^{(1), [N_k^2]}(N_k t_1) \right)^2 \right] \\
= & \mathbb{E} \left[ (1 + K_0) \left( \Delta_i^{(1), [N_k^2]}(0) \right)^2 + K_1 \left( \delta_{y_1, i}^{(1), [N_k^2]}(0) \right)^2 + K_2 \left( \delta_{y_2, i}^{(1), [N_k^2]}(0) \right)^2 \right] \\
& - 2c_1 \int_0^{t_1} \mathbb{E} \left[ \left( \Delta_i^{(1), [N_k^2]}(N_k s) \right)^2 \right] ds \\
& - 2K_1 e_1 \int_0^{t_1} \mathbb{E} \left[ \left( \Delta_i^{(1), [N_k^2]}(N_k s) - \delta_{y_1, i}^{(1), [N_k^2]}(N_k s) \right)^2 \right] ds \\
& + 2c_1 \int_0^{t_1} \mathbb{E} \left[ \Delta_i^{(1), [N_k^2]}(N_k s) \left( \Theta^{(2)}(t_2) - \frac{1}{N_k^2} \sum_{j \in [N_k^2]} x_j^{[N_k^2]}(N_k^2 t_2 + N_k s) \right) \right] ds \\
& + (K_1 e_1 + c_1) \int_0^{t_1} \mathbb{E} \left[ \Delta_i^{(1), [N_k^2]}(N_k s) \right. \\
& \quad \left. \times \left[ \frac{1}{N_k} \sum_{j \in [N_k]_i} x_j^{[N_k^2]}(N_k^2 t_2 + N_k s) - \bar{\Theta}_i^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k s) \right] \right] ds \\
& + 2K_1 e_1 \int_0^{t_1} \mathbb{E} \left[ \delta_{y_1, i}^{(1), [N_k^2]}(N_k s) \right. \\
& \quad \left. \times \left[ \bar{\Theta}_i^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k s) - \frac{1}{N_k} \sum_{j \in [N_k]_i} x_j^{[N_k^2]}(N_k^2 t_2 + N_k s) \right] \right] ds \\
& + 2 \frac{K_2 e_2}{N_k} \int_0^{t_1} \mathbb{E} \left[ \left[ \delta_{y_1, i}^{(1), [N_k^2]}(N_k s) - \Delta_i^{(1), [N_k^2]}(N_k^2 t_2 + N_k s) \right] \right. \\
& \quad \left. \times \left[ \frac{1}{N_k} \sum_{j \in [N_k]_i} x_j^{[N_k^2]}(N_k^2 s) - \Theta_{y_2, i}^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k t_1) \right] \right] ds \\
& + (1 + K_0)^2 \int_0^{t_1} \mathbb{E} \left[ \left( \sqrt{\mathcal{F}g}(x_1^{\text{aux}}(s)) - \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i^{[N_k^2]}(N_k^2 t_2 + N_k s))} \right)^2 \right] ds.
\end{aligned}
\tag{11.127}$$

Note that  $|\Delta_i^{(1), [N_k^2]}| \leq 1$  and  $|\delta_{y_1, i}^{(1), [N_k^2]}| \leq 1$ . Note that the first term tends to 0 by (11.126). We show by dominated convergence that also all other positive terms in the right-hand side of (11.127) tend to 0 as  $k \rightarrow \infty$ .

For the third term, we can estimate

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \mathbb{E} \left[ \frac{c_1}{1+K_0} \left| \frac{1}{N_k^2} \sum_{i \in [N_k^2]} x_j^{[N_k^2]}(N_k^2 t_2 + N_k s) - \bar{\Theta}^{(2)}(t_2) \right| \right] \\
& \leq \lim_{k \rightarrow \infty} \mathbb{E} \left[ \frac{c_1}{1+K_0} \left| \Theta_x^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) \right. \right. \\
& \quad \left. \left. - \frac{\Theta_x^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) + K_0 \Theta_{y_0}^{(2), [N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0} \right| \right] \\
(11.128) \quad & + \mathbb{E} \left[ \frac{c_1}{1+K_0} \left| \frac{\Theta_x^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) + K_0 \Theta_{y_0}^{(2), [N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0} \right. \right. \\
& \quad \left. \left. - \frac{\Theta_x^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) + K_0 \Theta_{y_0}^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) + K_1 \Theta_{y_1}^{(2), [N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0+K_1} \right| \right] \\
& + \mathbb{E} \left[ \frac{c_1}{1+K_0} \left| \frac{\Theta_x^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) + K_0 \Theta_{y_0}^{(2), [N_k^2]}(N_k^2 t_2 + N_k s) + K_1 \Theta_{y_1}^{(2), [N_k^2]}(N_k^2 t_2 + N_k s)}{1+K_0+K_1} \right. \right. \\
& \quad \left. \left. - \bar{\Theta}^{(2)}(t_2) \right| \right].
\end{aligned}$$

The first term in (11.128) tends to zero by Lemma 11.9, the second term tends to zero by Lemma 11.16, while the third term tends to zero by Lemma 11.17 and Lemma 11.5, which is the third assumption in (11.101). Hence the third term in (11.127) tends to zero by dominated convergence as  $k \rightarrow \infty$ .

The fourth and fifth term in (11.127) tend to zero by Lemma 11.9 and dominated convergence. The sixth term in (11.127) 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 (11.127) tends to zero, recall that the subsequence  $N_k$  is chosen such that

$$(11.129) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ (\Theta^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N t_1))_{t_1 > 0} \right]$$

exists. Note that

$$\begin{aligned}
(11.130) \quad & \mathbb{E} \left[ \left( \sqrt{(\mathcal{F}g)(x_1^{\text{aux}}(s))} - \sqrt{\frac{1}{N} \sum_{i \in [N]} g(x_i^{[N_k^2]}(N_k^2 t_2 + N_k s))} \right)^2 \right] \\
& \leq \mathbb{E} \left[ \left| (\mathcal{F}g)(x_1^{\text{aux}}(s)) - \frac{1}{N} \sum_{i \in [N]} g(x_i^{[N_k^2]}(N_k^2 t_2 + N_k s)) \right| \right],
\end{aligned}$$

and hence we can apply a similar reasoning as in (9.97) to see that (11.130) tends to zero as  $k \rightarrow \infty$ . Therefore we obtain

$$\begin{aligned}
(11.131) \quad & \lim_{k \rightarrow \infty} \mathbb{E} \left[ (1+K_0)(\Delta_i^{(1), [N_k^2]}(N_k t_1))^2 + K_1(\delta_{y_1, i}^{(1), [N_k^2]}(N_k t_1))^2 + K_2(\delta_{y_2, i}^{(1), [N_k^2]}(N_k t_1))^2 \right] \\
& = 0.
\end{aligned}$$

To prove (11.124), note that (11.131) implies convergence of the finite-dimensional distributions of  $(\Theta^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k t_1))_{t_1 > 0}$  by a similar argument as given below (9.36). By Lemma 11.6 we see that the laws of the processes

$$(11.132) \quad \left( \mathcal{L} \left[ (\Theta^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k t_1))_{t_1 > 0} \right] \right)_{k \in \mathbb{N}_0}$$

are tight. Therefore (11.124) indeed holds.  $\square$

**Remark 11.19.** Note that in the proof of Lemma 10.15 we could have proceeded as in the proof of Lemma 11.18, instead of using the criterion in [JM86, Theorem 3.3.1].  $\blacksquare$



**Lemma 11.20.** [Proof of (11.102)] *Under the assumptions in Proposition 11.13,*

$$(11.133) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2) \right] = \int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv).$$

*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  $\lim_{n \rightarrow \infty} t_n = \infty$  and  $\lim_{n \rightarrow \infty} t_n/n = 0$ . For each  $t_n$ , let  $\mu_{N,t_n}^{(1)}$  be the law obtained by periodic continuation of the configuration of  $(\Theta_i^{\text{aux},(1),[N^2]}(N^2 t_2 - N t_n))_{i \in [N]}$ . Recall that, since our state space is compact, the sequence  $(\mu_{N,t_n}^{(1)})_{N \in \mathbb{N}}$  is tight. Let  $\mu_{t_n}^{(1)}$  be any weak limit point of the sequence  $(\mu_{N,t_n}^{(1)})_{N \in \mathbb{N}}$ .

Let  $\mathcal{L}[\mathbf{z}_1^{\text{aux},n}(0)]$  be the law obtained by periodic continuation of  $\Theta_i^{\text{aux},(1),[N^2]}(N^2 t_2 - N t_n)$ . By Lemma 11.6 we know that the sequence

$$(11.134) \quad \left( \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),[N^2]}(N^2 t_2 - N t_n + N t_1) \right)_{t_1 > 0, i \in [N]} \right] \right)_{N \in \mathbb{N}}$$

is tight and hence for each  $t_n$  we can pass to a subsequence such that

$$(11.135) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_1) \right)_{t_1 > 0, i \in [N]} \right]$$

exists. By Lemmas 11.5–11.17, we obtain for all  $t_n$  that

$$(11.136) \quad \lim_{N \rightarrow \infty} \bar{\Theta}_i^{(2),[N_k^2]}(N_k^2 t_2 - N t_n) = \bar{\Theta}_i^{(2)}(t_2) \text{ in probability.}$$

Then, by (11.127) in the proof of Lemma 11.18, for fixed  $t_n$  and all  $i \in [N]$ ,

$$(11.137) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathbb{E} \left[ (1 + K_0) \left( x_{1,i}^{\text{aux},n}(t_n) - \bar{\Theta}_i^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_n) \right)^2 \right. \\ & + K_1 \left( y_{1,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{1,i}}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_n) \right)^2 \\ & \left. + K_2 \left( y_{2,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{2,i}}^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 - N_k t_n + N_k t_n) \right)^2 \right] = 0. \end{aligned}$$

By contradiction we can argue that

$$(11.138) \quad \begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{E} \left[ (1 + K_0) \left( x_{1,i}^{\text{aux},n}(t_n) - \bar{\Theta}_i^{\text{aux},(1),[N^2]}(N^2 t_2 - N t_n + N t_n) \right)^2 \right. \\ & + K_1 \left( y_{1,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{1,i}}^{\text{aux},(1),[N^2]}(N^2 t_2 - N t_n + N t_n) \right)^2 \\ & \left. + K_2 \left( y_{2,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{2,i}}^{\text{aux},(1),[N^2]}(N^2 t_2 - N t_n + N t_n) \right)^2 \right] = 0. \end{aligned}$$

To see why, suppose that (11.138) does not hold. Then for any  $\delta > 0$  we can construct a sequence  $(N_l)_{l > 0}$  such that, for  $l \in \mathbb{N}$ ,

$$(11.139) \quad \begin{aligned} & \mathbb{E} \left[ (1 + K_0) \left( x_{1,i}^{\text{aux},n}(t_n) - \bar{\Theta}_i^{\text{aux},(1),[N_l^2]}(N_l^2 t_2 - N_l t_n + N_l t_n) \right)^2 \right. \\ & + K_1 \left( y_{1,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{1,i}}^{\text{aux},(1),[N_l^2]}(N_l^2 t_2 - N_l t_n + N_l t_n) \right)^2 \\ & \left. + K_2 \left( y_{2,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{2,i}}^{\text{aux},(1),[N_l^2]}(N_l^2 t_2 - N_l t_n + N_l t_n) \right)^2 \right] > \delta. \end{aligned}$$

However, also the sequence

$$(11.140) \quad \left( \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),[N_l^2]}(N_l^2 t_2 - N_l t_n + N_l t_1) \right)_{t_1 > 0, i \in [N]} \right] \right)_{l \in \mathbb{N}}$$

is tight. Hence we can pass to a further subsequence  $(N_{\bar{l}})_{\bar{l} \in \mathbb{N}}$  for which (11.137) holds. But this contradicts (11.139). We conclude that (11.138) indeed holds. Moreover the argument holds for all  $t_n$ , so that (11.138) holds for all  $t_n$ .

Hence for every  $t_n$  there exists a  $N_n$  such that, for all  $N \geq N_n$ ,

$$(11.141) \quad \mathbb{E} \left[ (1 + K_0) \left( x_{1,i}^{\text{aux},n}(t_n) - \bar{\Theta}_i^{\text{aux},(1),[N_n^2]}(N_n^2 t_2) \right)^2 + K_1 \left( y_{1,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{1,i}}^{\text{aux},(1),[N_n^2]}(N_n^2 t_2) \right)^2 + K_2 \left( y_{2,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{2,i}}^{\text{aux},(1),[N_n^2]}(N_n^2 t_2) \right)^2 \right] < \frac{1}{n}.$$

In particular, we may require that  $N_n > N_{n-1}$ . Setting  $N = N_n$ , we obtain

$$(11.142) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[ (1 + K_0) \left( x_{1,i}^{\text{aux},n}(t_n) - \bar{\Theta}_i^{\text{aux},(1),[N_n^2]}(N_n^2 t_2) \right)^2 + K_1 \left( y_{1,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{1,i}}^{\text{aux},(1),[N_n^2]}(N_n^2 t_2) \right)^2 + K_2 \left( y_{2,1,i}^{\text{aux},n}(t_n) - \Theta_{y_{2,i}}^{\text{aux},(1),[N_n^2]}(N_n^2 t_2) \right)^2 \right] = 0.$$

If we can prove that

$$(11.143) \quad \lim_{n \rightarrow \infty} \mathcal{L}[z_i^{\text{aux},n}(t_n)] = \Gamma_{\bar{\Theta}(t_2)}^{(1)},$$

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

$$(11.144) \quad \left| \mathbb{E}[f(\Theta_i^{\text{aux},(1),[N^2]}(N^2 t_2))] - \mathbb{E}^{\Gamma_{\bar{\Theta}(t_2)}^{(1)}}[f] \right| \leq \left| \mathbb{E}[f(\Theta_i^{\text{aux},(1),[N^2]}(N^2 t_2)) - f(z_{1,i}^{\text{aux},n}(t_n))] \right| + \left| \mathbb{E}[f(z_{1,i}^{\text{aux},n}(t_n))] - \mathbb{E}^{\Gamma_{\bar{\Theta}(t_2)}^{(1)}}[f] \right|.$$

Therefore if (11.143) holds, then for all  $\epsilon > 0$  we can choose  $\bar{n}$  such that, for all  $n > \bar{n}$ ,

$$(11.145) \quad \left| \mathbb{E}[f(z_{1,i}^{\text{aux},n}(t_n))] - \mathbb{E}^{\Gamma_{\bar{\Theta}(t_2)}^{(1)}}[f] \right| < \frac{\epsilon}{2}.$$

By (11.142) we can find a  $\hat{n} > \bar{n}$  such that for all  $n > \hat{n}$

$$(11.146) \quad \left| \mathbb{E}[f(\Theta_i^{\text{aux}}(N_n^2 t_2)) - f(z_{1,i}^{\text{aux},n}(t_{\hat{n}}))] \right| < \frac{\epsilon}{2}.$$

Using (11.143) and the fact that the Lipschitz functions are dense in  $f \in \mathcal{C}_b([0, 1] \times [0, 1]^2)$ , we obtain (11.133).

**Proof of (11.143).** We use that any two systems  $(\mathbf{z}^{\text{aux},1}(t_1))_{t_1 > 0}$  and  $(\mathbf{z}^{\text{aux},2}(t_1))_{t_1 > 0}$  evolving according to (11.28), and having the same  $y_{2,1}$ -components and the same  $\bar{\Theta}(t_2)$ , can be constructed on one probability space and can be coupled by their Brownian motions. We obtain, for a component  $i \in \mathbb{N}_0$ ,

$$(11.147) \quad \begin{aligned} & \mathbb{E}[|x_{1,i}^{\text{aux},1}(t_n) - x_{1,i}^{\text{aux},2}(t_n)| + K_1 |y_{1,1,i}^{\text{aux},1}(t_n) - y_{1,1,i}^{\text{aux},2}(t_n)| + K_2 |y_{2,1,i}^{\text{aux},1}(t_n) - y_{2,1,i}^{\text{aux},2}(t_n)|] \\ &= \mathbb{E}[|x_{1,i}^{\text{aux},1}(0) - x_{1,i}^{\text{aux},2}(0)| + K_1 |y_{1,1,i}^{\text{aux},1}(0) - y_{1,1,i}^{\text{aux},2}(0)| + K_2 |y_{2,1,i}^{\text{aux},1}(0) - y_{2,1,i}^{\text{aux},2}(0)|] \\ &- c \int_0^{t_n} \mathbb{E}[|x_{1,i}^{\text{aux},1}(s) - x_{1,i}^{\text{aux},2}(s)|] ds \\ &- 2K_1 e_1 \int_0^{t_n} \mathbb{E}[|x_{1,i}^{\text{aux},1}(s) - x_{1,i}^{\text{aux},2}(s)| + K_1 |y_{1,1,i}^{\text{aux},1}(s) - y_{1,1,i}^{\text{aux},2}(s)| \\ &\quad \times \mathbf{1}_{\{\text{sign}(x_{1,i}^{\text{aux},1}(s) - x_{1,i}^{\text{aux},2}(s)) \neq \text{sign}(y_{1,1,i}^{\text{aux},1}(s) - y_{1,1,i}^{\text{aux},2}(s))\}}] ds. \end{aligned}$$

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

$$(11.148) \quad (\mathcal{L}[z_i^{\text{aux},n}(0)])_{n \in \mathbb{N}}$$

is tight. Therefore we can find converging subsequences such that

$$(11.149) \quad \lim_{k \rightarrow \infty} \mathcal{L}[z_i^{\text{aux},n_k}(0)] = \mu$$

for some probability measure  $\mu$  on  $[0, 1] \times [0, 1]^2$ .

Let  $(z^{\text{aux},0}(t_1))_{t_1 > 0}$  be the limiting system evolving according to (11.28) and starting from initial distribution  $\mu$ . By Skorohod's theorem, we can construct the sequence of limiting systems  $((z^{\text{aux},n_k}(t_1))_{t_1 > 0})_{k \in \mathbb{N}}$  and  $(z^{\text{aux},0}(t_1))_{t_1 > 0}$  on one probability space such that

$$(11.150) \quad \lim_{k \rightarrow \infty} z^{\text{aux},n_k}(0) = z^{\text{aux},0}(0) \quad a.s.$$

Use the coupling of Brownian motions to obtain

$$(11.151) \quad \begin{aligned} & \mathbb{E}[|x_{1,i}^{\text{aux},n_k}(t_{n_k}) - x_{1,i}^{\text{aux},0}(t_{n_k})| + K_1 |y_{1,1,i}^{\text{aux},n_k}(t_{n_k}) - y_{1,1,i}^{\text{aux},0}(t_{n_k})| \\ & \quad + K_2 |y_{2,1,i}^{\text{aux},n_k}(t_{n_k}) - y_{2,1,i}^{\text{aux},0}(t_{n_k})|] \\ & = \mathbb{E}[|x_{1,i}^{\text{aux},n_k}(0) - x_{1,i}^{\text{aux},0}(0)| + K_1 |y_{1,1,i}^{\text{aux},n_k}(0) - y_{1,1,i}^{\text{aux},0}(0)| + K_2 |y_{2,1,i}^{\text{aux},n_k}(0) - y_{2,1,i}^{\text{aux},0}(0)|] \\ & - c \int_0^{t_{n_k}} \mathbb{E}[|x_{1,i}^{\text{aux},n_k}(s) - x_{1,i}^{\text{aux},0}(s)|] ds \\ & - 2K_1 e_1 \int_0^{t_{n_k}} \mathbb{E}[|x_{1,i}^{\text{aux},n_k}(s) - x_{1,i}^{\text{aux},0}(s)| + K_1 |y_{1,1,i}^{\text{aux},n_k}(s) - y_{1,1,i}^{\text{aux},0}(s)| \\ & \quad \times \mathbf{1}_{\{\text{sign}(x_{1,i}^{\text{aux},n_k}(s) - x_{1,i}^{\text{aux},0}(s)) \neq \text{sign}(y_{1,1,i}^{\text{aux},n_k}(s) - y_{1,1,i}^{\text{aux},0}(s))\}}] ds. \end{aligned}$$

Taking the limit  $k \rightarrow \infty$  on both sides of (11.151), we obtain

$$(11.152) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathbb{E}[|x_{1,i}^{\text{aux},n_k}(t_{n_k}) - x_{1,i}^{\text{aux},0}(t_{n_k})| + K_1 |y_{1,1,i}^{\text{aux},n_k}(t_{n_k}) - y_{1,1,i}^{\text{aux},0}(t_{n_k})| \\ & \quad + K_2 |y_{2,1,i}^{\text{aux},n_k}(t_{n_k}) - y_{2,1,i}^{\text{aux},0}(t_{n_k})|] = 0. \end{aligned}$$

Note that  $\lim_{n \rightarrow \infty} t_n = \infty$  implies that  $\lim_{k \rightarrow \infty} t_{n_k} = \infty$ , so  $z_i^{\text{aux},0}$  is the limiting system in (11.28) with  $\theta$  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 (11.80), to obtain

$$(11.153) \quad \lim_{k \rightarrow \infty} \mathcal{L}[z_i^{\text{aux},0}(t_{n_k})] = \int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv).$$

Hence we conclude that

$$(11.154) \quad \lim_{k \rightarrow \infty} \mathcal{L}[z_i^{\text{aux},n_k}(t_{n_k})] = \int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv).$$

Equation (11.154) holds for all subsequences along which the initial distribution converges,

$$(11.155) \quad \lim_{k \rightarrow \infty} z_i^{\text{aux},n_k}(0) = z_i^{\text{aux},0}(0) \quad a.s.$$

We will show that this implies (11.143).

Suppose that

$$(11.156) \quad \lim_{n \rightarrow \infty} \mathcal{L}[z_i^{\text{aux},n}(t_n)] \neq \int_{[0,1]^2} \int_{[0,1]} \Gamma_{u,y_{2,1}}^{\text{aux},(1)} P^{(u,v)}(dy_{2,1}) P_{t_2}(du, dv).$$

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

$$(11.157) \quad \left| \mathbb{E}[f(z_i^{\text{aux},n}(t_n))] - \mathbb{E}^{\Gamma_{\bar{\Theta}(t_2)}^{(1)}}[f] \right| > \delta.$$

Hence we can construct a subsequence  $(z_i^{\text{aux}, n_k}(t_1))_{t_1 > 0, k \in \mathbb{N}}$  such that (11.157) holds for each  $k \in \mathbb{N}$ . However, also for this sequence  $(\mathcal{L}[z_i^{\text{aux}, n_k}(0)])_{k \in \mathbb{N}}$  is tight. Passing to a possibly further subsequence of converging initial distributions, we argue like before to obtain that along this subsequence

$$(11.158) \quad \lim_{k \rightarrow \infty} \left| \mathbb{E}[f(z_i^{\text{aux}, n_k}(t_{n_k}))] - \mathbb{E}^{\Gamma_{\Theta(t_2)}^{(1)}}[f] \right| = 0.$$

This contradicts (11.157) and so (11.143) is indeed true.  $\square$

### Proof of Proposition 11.13

*Proof.* Lemma 11.20 implies (11.102). Therefore Lemma 11.18 implies (11.103).  $\square$

#### 11.3.7 Convergence of 2-block process

In this section we derive the limiting evolution of the effective 2-block process.

**Lemma 11.21. [Convergence of the 2-block averages]** *Assume that  $(N_k)_{k \in \mathbb{N}} \subset \mathbb{N}$  is a subsequence satisfying*

$$(11.159) \quad \begin{aligned} & \lim_{k \rightarrow \infty} \mathcal{L} \left[ y_{2,1}^{[N_k^2]}(N_k t_2) \middle| \Theta^{(2), [N_k^2]}(N_k^2 t_2) \right] = P^{z_2}(t_2), \\ & \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( Y_{1,0}^{[N_k^2]}(N_k^2 t_2 + N_k t_1), Y_{2,0}^{[N_k^2]}(N_k^2 t_2) \right) \middle| \Theta^{\text{aux}, (1), [N_k^2]}(N_k^2 t_2 + N_k t_1) \right] \\ & \quad = P^{z_1^{\text{eff}}}(t_1). \end{aligned}$$

Then, for the effective 2-block estimator process defined in (11.23),

$$(11.160) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff}, (2), [N_k^2]}(N_k^2 t_2) \right)_{t_2 > 0} \right] = \mathcal{L} \left[ (z_2^{\text{eff}}(t_2))_{t_2 > 0} \right],$$

where the limit is determined by the unique solution of the SSDE (11.35) with initial state

$$(11.161) \quad z_2^{\text{eff}}(0) = (x_2^{\text{eff}}(0), y_2^{\text{eff}}(0)) = (\vartheta_1, \theta_{y_2}).$$

*Proof.* Again we use [JM86, Theorem 3.3.1]. By a similar argument as used in the proof of Lemma 10.15 we can show that

$$(11.162) \quad \lim_{t_2 \downarrow 0} \mathcal{L} \left[ \Theta^{\text{eff}, (2), [N_k^2]}(N_k^2 t_2) \right] = \delta_{(\vartheta_1, \theta_{y_2})}.$$

Note that by steps 1-4 of the scheme in Section 11.2 we can choose the subsequence  $(N_k)_{k \in \mathbb{N}}$  such that both (11.80) and (11.101) hold. Since we already established the tightness of the 2-block in Lemma 11.4, we are left to show that, for all  $t_2 > 0$ ,

$$(11.163) \quad \begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left[ G_{\dagger}^{(2), [N_k^2]}(f, \Theta^{\text{eff}, (2), [N_k^2]}(N_k^2 t_2), t_2, \omega) \right. \right. \\ & \quad \left. \left. - G^{(2)} f \left( \Theta^{\text{eff}, (2), [N_k^2]}(N_k^2 t_2) \right) \right] \right] = 0, \end{aligned}$$

where  $G_{\dagger}^{(2), [N_k^2]}$  is the  $\mathcal{D}$ -semi-martingale operator defined in (11.68),  $G^{(2)}$  is the generator of the process  $(z_2^{\text{eff}}(t_2))_{t_2 > 0}$  defined in (11.35), and both generators work on a probability space driven

by one set of Brownian motions. Note that, for all  $t_2 > 0$ ,

$$\begin{aligned}
& \mathbb{E} \left[ \left| G_{\dagger}^{(2), [N_k^2]} \left( f, \left( \bar{\Theta}^{(2), [N_k^2]}(N_k^2 t_2), \Theta_{y_2}^{(2), [N_k^2]}(N_k^2 t_2) \right), t_2, \omega \right) \right. \right. \\
& \quad \left. \left. - G^{(2)} f \left( \bar{\Theta}^{(2), [N_k^2]}(N_k^2 t_2), \Theta_{y_2}^{(2), [N_k^2]}(N_k^2 t_2) \right) \right| \right] \\
(11.164) \quad & \leq \frac{K_2 e_2}{1 + K_0 + K_1} \mathbb{E} \left[ \left| \bar{\Theta}^{(2), [N_k^2]}(N_k^2 t_2) - \Theta_x^{(2), [N_k^2]}(N_k^2 t_2, \omega) \right| \left| \frac{\partial f}{\partial x} \right| \right] \\
& \quad + e_2 \mathbb{E} \left[ \left| \Theta_x^{(2), [N_k^2]}(N_k^2 t_2, \omega) - \bar{\Theta}^{(2), [N_k^2]}(N_k^2 t_2) \right| \left| \frac{\partial f}{\partial y} \right| \right] \\
& \quad + \frac{1}{(1 + K_0 + K_1)^2} \\
& \quad \times \mathbb{E} \left[ \left| \frac{1}{N_k^2} \sum_{i \in [N_k^2]} g(x_i^{[N_k^2]}(N_k^2 t_2, \omega)) - (\mathcal{F}^{(2)} g)(\bar{\Theta}^{(2), [N_k^2]}(N_k^2 t_2)) \right| \left| \frac{\partial^2 f}{\partial x^2} \right| \right].
\end{aligned}$$

The first and second term on the right-hand side tend to 0 as  $k \rightarrow \infty$  by a similar argument as used in (11.128) 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  $Z^{[N_k^2]}(N_k^2 t_2 - L(N) + t)_{t \geq 0}$  and the limiting system  $(z^{\nu_{\bar{\Theta}_i^{(1)}}}(t))_{t > 0}$  on one probability space, such that by Skorohod'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

$$\begin{aligned}
& \mathbb{E} \left[ \left| \frac{1}{N_k^2} \sum_{i \in [N_k^2]} g(x_i^{[N_k^2]}(N_k^2 t_2, \omega)) - (\mathcal{F}^{(2)} g)(\bar{\Theta}^{(2), [N_k^2]}(N_k^2 t_2)) \right| \left| \frac{\partial^2 f}{\partial x^2} \right| \right] \\
& \leq \frac{1}{N_k} \sum_{i \in [N]} \mathbb{E} \left[ \left| \frac{1}{N_k} \sum_{j \in [N_k^2]} g(x_j^{[N_k^2]}(N_k^2 t_2, \omega)) - \frac{1}{N_k} \sum_{j \in [N_k^2]} g(x_j^{\nu_{\bar{\Theta}_i^{(1)}}}(L(N))) \right| \left| \frac{\partial^2 f}{\partial x^2} \right| \right] \\
(11.165) \quad & + \frac{1}{N_k} \sum_{i \in [N]} \mathbb{E} \left[ \left| \frac{1}{N_k} \sum_{j \in [N_k^2]} g(x_j^{\nu_{\bar{\Theta}_i^{(1)}}}(L(N))) - (\mathcal{F}^{(1)} g)(\bar{\Theta}_i^{(1)}) \right| \left| \frac{\partial^2 f}{\partial x^2} \right| \right] \\
& + \mathbb{E} \left[ \left| \frac{1}{N_k} \sum_{i \in [N]} (\mathcal{F}^{(1)} g)(\bar{\Theta}_i^{(1)}) - (\mathcal{F}^{(2)} g)(\bar{\Theta}_i^{(2)}) \right| \left| \frac{\partial^2 f}{\partial x^2} \right| \right] \\
& + \mathbb{E} \left[ \left| (\mathcal{F}^{(2)} g)(\bar{\Theta}_i^{(2)}) - (\mathcal{F}^{(2)} g)(\bar{\Theta}_i^{(2)}(N_k^2 t_2)) \right| \left| \frac{\partial^2 f}{\partial x^2} \right| \right].
\end{aligned}$$

The first term on the right-hand side tends to zero by Lipschitz continuity for  $g$  and Corollary 11.12. 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  $\bar{\Theta}_i^{(1)}$ . The third term tends to zero since by Proposition 11.13 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 10.11 that

$$(11.166) \quad \lim_{k \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}_i^{(2)} - \bar{\Theta}_i^{(2)}(N_k^2 t_2) \right| \right] = 0.$$

Hence the last term tends to zero by the Lipschitz property of  $\mathcal{F}^{(2)} g$ .  $\square$

**Remark 11.22.** Instead of [JM86, Theorem 3.3.1] we could have used a similar strategy as in the proof of Lemma 11.18 to obtain Lemma 11.21.  $\blacksquare$

### 11.3.8 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 11.23.** [Slow seed-banks] Fix  $t_2, t_1 > 0$ , for  $i \in [N^2]$  and all  $t_0 \geq 0$ ,

$$(11.167) \quad \lim_{N \rightarrow \infty} \left[ y_{i,1}^{[N^2]}(N^2 t_2 + N t_1 + t_0) - \Theta_{y_{i,1}^{[N^2]}}^{(1),[N^2]}(N^2 t_2 + N t_1 + t_0) \right] = 0 \quad a.s.,$$

where  $\Theta_{y_{i,1}^{[N^2]}}^{(1),[N^2]}$  is the 1-block average to which  $y_{i,1}^{[N^2]}$  contributes.

To prove Lemma 11.23, we need the kernel  $b^{[N^2]}(\cdot, \cdot)$  defined in 2.30, which becomes in the current setting

$$(11.168) \quad b^{[N^2]}((i, R_i), (j, R_j)) = \begin{cases} 1_{\{d_{[N^2]}(i,j) \leq 1\}} \frac{c_0}{N} + \frac{c_1}{N^3}, & \text{if } R_i = R_j = A, \\ K_m \frac{e_m}{N^m}, & \text{if } i = j, R_i = A, R_j = D_m, m \in \{0, 1, 2\}, \\ \frac{e_m}{N^m}, & \text{if } i = j, R_i = D_m, R_j = A, m \in \{0, 1, 2\}, \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding semigroup of the kernel  $b^{[N^2]}(\cdot, \cdot)$  is denoted by  $b_t^{[N^2]}(\cdot, \cdot)$ .

To prove Lemma 11.23 we will use the following lemma, which was proved in [GHO20][Lemma 6.1] and for our setting reads as follows.

**Lemma 11.24.** [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  $(i, R_i), (j, R_j) \in [N^2] \times \{A, D_0, D_1, D_2\}$ ,

$$(11.169) \quad \mathbb{E}_{z^{[N^2]}}[z_{(i, R_i)}^{[N^2]}(t)] = \sum_{\substack{(k, R_k) \in \\ \Omega_N \times \{A, D_0, D_1, D_2\}}} b_t^{[N^2]}((i, R_i), (k, R_k)) z_{(k, R_k)}^{[N^2]}$$

and

$$(11.170) \quad \begin{aligned} & \mathbb{E}_{z^{[N^2]}}[z_{(i, R_i)}^{[N^2]}(t) z_{(j, R_j)}^{[N^2]}(t)] \\ &= \sum_{\substack{(k, R_k), (l, R_l) \in \\ \Omega_N \times \{A, D_0, D_1, D_2\}}} b_t^{[N^2]}((i, R_i), (k, R_k)) b_t^{[N^2]}((j, R_j), (l, R_l)) z_{(k, R_k)}^{[N^2]} z_{(l, R_l)}^{[N^2]} \\ &+ 2 \int_0^t ds \sum_{k \in \Omega_N} b_{(t-s)}^{[N^2]}((i, R_i), (k, A)) b_{(t-s)}^{[N^2]}((j, R_j), (k, A)) \mathbb{E}_z^{[N^2]}[g(x_k^{[N^2]}(s))]. \end{aligned}$$

*Proof of Lemma 11.23.* 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

$$(11.171) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left( y_{i,1}^{[N^2]}(\bar{t}(N)) - y_{j,1}^{[N^2]}(\bar{t}(N)) \right)^2 \right] = 0.$$

This implies (11.167). By Lemma 11.24, we can write

$$\begin{aligned}
& \mathbb{E} \left[ \left( y_{i,1}^{[N^2]}(\bar{t}(N)) - y_{j,1}^{[N^2]}(\bar{t}(N)) \right)^2 \right] \\
&= \sum_{\substack{(k,R_k),(l,R_l) \in \\ [N^2] \times \{A,D_0,D_1,D_2\}}} \left( b_{\bar{t}(N)}^{[N^2]}((i,D_1),(k,R_k)) - b_{\bar{t}(N)}^{[N^2]}((j,D_1),(k,R_k)) \right) \\
(11.172) \quad & \times \left( b_{\bar{t}(N)}^{[N^2]}((i,D_1),(l,R_l)) - b_{\bar{t}(N)}^{[N^2]}((j,D_1),(l,R_l)) \right) \mathbb{E}[z_{(k,R_k)}^{[N^2]} z_{(l,R_l)}^{[N^2]}] \\
& + 2 \int_0^{\bar{t}(N)} ds \sum_{k \in [N^2]} \left( b_{\bar{t}(N)-s}^{[N^2]}((i,D_1),(k,A)) - b_{\bar{t}(N)-s}^{[N^2]}((j,D_1),(k,A)) \right)^2 \\
& \times \mathbb{E}[g(x_k^{[N^2]}(s))].
\end{aligned}$$

Using a coupling argument, we show that both terms in (11.172) tend to 0 as  $N \rightarrow \infty$ . To prove that the first term tends to 0, we will show that

$$(11.173) \quad \lim_{N \rightarrow \infty} \sum_{\substack{(k,R_k) \in \\ [N^2] \times \{A,D_0,D_1,D_2\}}} \left| b_{\bar{t}(N)}^{[N^2]}((i,D_1),(k,R_k)) - b_{\bar{t}(N)}^{[N^2]}((j,D_1),(k,R_k)) \right| = 0.$$

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  $b_t^{[N^2]}(\cdot, \cdot)$  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

$$(11.174) \quad ((i, R_i), (j, R_j)) \rightarrow \begin{cases} ((k, A), (l, A)) & \text{if } R_i = R_j = A \text{ and } d_{[N^2]}(i, k) = d_{[N^2]}(j, l) \\ & \text{at rate } 1_{\{d_{[N^2]}(i,j) \leq 1\}} \frac{c_0}{N} + \frac{c_1}{N^3}, \\ ((i, D_m), (j, D_m)) & \text{if } R_i = R_j = A \text{ at rate } \frac{K_m c_m}{N^m}, m \in \{0, 1, 2\}, \\ ((i, A), (j, A)) & \text{if } R_i = R_j = D_m \text{ at rate } \frac{c_m}{N^m}, m \in \{0, 1, 2\}. \end{cases}$$

Define the event

$$(11.175) \quad H_t^{[N^2]} = \{RW^{[N^2]} \text{ has migrated at least once up to time } t\}.$$

Note that if  $H_t^{[N^2]}$  has happened, then also  $RW'^{[N^2]}$  has migrated. Hence

$$\begin{aligned}
& b_{\bar{t}(N)}^{[N^2]}((i, D_1), (k, R_k)) - b_{\bar{t}(N)}^{[N^2]}((j, D_1), (k, R_k)) \\
&= \mathbb{P}_{(i, D_1)} \left( RW^{[N^2]}(\bar{t}(N)) = (k, R_k) \right) - \mathbb{P}_{(j, D_1)} \left( RW'^{[N^2]}(\bar{t}(N)) = (k, R_k) \right) \\
&= \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW^{[N^2]}(\bar{t}(N)) = (k, R_k), H_{\bar{t}(N)}^{[N^2]} \right) \\
(11.176) \quad & + \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW^{[N^2]}(\bar{t}(N)) = (k, R_k), (H_{\bar{t}(N)}^{[N^2]})^c \right) \\
& - \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW'^{[N^2]}(\bar{t}(N)) = (k, R_k), H_{\bar{t}(N)}^{[N^2]} \right) \\
& - \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW'^{[N^2]}(\bar{t}(N)) = (k, R_k), (H_{\bar{t}(N)}^{[N^2]})^c \right) \\
&= \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW^{[N^2]}(\bar{t}(N)) = (k, R_k), (H_{\bar{t}(N)}^{[N^2]})^c \right) \\
& - \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW'^{[N^2]}(\bar{t}(N)) = (k, R_k), (H_{\bar{t}(N)}^{[N^2]})^c \right),
\end{aligned}$$

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 over  $[N^2] \times A$  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

$$(11.177) \quad \sum_{\substack{(k, R_k) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \left| b_{\bar{t}(N)}^{[N^2]}((i, D_1), (k, R_k)) - b_{\bar{t}(N)}^{[N^2]}((j, D_1), (k, R_k)) \right| \leq 2\tilde{\mathbb{P}}((H_{\bar{t}(N)}^{[N^2]})^c)$$

and we are left to show that

$$(11.178) \quad \lim_{N \rightarrow \infty} \tilde{\mathbb{P}}((H_{\bar{t}(N)}^{[N^2]})^c) = 0.$$

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

$$(11.179) \quad \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)} \left( RW^{[N^2]} \text{ does not wake up before } \bar{t}(N) \right) = e^{-\frac{e_1}{N} \bar{t}(N)} = e^{-\frac{e_1(N^2 t_2 + N t_1 + t_0)}{N}}$$

and hence

$$(11.180) \quad \lim_{N \rightarrow \infty} \mathbb{P}_{(i, D_1), (j, D_1)} \left( RW^{[N^2]} \text{ does not wake up before } \bar{t}(N) \right) = 0.$$

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

$$(11.181) \quad \begin{aligned} C^{[N^2]}(t) &= \{\# \text{ times } RW^{[N^2]} \text{ gets active before time } t\}, \\ T_A^{[N^2]}(t) &= \{\text{total time } RW^{[N^2]} \text{ is active up to time } t\}, \\ T_D^{[N^2]}(t) &= \{\text{total time } RW^{[N^2]} \text{ is dormant up to time } t\}. \end{aligned}$$

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  $n$ th cycle. Define

$$(11.182) \quad \chi = K_0 e_0 + \frac{K_1 e_1}{N} + \frac{K_2 e_2}{N^2},$$

so  $\chi$  is the total rate at which  $RW$  and  $RW'$  become dormant when they are active. Define

$$(11.183) \quad c = c_0 + \frac{c_1}{N},$$

so  $c$  is the total rate at which  $RW$  and  $RW'$  migrate when they are active. Then

$$(11.184) \quad T_A^{[N^2]}(t) = \sum_{n=1}^{C^{[N^2]}(t)} T_{A,n}^{[N^2]}, \quad T_D^{[N^2]}(t) = \sum_{n=1}^{C^{[N^2]}(t)} T_{D,n}^{[N^2]},$$

where  $T_{A,n}^{[N^2]} \stackrel{d}{=} \exp(\chi)$  and  $T_{D,n}^{[N^2]} \stackrel{d}{=} \frac{1}{\chi} K_0 e_0 \exp(e_0) + \frac{1}{\chi} \frac{K_1 e_1}{N} \exp(\frac{e_1}{N}) + \frac{1}{\chi} \frac{K_2 e_2}{N^2} \exp(\frac{e_2}{N^2})$ . Once awake,  $RW^{[N^2]}$  migrates at rate  $c$  and hence the probability to migrate before time  $\bar{t}(N)$  is given by  $1 - e^{-c T_A^{[N^2]}(\bar{t}(N))}$ . Therefore we are left to show that

$$(11.185) \quad \lim_{N \rightarrow \infty} c T_A^{[N^2]}(\bar{t}(N)) = \lim_{N \rightarrow \infty} c \sum_{n=1}^{C^{[N^2]}(\bar{t}(N))} T_{A,n}^{[N^2]} = \infty, \quad a.s.$$



Since  $T_{A,n}^{[N^2]} \stackrel{d}{=} \exp(\chi)$ , it is enough to show that

$$(11.186) \quad \lim_{N \rightarrow \infty} C^{[N^2]}(\bar{t}(N)) = \infty \quad a.s.$$

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

$$(11.187) \quad \mathbb{P}_{(i,D_1)}(C^{[N^2]}(\bar{t}(N)) \leq R) > 0.$$

Let  $L(N)$  be such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ . Note that, by (11.180), we can condition on the first wake-up time and estimate

$$(11.188) \quad \begin{aligned} \mathbb{P}_{(i,D_1)}(C^{[N^2]}(\bar{t}(N)) \leq R) &= \int_0^{\bar{t}(N)} ds \mathbb{P}_{(i,A)}(C^{[N^2]}(\bar{t}(N) - s) \leq R) \frac{e_1}{N} e^{-\frac{e_1}{N}s} \\ &= \int_0^{\bar{t}(N)-L(N)} ds \mathbb{P}_{(i,A)}(C^{[N^2]}(\bar{t}(N) - s) \leq R) \frac{e_1}{N} e^{-\frac{e_1}{N}s} \\ &\quad + \int_{\bar{t}(N)-L(N)}^{\bar{t}(N)} ds \mathbb{P}_{(i,A)}(C^{[N^2]}(\bar{t}(N) - s) \leq R) \frac{e_1}{N} e^{-\frac{e_1}{N}s} \\ &\leq \mathbb{P}_{(i,A)}(C^{[N^2]}(L(N)) \leq R) + e^{-\frac{e_1}{N}\bar{t}(N)} \left[ e^{\frac{e_1}{N}L(N)} - 1 \right]. \end{aligned}$$

Note that the second term in the last inequality tends to 0 as  $N \rightarrow \infty$ . For the first term, note that we are now looking at time  $L(N)$ , i.e., time scale  $N^0$ . Since

$$(11.189) \quad \begin{aligned} &\lim_{N \rightarrow \infty} \mathbb{P}_{i,A} \left( RW^{[N^2]} \text{ jumps to } D_1 \text{ or } D_2 \text{ before time } L(N) \right) \\ &= \lim_{N \rightarrow \infty} 1 - e^{-\left(\frac{K_1 e_1}{N} + \frac{K_2 e_2}{N^2}\right)L(N)} = 0. \end{aligned}$$

we have  $\lim_{N \rightarrow \infty} \mathbb{P}_{i,A} \left( \{RW^{[N^2]}(s) \in \{A, D_0\} \text{ for } s \in [0, L(N)]\} \right) = 1$ . Hence, conditioned on the event  $\{RW^{[N^2]} \in \{A, D_0\}\}$ ,  $T_{A,n}^{[N^2]} \stackrel{d}{=} \exp(K_0 e_0)$  and  $T_{D,n}^{[N^2]} \stackrel{d}{=} \exp(e_0)$ . We therefore obtain

$$(11.190) \quad \begin{aligned} \mathbb{P}_{(i,A)}(C^{[N^2]}(L(N)) \leq R) &= \mathbb{P}_{(i,A)} \left( \sum_{n=1}^R (T_{A,n}^{[N^2]} + T_{D,n}^{[N^2]}) \geq L(N) \right) \\ &\leq \frac{R}{L(N)} \mathbb{E}_{(i,A)} \left[ T_{A,n}^{[N^2]} + T_{D,n}^{[N^2]} \right] \\ &= \frac{R}{L(N)} \left[ \frac{1}{K_0 e_0} + \frac{1}{e_0} \right]. \end{aligned}$$

Taking the limit  $N \rightarrow \infty$  in (11.190) and combining this with (11.188), we conclude that (11.186) indeed holds. Hence also (11.178) and (11.173) hold.

We are left to show that

$$(11.191) \quad \begin{aligned} &\lim_{N \rightarrow \infty} 2 \int_0^{\bar{t}(N)} ds \sum_{k \in [N^2]} \left( b_{(\bar{t}(N)-s)}^{[N^2]}((i, D_1), (k, A)) - b_{(\bar{t}(N)-s)}^{[N^2]}((j, D_1), (k, A)) \right)^2 \\ &\quad \times \mathbb{E}[g(x_k(s))] = 0. \end{aligned}$$

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 (11.191) tends to 0 as  $N \rightarrow \infty$ .

Note that, by (11.172),

$$(11.192) \quad \begin{aligned} &\left| 2 \int_0^{\bar{t}(N)} ds \sum_{k \in [N^2]} \left( b_{(\bar{t}(N)-s)}^{[N^2]}((i, D_1), (k, A)) - b_{(\bar{t}(N)-s)}^{[N^2]}((j, D_1), (k, A)) \right)^2 \right. \\ &\quad \left. \times \mathbb{E}[g(x_k^{[N^2]}(s))] \right| \leq 2. \end{aligned}$$

We will again use the coupling in (11.175). Define

$$(11.193) \quad \tau^{[N^2]} = \inf\{t \geq 0 : RW^{[N^2]}(t) = (k, A) \text{ for some } k \in [N^2]\}.$$

Then, for all  $l \in [N^2]$ ,

$$(11.194) \quad \mathbb{P}_{(l, D_1)}(\tau^{[N^2]} \leq t) = 1 - e^{-\frac{\varepsilon_1}{N}t}.$$

Setting  $s = \bar{t}(N) - s$ , we can rewrite the integral in (11.191) as

$$(11.195) \quad \begin{aligned} & 2 \int_0^{\bar{t}(N)} ds \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)}^{[N^2]} \left( RW^{[N^2]}(s) = (k, A) \right) \right. \\ & \quad \left. - \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)}^{[N^2]} \left( RW'^{[N^2]}(s) = (k, A) \right) \right)^2 \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N) - s))] \\ & = 2 \int_0^{\bar{t}(N)} ds \sum_{k \in [N^2]} \left[ \int_0^s dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \right. \\ & \quad \left. \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(s - r) = (k, A) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(s - r) = (k, A) \right) \right) \right] \\ & \quad \times \left( \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)}^{[N^2]} \left( RW^{[N^2]}(s) = (k, A) \right) - \tilde{\mathbb{P}}_{(i, D_1), (j, D_1)}^{[N^2]} \left( RW'^{[N^2]}(s) = (k, A) \right) \right) \\ & \quad \times \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N) - s))]. \end{aligned}$$

In what follows we will abbreviate

$$(11.196) \quad \begin{aligned} & P_{\Delta_{(i, A), (j, A)}^{[N^2]}}(t, ((l, R_l), (j, R_j))) \\ & = \tilde{\mathbb{P}}_{(i, A), (j, A)} \left( RW^{[N^2]}(t) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)} \left( RW'^{[N^2]}(t) = (j, R_j) \right), \end{aligned}$$

and similarly, for  $m \in \{0, 1, 2\}$ ,

$$(11.197) \quad \begin{aligned} & P_{\Delta_{(i, D_m), (j, D_m)}^{[N^2]}}(t, ((l, R_l), (j, R_j))) \\ & = \tilde{\mathbb{P}}_{(i, D_m), (j, D_m)} \left( RW^{[N^2]}(t) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i, D_m), (j, D_m)} \left( RW'^{[N^2]}(t) = (j, R_j) \right). \end{aligned}$$

By (11.192), we can use Fubini to swap the order of integration and subsequently substitute  $v = s - r$ , to obtain

$$(11.198) \quad \begin{aligned} & 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \int_0^{\bar{t}(N) - r} dv \sum_{k \in [N^2]} P_{\Delta_{(i, A), (j, A)}^{[N^2]}}(v, ((k, A), (k, A))) \\ & \quad \times P_{\Delta_{(i, D_1), (j, D_1)}^{[N^2]}}(v + r, ((k, A), (k, A))) \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N) - r - v))] \\ & = 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \\ & \quad \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \left( \mathbb{P}_{(i, D_1)} \left( RW^{[N^2]}(r) = (l, R_l) \right) - \mathbb{P}_{(j, D_1)} \left( RW'^{[N^2]}(r) = (l, R_l) \right) \right) \\ & \quad \times \int_0^{\bar{t}(N) - r} dv \sum_{k \in [N^2]} P_{\Delta_{(i, A), (j, A)}^{[N^2]}}(v, ((k, A), (k, A))) \\ & \quad \times \mathbb{P}_{(l, R_l)} \left( RW^{[N^2]}(v) = (k, A) \right) \mathbb{E}[g(x_k^{[N^2]}(\bar{t}(N) - r - v))], \end{aligned}$$

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)}} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \\
& \quad \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \left( \mathbb{P}_{(i, D_1)}^{[N^2]} \left( RW^{[N^2]}(r) = (l, R_l) \right) - \mathbb{P}_{(j, D_1)}^{[N^2]} \left( RW^{[N^2]}(r) = (l, R_l) \right) \right) \\
& \quad \times \int_0^{\bar{t}^{(N)} - r} dv \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \mathbb{P}_{(l, R_l)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \mathbb{E}[g(x_k^{[N^2]})(\bar{t}^{(N)} - r - v)] \\
& = 2 \int_0^{\bar{t}^{(N)}} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \int_0^r du \tilde{\mathbb{P}}(\tau^{[N^2]} = u) \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \\
& \quad \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(r - u) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(r - u) = (l, R_l) \right) \right) \\
& \quad \times \int_0^{\bar{t}^{(N)} - r} dv \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \mathbb{P}_{(l, R_l)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \mathbb{E}[g(x_k^{[N^2]})(\bar{t}^{(N)} - r - v)] \\
& + 2 \int_0^{\bar{t}^{(N)}} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \quad \times \int_0^{\bar{t}^{(N)} - r} dv \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \left( \mathbb{P}_{(i, D_1)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) - \mathbb{P}_{(j, D_1)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \mathbb{E}[g(x_k^{[N^2]})(\bar{t}^{(N)} - r - v)].
\end{aligned} \tag{11.199}$$

We will show that both terms in the last equality of (11.199) tends to 0 as  $N \rightarrow \infty$ .

For the first term note that, by (11.170), (11.172) and (11.194), we have

$$\begin{aligned}
& 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \int_0^r du \tilde{\mathbb{P}}(\tau^{[N^2]} = u) \\
& \quad \times \left[ \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(r - u) = (l, R_l) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(r - u) = (l, R_l) \right) \right] \\
& \quad \times \int_0^{\bar{t}(N) - r} dv \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \mathbb{P}_{(l, R_l)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) \mathbb{E}[g(x_k(\bar{t}(N) - r - v))] \\
& \leq 4 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \int_0^r du \tilde{\mathbb{P}}(\tau^{[N^2]} = u) \\
& \quad \times \left| \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(r - u) = (l, R_l) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(r - u) = (l, R_l) \right) \right| \\
(11.200) \quad & \leq 4 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \int_0^{r - L(N)} du \tilde{\mathbb{P}}(\tau^{[N^2]} = u) \\
& \quad \times \left| \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(r - u) = (l, R_l) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(r - u) = (l, R_l) \right) \right| \\
& + 8 \int_{L(N)}^{\bar{t}(N)} dr \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)]] \\
& \leq 4 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \int_0^{r - L(N)} du \tilde{\mathbb{P}}(\tau^{[N^2]} = u) \\
& \quad \times \left| \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(r - u) = (l, R_l) \right) \right. \\
& \quad \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(r - u) = (l, R_l) \right) \right| \\
& + 16[1 - e^{-\frac{c_1}{N}L(N)}].
\end{aligned}$$

Hence the last term in the last inequality tends to 0 as  $N \rightarrow \infty$ .

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 (11.175). Note that we can rewrite the sum as

$$(11.201)$$

$$\begin{aligned}
& \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(r - u) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(r - u) = (l, R_l) \right) \\
& \leq \sum_{\substack{(l, R_l) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \sum_{\substack{(l', R_{l'}) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \left( \mathbb{P}_{(i, A)}^{[N^2]} \left( RW^{[N^2]}(L(N)) = (l', R_{l'}) \right) \right. \\
& \quad \times \mathbb{P}_{(l', R_{l'})}^{[N^2]} \left( RW^{[N^2]}(r - u - L(N)) = (l, R_l) \right) \\
& \leq \sum_{\substack{(l', R_{l'}) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \left( \mathbb{P}_{(i, A)}^{[N^2]} \left( RW^{[N^2]}(L(N)) = (l', R_{l'}) \right) - \mathbb{P}_{(j, A)}^{[N^2]} \left( RW'^{[N^2]}(L(N)) = (l', R_{l'}) \right) \right) \\
& = \sum_{\substack{(l', R_{l'}) \in \\ [N^2] \times \{A, D_0, D_1, D_2\}}} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(L(N)) = (l', R_{l'}), (H_t^{[N^2]})^c \right) \right. \\
& \quad \left. - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(L(N)) = (l', R_{l'}), (H_t^{[N^2]})^c \right) \right) \\
& \leq 2\tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( (H_t^{[N^2]})^c \right).
\end{aligned}$$

To show that

$$(11.202) \quad \lim_{n \rightarrow \infty} \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( (H_t^{[N^2]})^c \right) = 0,$$

we can use a similar strategy as between (11.188) and (11.190), but note that we now start from two active sites instead of two 1-dormant sites. Therefore (11.186) directly follows from (11.189) and (11.190).

To show the second term in (11.199) tends to 0, we write it as

$$\begin{aligned}
& 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \quad \times \int_0^{\bar{t}(N) - r} dv \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \left[ \mathbb{P}_{(i, D_1)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) - \mathbb{P}_{(j, D_1)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right] \mathbb{E}[g(x_k^{[N^2]})(\bar{t}(N) - r - v)] \\
(11.203) \quad & = 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \quad \times \int_0^{\bar{t}(N) - r} dv \sum_{k \in [N^2]} \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right) \\
& \quad \times \int_0^v du \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \\
& \quad \left( \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW^{[N^2]}(v - u) = (k, A) \right) - \tilde{\mathbb{P}}_{(i, A), (j, A)}^{[N^2]} \left( RW'^{[N^2]}(v - u) = (k, A) \right) \right) \\
& \quad \times \mathbb{E}[g(x_k^{[N^2]})(\bar{t}(N) - r - v)].
\end{aligned}$$

Changing the order of integration and setting  $w = v - u$ , we obtain

$$\begin{aligned}
& 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \times \int_0^{\bar{t}(N)-r} du \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \int_u^{\bar{t}(N)-r} dv \\
& \sum_{k \in [N^2]} \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(v) = (k, A) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(v) = (k, A) \right) \right] \\
& \times \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(v-u) = (k, A) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(v-u) = (k, A) \right) \right] \\
(11.204) \quad & \times \mathbb{E}[g(x_k^{[N^2]})(\bar{t}(N) - r - v)] \\
& = 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \times \int_0^{\bar{t}(N)-r} du \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \int_0^{\bar{t}(N)-r-u} dw \\
& \sum_{k \in [N^2]} \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(w+u) = (k, A) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(w+u) = (k, A) \right) \right] \\
& \times \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(w) = (k, A) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(w) = (k, A) \right) \right] \\
& \times \mathbb{E}[g(x_k^{[N^2]})(\bar{t}(N) - r - u - w)].
\end{aligned}$$

This can be rewritten as

$$\begin{aligned}
& 2 \int_0^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \int_0^{\bar{t}(N)-r} du \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \\
& \times \sum_{(l, R_l) \in [N^2]} \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(u) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(u) = (l, R_l) \right) \right] \\
& \times \int_0^{\bar{t}(N)-r-u} dw \sum_{k \in [N^2]} \left[ \tilde{\mathbb{P}}_{(l, R_l)}^{[N^2]} \left( RW^{[N^2]}(w) = (k, A) \right) \right] \\
& \times \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(w) = (k, A) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(w) = (k, A) \right) \right] \\
& \times \mathbb{E}[g(x_k^{[N^2]})(\bar{t}(N) - r - u - w)] \\
(11.205) \quad & \leq 8 \int_{L(N)}^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \int_{L(N)}^{\bar{t}(N)-r} du \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \\
& \times \sum_{(l, R_l) \in [N^2]} \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(u) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(u) = (l, R_l) \right) \right] \\
& + 8 \int_{L(N)}^{\bar{t}(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r) \\
& \int_0^{L(N)} du \tilde{\mathbb{P}}^{[N^2]}(\tau^{[N^2]} = u) \\
& \times \sum_{(l, R_l) \in [N^2]} \left[ \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW^{[N^2]}(u) = (l, R_l) \right) - \tilde{\mathbb{P}}_{(i,A),(j,A)}^{[N^2]} \left( RW'^{[N^2]}(u) = (l, R_l) \right) \right] \\
& + 16 \int_0^{L(N)} dr \tilde{\mathbb{P}}(\tau^{[N^2]} = r) \tilde{\mathbb{P}}(\tau^{[N^2]} \geq r).
\end{aligned}$$

This tends to 0 by (11.201) and the reasoning below (11.202).  $\square$

### 11.3.9 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  $N \rightarrow \infty$ . Recall that Lemma 11.23 tells us that all single 1-dormant colonies equal the value of the 1-dormant 1-block average. Therefore the second assumption in (11.80) in Proposition (11.8) can be replaced by

$$(11.206) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( Y_{2,0}^{[N_k^2]}(N_k^2 t_2) \right) \middle| \Theta^{\text{aux},(1),[N_k^2]}(N_k^2 t_2 + N_k t_1) \right] = P^{z_1^{\text{eff}}(t_1)},$$

since, by Lemma 11.23, the limiting law

$$(11.207) \quad \lim_{k \rightarrow \infty} \mathcal{L} \left[ \left( Y_{1,0}^{[N_k^2]}(N_k^2 t_2) \right) \right]$$

is completely determined by the first line in (11.80). Hence, the assumptions in Proposition 11.13 and Lemma 11.21 can be weakened in the same way. Using that in Proposition 11.1 we assume (11.37) and (11.38), we find that the 2-block convergence stated in Lemma 11.21 holds along all subsequences we choose in Step 5. We conclude that Proposition 11.1(a) is indeed true. Combining Proposition 11.1(a) with steps 1-4 of the scheme and Lemma 11.23, we find that the assumptions in Proposition 11.13 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 11.1(b). Finally, combining Proposition 11.1(a), steps 1-4, and the fact that Proposition 11.13 is true along all subsequences, we obtain Proposition 11.1(c) and (f).

### 11.3.10 Convergence in the Meyer-Zheng topology

In this section we show how the results on the effective and estimator processes can be used to show convergence of the full 1- and 2-block processes.

**Lemma 11.25** ([Convergence of 1 process in the Meyer-Zheng topology]). *Assume that for the 1-block estimator process defined in (11.21)*

$$(11.208) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{aux},(1),[N^2]}(N^2 t_2 + N t_1) \right)_{t_1 > 0} \right] = \mathcal{L} [(z_1^{\text{aux}}(t_1))_{t_1 > 0}],$$

where, conditional on  $x_2^{\text{eff}}(t_2) = u$ , the limit process is the unique solution of the SSDE in (11.29) with  $\theta$  replaced by  $u$  and with initial measure  $\Gamma_u^{\text{eff},(1)}$ . Then

$$(11.209) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(1),[N^2]}(N^2 t_2 + N t_1) \right)_{t_1 > 0} \right] = \mathcal{L} [(z_1^{\Gamma^{(1)}(t_2)}(t_1))_{t_1 > 0}]$$

in the Meyer-Zheng topology,

where  $\Gamma^{(1)}(t_2)$  is defined as in (11.45) and  $(z_1^{\Gamma^{(1)}(t_2)}(t_1))_{t_1 > 0}$  is the process moving according to (11.27) with initial measure  $\Gamma^{(1)}(t_2)$ .

*Proof.* By assumption 11.208 and Lemma 11.9, we can proceed as in the proof of Proposition 10.16 to find (11.209).  $\square$

**Lemma 11.26** ([Convergence of 2-block process in the Meyer-Zheng topology]).

*Assume that for the effective 2-block process defined in (11.23)*

$$(11.210) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{eff},(2),[N^2]}(N^2 t_2) \right)_{t_2 > 0} \right] = \mathcal{L} [(z_2^{\text{aux}}(t_2))_{t_2 > 0}],$$

where  $(z_2^{\text{eff}}(t_2))_{t_2 > 0}$  is the process evolving according to (11.35) and starting from  $(\vartheta_1, \theta_{y_2})$ . Then for the 2-block estimator process defined in (11.23)

$$(11.211) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(2),[N^2]}(N^2 t_2) \right)_{t_2 > 0} \right] = \mathcal{L} [(z_2(t_2))_{t_2 > 0}]$$

in the Meyer-Zheng topology,

where  $(z_2(t_2))_{t_2 > 0}$  is the process evolving according to (11.34) and starting in state  $(\vartheta_1, \vartheta_1, \vartheta_1, \theta_{y_2})$ .

*Proof.* Combining Lemmas 11.9 and 11.16, we find for  $t_2 > 0$

$$(11.212) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \Theta_x^{(2),[N^2]}(N^2 t_2) \right| \right] &= 0, \\ \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \Theta_{y_0}^{(2),[N^2]}(N^2 t_2) \right| \right] &= 0, \\ \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(2),[N^2]}(N^2 t_2) - \Theta_{y_1}^{(2),[N^2]}(N^2 t_2) \right| \right] &= 0. \end{aligned}$$

Therefore we can again proceed as in the proof of Proposition 10.16 to find (11.211).  $\square$

### 11.3.11 Proof of the two-level three-colour mean-field finite-systems scheme

In Section 11.3.9 we already proved Proposition 11.1(a),(b),(c) and (f). The proof of Proposition 11.1(d) follows from Proposition 11.1(a) by applying Lemma 11.26. The proof of Proposition 11.1(e) follows from Proposition 11.1(b) by applying Lemma 11.25. This completes the proof of Proposition 11.1.

## 12 Proofs: $N \rightarrow \infty$ , hierarchical multi-scale limit theorems

In this section we prove the hierarchical multi-scale limit theorems stated in Theorem 4.2 and Theorem 4.4. In Section 12.1 we first introduce the finite-level mean-field finite-systems scheme. In Section 12.2 we given an outline of how to prove the finite-level mean-field finite-systems scheme. In Section 12.3 we show how Theorems 4.2–4.4 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 11.2 once we incorporate more levels. Since the proofs for the finite-level mean-field systems scheme are similar as the proofs in Section 11.3, we will not write out the full proof, but only give an outline and a sketch.

### 12.1 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  $\Omega_N^{k+1}$  obtained by truncating the hierarchical group  $\Omega_N$  (recall (2.2)) after hierarchical level  $k+1$ , i.e.,  $\Omega_N^{k+1} = B_{k+1}(0)$  the  $(k+1)$ -block centred at the origin (recall (2.2), (2.4) and Fig. 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  $\{0, \dots, k\} \cup \{k+1\}$ . On this space we again consider a restricted version of the SSDE in (2.20) to the geographic space  $\Omega_N^{k+1}$ . The migration kernel  $a^{\Omega_N}(\cdot, \cdot)$  is restricted to  $\Omega_N^{k+1}$  by setting all migration outside  $B_{k+1}(0)$  equal to 0, i.e.,

$$(12.1) \quad a^{[\Omega_N^{k+1}]}(\xi, \eta) = \sum_{l=1}^{k+1} \mathbf{1}_{\{d_{\Omega_N^{k+1}}(\xi, \eta) \leq l\}} \frac{c_l}{N^{l-1}} \frac{1}{N^l},$$

where  $d_{\Omega_N^{k+1}}$  is the hierarchical distance  $d_{\Omega_N}$  restricted to the space  $\Omega_N^{k+1}$ . The colour- $l$  dormant population exchanges individuals with the active population at rates  $\frac{c_l}{N^l}$ ,  $\frac{K_l e_l}{N^l}$  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

$$(12.2) \quad S = (\mathfrak{s}^{k+1})^{\Omega_N^{k+1}}, \quad \mathfrak{s}^{k+1} = [0, 1] \times [0, 1]^{k+2},$$



and the system is denoted by

$$(12.3) \quad \begin{aligned} Z_N^{\Omega^{k+1}} &= (Z_N^{\Omega^{k+1}}(t))_{t \geq 0}, & Z_N^{\Omega^{k+1}}(t) &= (z_\xi^{\Omega^{k+1}}(t))_{\xi \in \Omega_N^{k+1}}, \\ z_\xi^{\Omega^{k+1}}(t) &= (x_\xi^{\Omega^{k+1}}(t), (y_{\xi,m}^{\Omega^{k+1}}(t))_{m=0}^{k+1}). \end{aligned}$$

The components evolve according to the SSDE

$$(12.4) \quad \begin{aligned} dx_\xi^{\Omega^{k+1}}(t) &= \sum_{l=1}^k \frac{c_{l-1}}{N^{l-1}} \frac{1}{N^l} \sum_{\eta \in B_l(\xi)} \left[ x_\eta^{\Omega^{k+1}}(t) - x_\xi^{\Omega^{k+1}}(t) \right] dt \\ &\quad + \sqrt{g(x_\xi^{\Omega^{k+1}}(t))} dw_\xi(t) + \sum_{m=0}^k \frac{K_m e_m}{N^m} \left[ y_{\xi,m}^{\Omega^{k+1}}(t) - x_\xi^{\Omega^{k+1}}(t) \right] dt, \\ dy_{\xi,m}^{\Omega^{k+1}}(t) &= \frac{e_m}{N^m} \left[ x_\xi^{\Omega^{k+1}}(t) - y_{\xi,m}^{\Omega^{k+1}}(t) \right] dt, \quad 0 \leq m \leq k, \\ dy_{\xi,m}^{\Omega^{k+1}}(t) &= 0, \quad \xi \in \Omega_N^{k+1}, \end{aligned}$$

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 (2.20) 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 [SS80] the martingale problem associated with (12.4) is well-posed, and for every initial state in  $S$  the SSDE has a unique strong solution. We will analyse (12.4) on time scales  $1, N, N^2, \dots, N^k$ . 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 (2.3). For the  $k$ -level mean-field system these are given by

$$(12.5) \quad \begin{aligned} x_l^{\Omega^{k+1}}(t) &= \frac{1}{N^l} \sum_{\eta \in B_l(0)} x_\eta^{\Omega^{k+1}}(N^l t), \\ y_{m,l}^{\Omega^{k+1}}(t) &= \frac{1}{N^l} \sum_{\eta \in B_l(0)} y_{\eta,m}^{\Omega^{k+1}}(N^l t), \quad 0 \leq m \leq k+1, \quad 0 \leq l \leq k. \end{aligned}$$

For  $0 \leq l \leq k$  these block averages evolve according to the SSDE

$$(12.6) \quad \begin{aligned} dx_l^{\Omega^{k+1}}(t) &= \sum_{n=1}^{k-(l-1)} \frac{c_{l+n-1}}{N^{n-1}} \left[ x_{l+n}^{\Omega^{k+1}}(N^{-n}t) - x_l^{\Omega^{k+1}}(t) \right] dt \\ &\quad + \sqrt{\frac{1}{N^l} \sum_{i \in B_l(0)} g(x_i(N^l t))} dw_l(t) \\ &\quad + \sum_{m=0}^k N^l \frac{K_m e_m}{N^m} \left[ y_{m,l}^{\Omega^{k+1}}(t) - x_l^{\Omega^{k+1}}(t) \right] dt, \end{aligned}$$

$$(12.7) \quad dy_{m,l}^{\Omega^{k+1}}(t) = N^l \frac{e_m}{N^m} \left[ x_l^{\Omega^{k+1}}(t) - y_{m,l}^{\Omega^{k+1}}(t) \right] dt, \quad 0 \leq m \leq k,$$

$$(12.8) \quad dy_{k+1,l}^{\Omega^{k+1}}(t) = 0,$$

In the limit as  $N \rightarrow \infty$ , 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  $y_{m,l}^{\Omega^{k+1}}$  with  $m > l$  vanishes in the limit as  $N \rightarrow \infty$ . 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  $y_{l,l}^{\Omega^{k+1}}$  has a non-trivial drift towards the active  $l$ -block average, written  $x_l^{\Omega^{k+1}}$ . 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  $\lim_{N \rightarrow \infty} y_{m,l}^{\Omega_N^{k+1}} = \lim_{N \rightarrow \infty} x_l^{\Omega_N^{k+1}}$ . 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 (12.6)–(12.6), we see that if we consider the quantity

$$(12.9) \quad \frac{x_l^{\Omega_N^{k+1}}(t) + \sum_{m=0}^{l-1} K_m y_{l,m}^{\Omega_N^{k+1}}(t)}{1 + \sum_{m=0}^{l-1} K_m},$$

then all infinite rates cancel out. Therefore

$$(12.10) \quad \left( \frac{x_l^{\Omega_N^{k+1}}(t) + \sum_{m=0}^{l-1} K_m y_{l,m}^{\Omega_N^{k+1}}(t)}{1 + \sum_{m=0}^{l-1} K_m}, y_{l,l}^{\Omega_N^{k+1}}(t) \right)_{t>0}$$

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 (12.6)–(12.7) and the two-level three-colour mean-field system in Section (11.1) 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 (4.5) and (4.3). For  $0 \leq l \leq k$ , let

$$(12.11) \quad (z_{l,(\theta, (y_{m,l})_{m=0}^{k+1})}(t))_{t \geq 0} = (x_l(t), (y_{m,l}(t))_{m=0}^{k+1})_{t \geq 0}$$

be the process evolving according to

$$(12.12) \quad \begin{aligned} dx_l(t) &= E_l \left[ c_l [\theta - x_l(t)] dt + \sqrt{\mathcal{F}^{(l)} g(x_l(t))} dw(t) + K_l e_l [y_{l,l}(t) - x_l(t)] dt \right], \\ y_{m,l}(t) &= x_l(t), \quad \text{for } 0 \leq m < l \\ dy_{l,l}(t) &= e_l [x_l(t) - y_{l,l}(t)] dt, \\ y_{m,l}(t) &= y_{m,l}, \quad \text{for } l < m \leq k+1, \end{aligned}$$

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

$$(12.13) \quad (z_{l,(\theta, (y_{m,l})_{m=l+1}^{k+1})}(t))_{t \geq 0} = (x_l^{\text{aux}}(t), (y_{m,l}^{\text{aux}}(t))_{m=l+1}^{k+1})_{t \geq 0}$$

be the process evolving according to

$$(12.14) \quad \begin{aligned} dx_l^{\text{aux}}(t) &= E_l \left[ c_l [\theta - x_l^{\text{aux}}(t)] dt + \sqrt{\mathcal{F}^{(l)} g(x_l^{\text{aux}}(t))} dw(t) + K_l e_l [y_{l,l}^{\text{aux}}(t) - x_l^{\text{aux}}(t)] dt \right], \\ dy_{l,l}^{\text{aux}}(t) &= e_l [x_l^{\text{aux}}(t) - y_{l,l}^{\text{aux}}(t)] dt, \\ y_{m,l}^{\text{aux}}(t) &= y_{m,l}, \quad \text{for } l < m \leq k+1, \end{aligned}$$

where  $\theta \in [0, 1]$  and  $y_{m,l}^{\text{aux}} \in [0, 1]$ , for  $l < m \leq k+1$ .

For  $0 \leq l \leq k$ , let

$$(12.15) \quad (z_{l,\theta}^{\text{eff}}(t))_{t \geq 0} = (x_l^{\text{eff}}(t), y_{l,l}^{\text{eff}}(t))_{t \geq 0}$$

be the *effective process* evolving according to

$$(12.16) \quad \begin{aligned} dx_l^{\text{eff}}(t) &= E_l \left[ c_l [\theta - x_l^{\text{eff}}(t)] dt + \sqrt{(\mathcal{F}^{(l)}g)(x_l^{\text{eff}}(t))} dw(t) + K_l e_l [y_{l,l}^{\text{eff}}(t) - x_l^{\text{eff}}(t)] dt \right], \\ dy_{l,l}^{\text{eff}}(t) &= e_l [x_l^{\text{eff}}(t) - y_{l,l}^{\text{eff}}(t)] dt. \end{aligned}$$

Comparing (12.12) with (12.16), we see that the effective process looks at the non-trivial components of the full process. The auxiliary process in (12.14) 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:

1. For  $t > 0$  and for  $0 \leq l \leq k$ , define the  *$l$ -block estimators*

$$(12.17) \quad \begin{aligned} \bar{\Theta}^{(l), \Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{i \in B_l} \frac{x_i^{\Omega_N^{k+1}}(t) + \sum_{m=0}^{l-1} K_m y_{i,0}^{\Omega_N^{k+1}}(t)}{1 + K_0}, \\ \Theta_x^{(l), \Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{i \in B_l} x_i^{\Omega_N^{k+1}}(t), \\ \Theta_{y_m}^{(l), \Omega_N^{k+1}}(t) &= \frac{1}{N^l} \sum_{i \in B_l} y_{i,m}^{\Omega_N^{k+1}}(t), \quad 0 \leq m \leq k+1, \end{aligned}$$

and put

$$(12.18) \quad \begin{aligned} \Theta^{(l), \Omega_N^{k+1}}(t) &= \left( \Theta_x^{(l), \Omega_N^{k+1}}(t), (\Theta_{y_m}^{(l), \Omega_N^{k+1}}(t))_{m=0}^{k+1} \right), \\ \Theta^{\text{aux}, (l), \Omega_N^{k+1}}(t) &= \left( \bar{\Theta}^{(l), \Omega_N^{k+1}}(t), \left( \Theta_{y_l}^{(l), \Omega_N^{k+1}}(t) \right)_{m=l}^{k+1} \right), \\ \Theta^{\text{eff}, (l), \Omega_N^{k+1}}(t) &= \left( \bar{\Theta}^{(l), \Omega_N^{k+1}}(t), \Theta_{y_l}^{(l), \Omega_N^{k+1}}(t) \right). \end{aligned}$$

We call  $(\Theta^{(l), \Omega_N^{k+1}}(t))_{t>0}$  the  *$l$ -block estimator process*,  $(\Theta^{\text{aux}, (l), \Omega_N^{k+1}}(t))_{t>0}$  the *auxiliary  $l$ -block estimator process* and  $(\Theta^{\text{eff}, (l), \Omega_N^{k+1}}(t))_{t>0}$  the *effective  $l$ -block estimator process*.

2. For  $0 \leq l \leq k$ , define the *time scales*  $N^l$  such that

$$(12.19) \quad \mathcal{L}[\bar{\Theta}^{(l), \Omega_N^{k+1}}(N^l t_l - L(N)N^{l-1}) - \bar{\Theta}^{(l), \Omega_N^{k+1}}(N^l t_l)] = \delta_0$$

for all  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \text{inf ty}$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , 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.

3. The *invariant measure* for the evolution of the  $l$ -block average in (12.12), denoted by

$$(12.20) \quad \Gamma_{\theta, y_l}^{(l)}, \quad y_l = (y_{m,l})_{m=0}^{k+1}.$$

The *invariant measures* of the auxiliary  $l$ -block process in (12.15) and the effective  $l$ -block process in (12.16), denoted by, respectively,

$$(12.21) \quad \Gamma_{\theta, y_l}^{(l), \text{aux}}, \quad y_l = (y_{m,l})_{m=l+1}^{k+1}$$

and

$$(12.22) \quad \Gamma_{\theta}^{(l), \text{eff}}.$$

4. 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

$$(12.23) \quad (\mathcal{F}^{E_l, c_l, K_l, e_l} g)(\theta) = \int_{[0,1]^2} g(x) \Gamma_{\theta}^{(l)}(dx, (dy)_{m=0}^{k+1}), \quad \theta \in [0, 1],$$

and define the iterates  $\mathcal{F}^{(n)}$ ,  $0 \leq n \leq k$ , of the renormalisation transformation as the compositions

$$(12.24) \quad \mathcal{F}^{(l)} = \mathcal{F}^{E_{n-1}, c_{n-1}, K_{n-1}, e_{n-1}} \circ \dots \circ \mathcal{F}^{E_0, c_0, K_0, e_0}, \quad 0 \leq l \leq k.$$

(Recall (4.17).)

5. To give a detailed description of the multi-scale behaviour of the SSDE in (2.20), define the *interaction chain*

$$(12.25) \quad (M_{-l}^k)_{-l=-(k+1), -k, \dots, 0}$$

as the *time-inhomogeneous* Markov chain on  $[0, 1] \times [0, 1]^{k+1}$  with initial state

$$(12.26) \quad M_{-(k+1)}^k = (\vartheta_k, \overbrace{\vartheta_k, \dots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y, k+1})$$

that evolves according to the transition kernel  $Q^{[l]}$  from time  $-(l+1)$  to time  $-l$  given by

$$(12.27) \quad Q^{[l]}(u, dv) = \Gamma_u^{(l)}(dv), \quad 0 \leq l \leq k.$$

(Recall (4.18).)

We are now ready to state the scaling limit for the evolution of the averages in (10.7).

**Proposition 12.1. [Finite-level mean-field: finite-systems scheme]** *Suppose that the initial state of the system in (12.4) is given by  $\mu(0) = \mu^{\otimes [\Omega_N^{k+1}]}$  for some  $\mu \in \mathcal{P}([0, 1] \times [0, 1]^{k+2})$ . Let  $L(N)$  be such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ , 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  $t_k, \dots, t_0 \in (0, \infty)$ ,

$$(12.28) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(l), \Omega_N^{k+1}}(\bar{t}) \right)_{l=k+1, k, \dots, 0} \right] = \mathcal{L} \left[ (M_{-l}^k)_{-l=-(k+1), -k, \dots, 0} \right],$$

where  $(M_{-l}^k)_{-l=-(k+1), -k, \dots, 0}$  is the interaction chain in (12.25) starting from

$$(12.29) \quad M_{-(k+1)}^k = (\vartheta_k, \overbrace{\vartheta_k, \dots, \vartheta_k}^{k+1 \text{ times}}, \theta_{y, k+1}).$$

- (b) For all  $0 \leq l \leq k$ ,

$$(12.30) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{(l), \Omega_N^{k+1}}(\bar{t} + t_l N^l) \right)_{t_l > 0} \right] = \mathcal{L} \left[ \left( z_{l, M_{-(l+1)}^k}(t_l) \right)_{t_l > 0} \right],$$

in the Meyer-Zheng topology,

where

$$(12.31) \quad (z_{l, M_{-(l+1)}^k}(t_l))_{t_l > 0}$$

is the processes defined (12.12) with  $\theta$ ,  $(y_m, l)_{m=l+1}^{k+1}$  given by the corresponding components in  $M_{-(l+1)}^k$  and with initial measure

$$(12.32) \quad \begin{aligned} \mathcal{L} \left[ z_{l, M_{-(l+1)}^k}(0) \right] &= \Gamma_{M_{-(l+1)}^k}^{(l)} \\ \Gamma_{M_{-(l+1)}^k}^{(l)} &= \int_{\mathfrak{S}^{k+1}} \dots \int_{\mathfrak{S}^{k+1}} \int_{\mathfrak{S}^{k+1}} \Gamma_{M_{-(k+1)}^k}^{(k)}(du_k) \Gamma_{u_k}^{(k-1)}(du_{k-1}) \dots \Gamma_{u_2}^{(l+1)}(du_{l+2}) \Gamma_{u_{l+1}}^{(l)}. \end{aligned}$$

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  $N \rightarrow \infty$  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 12.2.** In contrast to Propositions 10.2–11.1, there are no assumptions on the seed-bank behaviour in Proposition 12.1. 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 scale on which this seed-bank is effective. Hence the state of the slow seed-banks is determined by the interaction chain. ■

The proof of Proposition 12.1 will be given in Section 12.2.

## 12.2 Proof of the mean-field finite-systems scheme: finite-level

We give a sketch of the proof Proposition 12.1. The proof uses a similar scheme as the proof of Proposition 11.1. We state the scheme and indicate at each step how it can be proved.

1. Tightness of the auxiliary  $l$ -block estimator processes, for  $0 \leq l \leq k$ ,

$$(12.33) \quad \left( \left( \Theta^{\text{aux},(l),\Omega_N^{k+1}}(N^l t_l) \right)_{t_l > 0} \right)_{N \in \mathbb{N}}.$$

*Proof.* For each  $0 \leq l \leq k$  we use the tightness criterion in [JM86, Proposition 3.2.3]. □

2. Stability property of the 2-block estimators, i.e., for  $L(N)$  such that  $\lim_{N \rightarrow \infty} L(N) = \infty$  and  $\lim_{N \rightarrow \infty} L(N)/N = 0$ ,

$$(12.34) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l) - \bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l - N^{l-1} t) \right| = 0 \text{ in probability}$$

and, for all  $l \leq m \leq k + 1$ ,

$$(12.35) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \Theta_{y_m}^{(l),\Omega_N^{k+1}}(N^l t_l) - \Theta_{y_m}^{(l),\Omega_N^{k+1}}(N^l t_l - N^{l-1} t_l) \right| = 0 \text{ in probability.}$$

*Proof.* Use a similar computation as in the proof of Lemma 11.7. □

3. We analyse the behaviour of the slow seed-banks by proving the following lemma.

**Lemma 12.3. [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  $t_l > 0$ ,*

$$(12.36) \quad \lim_{N \rightarrow \infty} \left[ y_{i,m}^{\Omega_N^{k+1}}(\bar{t} + N^l t_l) - \Theta_{y_m,i}^{(m),\Omega_N^{k+1}}(\bar{t} + N^l t_l) \right] = 0 \text{ a.s.}$$

and hence

$$(12.37) \quad \lim_{N \rightarrow \infty} \left[ \Theta_{y_m,i}^{(l),\Omega_N^{k+1}}(\bar{t} + N^l t_l) - \Theta_{y_m,i}^{(m),\Omega_N^{k+1}}(\bar{t} + N^l t_l) \right] = 0 \text{ a.s.}$$

*Proof.* We can proceed as in the proof of Lemma 11.23, 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

$$(12.38) \quad H_t^{m,\Omega_N^{k+1}} = \{RW^{\Omega_N^{k+1}} \text{ has migrated over distance } m \text{ at least once up to time } t\}.$$

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  $\Omega_N^{k+1}$ . Since tightness of components implies tightness of the process, step 1 implies that for  $0 \leq l \leq k$  the full  $l$ -block processes

$$(12.39) \quad \left( \left( \Theta_i^{\text{aux},(l),\Omega_N^{k+1}}(\bar{t} + N^l t_l) \right)_{t_l > 0, i \in [N^{k+1-l}]} \right)_{N \in \mathbb{N}}$$

are tight. From the tightness in steps 1 we can construct a subsequence  $(N_n)_{n \in \mathbb{N}}$  along which, for all  $0 \leq l \leq k$ ,

$$(12.40) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t} + N_n^l t_l) \right)_{t_l > 0, i \in [N_n^{k+1-l}]} \right]$$

exists. Note that  $\bar{t}$  depends on the subsequence. For example, along the subsequence  $(N_{\bar{n}})_{\bar{n} \in \mathbb{N}}$ ,

$$(12.41) \quad \bar{t} = L(N)N^k + \sum_{n=0}^k t_n N_{\bar{n}}^n.$$

We define the measure

$$(12.42) \quad \nu_{\Theta}^{(0)} = \prod_{i \in \mathbb{N}_0} \Gamma_{\Theta_i}^{(0)}(\bar{t}),$$

where

$$(12.43) \quad \Theta_i \in \mathfrak{s}^{k+1}.$$

In this step we show that along the same subsequence the single components converge to the infinite system. We show that if

$$(12.44) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(\Theta_i^{\text{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i \in [N_n^k]}] = P^{(1)},$$

then

$$(12.45) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( Z^{\Omega_{N_n}^{k+1}}(\bar{t} + t_0) \right)_{t_0 \geq 0} \right] = \mathcal{L} \left[ \left( Z^{\nu^{(0)}(\bar{t})}(t_0) \right)_{t_0 \geq 0} \right],$$

where

$$(12.46) \quad \nu^{(0)}(\bar{t}) = \int \nu_u^{(0)} P^{(1)}(du).$$

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 (11.18), where  $\theta$  is now a random variable that inherits its law from

$$(12.47) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(\Theta_i^{\text{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i \in [N_n^k]}]$$

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

$$(12.48) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(\Theta_i^{\text{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i \in [N_n^k]}].$$

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 (12.44) exists.

*Proof.* Proceed as in the proof of Proposition 11.8. Note that the assumptions on the seed-banks in Proposition 11.8 follow from the choice of the subsequence and Lemma 12.3.  $\square$

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

$$(12.49) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(\Theta^{\text{aux},(l+1),\Omega_{N_n}^{k+1}}(\bar{t}))_{i \in [N_n^k]}] = P^{(l+1)},$$

then

$$(12.50) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{aux},(l),\Omega_{N_n}^{k+1}}(\bar{t} + N_n^l t_l) \right)_{t_l > 0, i \in [N_n^{k+1-l}]} \right] = \mathcal{L} \left[ (z_{l,\Theta^{(l+1)}}^{\text{aux}}(t))_{t \geq 0} \right],$$

where  $\Theta^{(l+1)} = (\Theta_x^{(l+1)}, (\Theta_{y_m,l}^{(l+1)})_{m=l+1}^{k+1}) \in [0, 1] \times [0, 1]^{k+2-(l+1)}$ ,

$$(12.51) \quad \begin{aligned} \mathcal{L} \left[ z_{l,\Theta^{(l+1)}}^{\text{aux}}(0) \right] &= \Gamma_{\Theta^{(l+1)}}^{(l),\text{aux}}, \\ \Gamma_{\Theta^{(l+1)}}^{(l),\text{aux}} &= \int_{[0,1] \times [0,1]^{k+2-(l+1)}} \Gamma_u^{(l),\text{aux}} P^{(l+1)}(du) \end{aligned}$$

and  $(z_{l,\Theta^{(l+1)}}^{\text{aux}}(t))_{t \geq 0}$  is the process evolving according to (12.14) with  $\theta, (y_{m,l})_{m=l}^{k+1}$  replaced by the random variables  $\Theta_x^{(l+1)}, (\Theta_{y_m,l}^{(l+1)})_{m=l}^{k+1}$ . 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 (12.49) exist.

*Proof.* The proof goes by induction. Using the convergence of the single components, we can proceed as in the proof of Proposition 11.13 to prove the convergence of the 1-blocks averages

$$(12.52) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(1),\Omega_{N_n}^{k+1}}(\bar{t} + N_n t_1) \right)_{t_1 > 0, i \in [N_n^{k+1-l}]} \right].$$

Then, assuming that we have the convergence for all  $0 \leq l \leq L$ , we get

$$(12.53) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(l),\Omega_{N_n}^{k+1}}(\bar{t} + N_n^l t_l) \right)_{t_l > 0, i \in [N_n^{k+1-l}]} \right],$$

and we prove the convergence of

$$(12.54) \quad \lim_{n \rightarrow \infty} \mathcal{L} \left[ \left( \Theta_i^{\text{aux},(L+1),\Omega_{N_n}^{k+1}}(\bar{t} + N_n^{(L+1)} t_{(L+1)}) \right)_{t_{(L+1)} > 0, i \in [N_n^{k+1-(L+1)}]} \right].$$

This is done using a similar proof strategy as in the proof of Proposition 11.13. In particular, we need to derive the  $l$ -level equivalent of Lemma 11.16. Since this lemma is also key to proving convergence in the Meyer-Zheng topology, we state it explicitly below.  $\square$

**Lemma 12.4.** [ $l$ -block averages] *Define*

$$(12.55) \quad \begin{aligned} &\Delta_{\Sigma}^{(l),\Omega_{N_n}^{k+1}}(N^{l-1} t_{l-1}) \\ &= \frac{\Theta_x^{(l),\Omega_{N_n}^{k+1}}(N^{l-1} t_{l-1}) + \sum_{m=0}^{l-2} K_m \Theta_{y_m}^{(l),\Omega_{N_n}^{k+1}}(N^{l-1} t_{l-1})}{1 + \sum_{m=0}^{l-2} K_m} - \Theta_{y_{l-1}}^{(l),\Omega_{N_n}^{k+1}}(N^{l-1} t_{l-1}) \end{aligned}$$

and

$$(12.56) \quad R_l = \frac{1 + \sum_{m=0}^{l-1} K_m}{1 + \sum_{m=0}^{l-2} K_m}.$$

For  $t \geq 0$  set  $\bar{\Theta}^{(0)}(t) = \Theta_x^{(0)}(t) = x_0(t)$ . Then, for  $1 \leq l \leq k$ ,

$$\begin{aligned}
& \mathbb{E} \left[ \left| \Delta_{\Sigma}^{(l), \Omega_N^{k+1}}(N^{l-1}t_{l-1}) \right| \right] \\
& \leq \sqrt{\mathbb{E} \left[ \left( \Delta_{\Sigma}^{(l), \Omega_N^{k+1}}(0) \right)^2 \right]} e^{-\epsilon_l R_l t_{l-1}} \\
(12.57) \quad & + \sqrt{\int_0^{t_1} ds 2\epsilon_l R_l e^{-2\epsilon_l R_l(t_1-s)} \mathbb{E} \left[ \left| \bar{\Theta}^{(l-1), \Omega_N^{k+1}}(N^{l-1}s) - \Theta_x^{(l-1), \Omega_N^{k+1}}(N^{l-1}s) \right|^2 \right]} \\
& + \sqrt{\frac{1}{N} \frac{1}{2\epsilon_l(1 + \sum_{m=0}^{l-1} K_m)} \left[ \sum_{n=l+1}^k \frac{c_{n-1}}{N^{n-(l+1)}} + \sum_{m=l}^k \frac{K_m \epsilon_m}{N^{m-l}} + E_{l-1} \|g\| \right]}.
\end{aligned}$$

*Proof.* Proceed like in the proof of Lemma 11.16, using the SSDE in (12.4) instead of the SSDE in (11.6).  $\square$

We obtain the following useful corollary from Lemma 12.4.

**Corollary 12.5.** For all  $1 \leq l \leq k$ ,  $s > 0$  and  $\tilde{l} \geq l$ ,

$$(12.58) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(l-1), \Omega_N^{k+1}}(N^{\tilde{l}-1}s) - \Theta_x^{(l-1), \Omega_N^{k+1}}(N^{\tilde{l}-1}s) \right|^2 \right] = 0.$$

*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

$$\begin{aligned}
& \mathbb{E} \left[ \left| \bar{\Theta}^{(L), \Omega_N^{k+1}}(N^L s) - \Theta_x^{(L), \Omega_N^{k+1}}(N^L s) \right|^2 \right] \\
(12.59) \quad & \leq \mathbb{E} \left[ \left| \bar{\Theta}^{(L), \Omega_N^{k+1}}(N^L s) - \frac{1}{N} \sum_{i=0}^{N-1} \bar{\Theta}_i^{(L-1), \Omega_N^{k+1}}(N^L s) \right|^2 \right] \\
& + \frac{1}{N} \sum_{i=0}^{N-1} \mathbb{E} \left[ \left| \bar{\Theta}_i^{(L-1), \Omega_N^{k+1}}(N^L s) - \Theta_{x,i}^{(L-1), \Omega_N^{k+1}}(N^L s) \right|^2 \right].
\end{aligned}$$

Note that the second term in the right-hand side of (12.59) tends to 0 as  $N \rightarrow \infty$  by the induction hypothesis. For the first term in the right-hand side of (12.59), note that

$$\begin{aligned}
& \mathbb{E} \left[ \left| \bar{\Theta}^{(L), \Omega_N^{k+1}}(N^L s) - \frac{1}{N} \sum_{i=0}^{N-1} \bar{\Theta}_i^{(L-1), \Omega_N^{k+1}}(N^L s) \right|^2 \right] \\
(12.60) \quad & = \mathbb{E} \left[ \left| \frac{\Theta_x^{(L), \Omega_N^{k+1}}(N^L s) + \sum_{m=0}^{L-1} K_m \Theta_{y_m}^{(L), \Omega_N^{k+1}}(N^L s)}{1 + \sum_{m=0}^{L-1} K_m} \right. \right. \\
& \quad \left. \left. - \frac{\Theta_x^{(L), \Omega_N^{k+1}}(N^L s) + \sum_{m=0}^{L-2} K_m \Theta_{y_m}^{(L), \Omega_N^{k+1}}(N^L s)}{1 + \sum_{m=0}^{L-2} K_m} \right|^2 \right] \\
& = \frac{K_{L-1}}{1 + \sum_{m=0}^{L-1} K_m} \mathbb{E} \left[ \left| \Theta_{y_{L-1}}^{(L), \Omega_N^{k+1}}(N^L s) - \bar{\Theta}^{(L), \Omega_N^{k+1}}(N^L s) \right|^2 \right].
\end{aligned}$$

Invoking Lemma 12.4 and using the induction hypothesis, we see that for  $s > 0$  and  $\tilde{l} \geq L$  indeed

$$(12.61) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(L), \Omega_N^{k+1}}(N^{\tilde{l}}s) - \Theta_x^{(L), \Omega_N^{k+1}}(N^{\tilde{l}}s) \right|^2 \right] = 0.$$

$\square$



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  $N \rightarrow \infty$ .

Define, for  $0 \leq l \leq k$ ,

$$(12.62) \quad \mathfrak{s}_l^{k+1} = [0, 1] \times [0, 1]^{k+2-l}.$$

We obtain, for  $0 \leq l \leq k-1$ ,

$$(12.63) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathcal{L}[(\Theta^{\text{aux},(l+1),\Omega_N^{k+1}}(\bar{t}))] &= \Gamma_{\Theta^{(l+2)}}^{(l+1),\text{aux}}, \\ \Gamma_{\Theta^{(l+2)}}^{(l+1),\text{aux}} &= \int_{\mathfrak{s}_{l+2}^{k+1}} \cdots \int_{\mathfrak{s}_k^{k+1}} \Gamma_{\vartheta_k}^{(k),\text{aux}}(du_k) \cdots \Gamma_{u_{l+3}}^{(l+2),\text{aux}}(du_{l+2}) \Gamma_{u_{l+2}}^{(l+1),\text{aux}}. \end{aligned}$$

Therefore by step 5

$$(12.64) \quad \lim_{N \rightarrow \infty} \mathcal{L} \left[ \left( \Theta^{\text{aux},(l),\Omega_N^{k+1}}(\bar{t} + N_n^l t_l) \right)_{t_l > 0, i \in [N^{k+1-l}]} \right] = \mathcal{L} \left[ (z_{l,\Theta^{(l+1)}}^{\text{aux}}(t))_{t \geq 0} \right],$$

where  $\Theta^{(l+1)} = (\Theta_x^{(l+1)}, (\Theta_{y_m,l}^{(l+1)})_{m=l}^{k+1}) \in \mathfrak{s}_l^{(k+1)}$  are random variables with law

$$(12.65) \quad \mathcal{L} \left[ \Theta^{(l+1)} \right] = \Gamma_{\Theta^{(l+2)}}^{(l+1),\text{aux}}.$$

The initial state of the limiting process in (12.64) is given by

$$(12.66) \quad \begin{aligned} \mathcal{L} \left[ z_{l,\Theta^{(l)}}^{\text{aux}}(0) \right] &= \Gamma_{\Theta^{(l+1)}}^{(l),\text{aux}}, \\ \Gamma_{\Theta^{(l+1)}}^{(l),\text{aux}} &= \int_{\mathfrak{s}_{l+2}^{k+1}} \cdots \int_{\mathfrak{s}_k^{k+1}} \Gamma_{\vartheta+k}^{(k),\text{aux}}(du_k) \cdots \Gamma_{u_{l+2}}^{(l+2),\text{aux}}(du_{l+1}) \Gamma_{u_{l+1}}^{(l+1),\text{aux}} \end{aligned}$$

and  $(z_{l,\Theta^{(l+1)}}^{\text{aux}}(t))_{t \geq 0}$  is the process evolving according to (12.14) with  $\theta, (y_{m,l})_{m=l+1}^{k+1}$  replaced by the random variables  $\Theta_x^{(l+1)}, (\Theta_{y_m}^{(l+1)})_{m=l+1}^{k+1}$ . Recall that, by Lemma 12.3, we have, for  $l+1 \leq m \leq k+1$ ,

$$(12.67) \quad \Theta_{y_m}^{(l+1)} = \Theta_{y_m}^{(m)}.$$

7. Use the Meyer-Zheng topology to obtain Proposition 12.1(b).

*Proof.* Note that Lemma 12.4 and Corollary 12.5 together imply

$$(12.68) \quad \begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l) - \Theta_x^{(l),\Omega_N^{k+1}}(N^l t_l) \right| \right] &= 0, \\ \lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \bar{\Theta}^{(l),\Omega_N^{k+1}}(N^l t_l) - \Theta_{y_m}^{(l),\Omega_N^{k+1}}(N^l t_l) \right| \right] &= 0, \text{ for } 0 \leq m \leq l-1. \end{aligned}$$

Combining the result obtained in step 6 with the proof strategy followed in Section 11.3.10, we get the claim.  $\square$

8. Finally, we prove Proposition 12.1(a).

*Proof.* Step 6 and step 7 yield the laws of the components  $\mathcal{L}[M_l^k]$  of the interaction chain  $(M_{-l}^k)_{l=-(k+1)}^0$ . Note that the state space  $([0, 1] \times [0, 1]^{k+2})^{k+2}$  is compact, and therefore the sequence of random variables

$$(12.69) \quad \left( \left( \Theta^{(l),\Omega_N^{k+1}}(\bar{t}) \right)_{l=k+1,k,\dots,0} \right)_{N \in \mathbb{N}}$$

is tight. For any

$$(12.70) \quad \begin{aligned} f &: ([0, 1] \times [0, 1]^{k+2})^{n+2} \rightarrow \mathbb{R}, \\ f(x) &= \prod_{i=1}^n f_i(x_i), \\ f_i &\in \mathcal{C}_b([0, 1] \rightarrow \mathbb{R}), \end{aligned}$$

we can use conditioning on the previous block average to obtain

$$(12.71) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ f \left( \left( \Theta^{(l), \Omega_N^{k+1}}(\bar{t}) \right)_{l=k+1, k, \dots, 0} \right) \right] = \mathbb{E} \left[ f \left( (M_{-l}^k)_{-l=-(k+1)}^0 \right) \right].$$

Using that the set of functions of the form (12.70) is separating, we obtain the claim.  $\square$

### 12.3 Proof: of the hierarchical multi-scale limit theorems.

In this section we prove Theorems 4.2 and 4.4. We start by proving Theorem 4.4. Theorem 4.2 will follow from Theorem 4.4 by projection onto the effective components.

#### Proof of Theorem 4.2

*Proof.* Recall the estimators in (4.11). Like for the finite-level hierarchical mean-field system, we can define the auxiliary estimator process by

$$(12.72) \quad \Theta^{(l), \text{aux}, \Omega_N}(t) = (\bar{\Theta}^{(l), \Omega_N}(t), (\Theta_{y_m}^{(l), \Omega_N}(t))_{m=l}^\infty).$$

For  $l, k \in \mathbb{N}$  the processes  $(\Theta^{(l), \text{aux}, \Omega_N}(\bar{t} + N^k t))_{t>0}$  evolve according to (recall, (4.54))

$$(12.73) \quad \begin{aligned} d\bar{\Theta}^{(l), \Omega_N}(N^k t) &= E_l \sum_{n=l+1}^\infty \frac{c_{n-1}}{N^{n-1-k}} \left[ \Theta_x^{(n), \Omega_N}(N^k t) - \Theta_x^{(l), \Omega_N}(N^k t) \right] dt \\ &\quad + E_l \sqrt{\frac{N^k}{N^{2l}} \sum_{\xi \in B_l} g(x_\xi(N^k t))} dw(t) \\ &\quad + E_l \sum_{m=l}^\infty \frac{K_m e_m}{N^{m-k}} \left[ \Theta_{y_m}^{(l), \Omega_N}(N^k t) - \Theta_x^{(l), \Omega_N}(N^k t) \right] dt, \end{aligned}$$

$$d\Theta_{y_m}^{(l), \Omega_N}(N^k t) = \frac{e_m}{N^{m-k}} \left[ \Theta_x^{(l), \Omega_N}(N^k t) - \Theta_{y_m}^{(l), \Omega_N}(N^k t) \right] dt, \quad l \leq m \leq \infty.$$

Therefore, for  $l > k$  and all  $\epsilon > 0$ ,

$$(12.74) \quad \begin{aligned} &\mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(l), \Omega_N}(\bar{t}) - \bar{\Theta}^{(l), \Omega_N}(\bar{t} + N^k t) \right| > \epsilon \right] \\ &= \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} E_l \left| \int_{\bar{t}}^{\bar{t} + N^k t} dr \sum_{n=l+1}^\infty \frac{c_{n-1}}{N^{n-1}} \left[ \Theta_x^{(n), \Omega_N}(r) - \Theta_x^{(l), \Omega_N}(r) \right] \right. \right. \\ &\quad \left. \left. + \int_{\bar{t}}^{\bar{t} + N^k t} dr \sum_{m=l}^\infty \frac{K_m e_m}{N^m} \left[ \Theta_{y_m}^{(l), \Omega_N}(r) - \Theta_x^{(l), \Omega_N}(r) \right] \right. \right. \\ &\quad \left. \left. + \int_{\bar{t}}^{\bar{t} + N^k t} dw_i(r) \sqrt{\frac{1}{N^{2l}} \sum_{\xi \in B_l} g(x_\xi(r))} \right| > \epsilon \right] \\ &\leq \mathbb{P} \left[ \sup_{0 \leq t \leq L(N)} E_l \left| \int_{\bar{t}}^{\bar{t} + N^k t} dw_i(r) \sqrt{\frac{1}{N^{2l}} \sum_{\xi \in B_l} g(x_\xi(r))} \right| \right. \\ &\quad \left. > \epsilon - t \left[ \sum_{n=l+1}^\infty \frac{c_{n-1}}{N^{n-1-k}} + \sum_{m=l}^\infty \frac{K_m e_m}{N^{m-k}} \right] \right]. \end{aligned}$$

Note that, since  $l > k$ ,

$$(12.75) \quad \lim_{N \rightarrow \infty} t \left[ \sum_{n=l+1}^{\infty} \frac{c_{n-1}}{N^{n-1-k}} + \sum_{m=l}^{\infty} \frac{K_m e_m}{N^{m-k}} \right] = 0.$$

Hence, like in the proof of Lemma 11.7, we can use an optional stopping argument to obtain

$$(12.76) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}^{(l), \Omega_N}(\bar{t}) - \bar{\Theta}^{(l), \Omega_N}(\bar{t} + N^k t) \right| = 0 \quad \text{in probability.}$$

Using a similar computation as in (12.74), we can show

$$(12.77) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq L(N)} \left| \bar{\Theta}_{y_m}^{(l), \Omega_N}(\bar{t}) - \bar{\Theta}_{y_m}^{(l), \Omega_N}(\bar{t} + N^k t) \right| = 0 \quad \text{in probability.}$$

Hence we obtain that, on time scale  $N^k$  as  $N \rightarrow \infty$ , the process  $(\Theta^{(l), \text{aux}, \Omega_N}(\bar{t} + N^k t))_{t > 0}$  does not evolve and therefore is still in its initial state  $(\Theta^{(l), \text{aux}, \Omega_N}(\bar{t}))$ .

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 12.1. Therefore we can proceed as in the proof of Proposition 12.1 to prove the second and third line in (4.29) in Theorem 4.2. The  $l$ -block estimator process  $(\Theta^{(l), \text{aux}, \Omega_N}(\bar{t} + N^l t))_{t > 0}$  evolves according to (12.73) 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 (4.66) are of order  $\mathcal{O}(1/N)$ . Therefore these terms vanish as  $N \rightarrow \infty$ , and we can just proceed as in the scheme of Section 12.2, to obtain the second and third line in (4.29) in Theorem 4.2. Using these results, we obtain that, for  $l > k$ ,

$$(12.78) \quad \Theta^{(l), \text{aux}, \Omega_N}(\bar{t}) = \delta_{M_{-(k+1)}^k}.$$

□

## 13 Proofs: $N \rightarrow \infty$ , orbit of renormalisation transformation

In this section we analyse the orbit of the renormalisation transformation and show that it has the Fisher-Wright diffusion as a global attractor. In Section 13.1 we write down moment relations for the equilibrium defined in (4.8) for single colonies (Proposition 13.1) and for block averages (Proposition 13.2). In Section 13.2 we derive the iterates of these moment relations for single colonies (Proposition 13.3) and for blocks (Proposition 13.4). In Section 13.3 we prove clustering (Propositions 13.5–13.6). In Section 13.4 we prove Theorems 5.1 and 5.3, and work out the dichotomy of a finite seed-bank ( $\rho < \infty$ ) versus infinite seed-bank ( $\rho = \infty$ ).

### 13.1 Moment relations

We use Itô-calculus to compute the mixed moments. Recall  $\theta_x, (\theta_{y_m})$  as defined in (2.21),  $\vartheta_k$  as defined in (4.3) and  $\bar{\vartheta}^{(k)}$  (5.2). Also recall  $E_k$  as defined in (4.5). Abbreviate

$$(13.1) \quad A_0^n = \frac{1}{2} \sum_{k=0}^n \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}, \quad n \in \mathbb{N},$$

and

$$(13.2) \quad B_0 = \frac{1}{2} \frac{E_0^2}{(E_0 c_0 + e_0) + E_0 K_0 e_0}.$$

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 13.1. [Moment relations: single colonies]**

Let  $\vartheta_0$  be as defined in (4.3), and let  $\Gamma_{(\vartheta_0, y_t)}^{(0)} = \Gamma_{(\vartheta_0, y_t)}^{g, c_0, E_0, K_0, e_0}$  be the equilibrium of (4.8) measure defined in (4.14) 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:

$$(13.3) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0 \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0) = \vartheta_0,$$

$$(13.4) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{0,0} \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0) = \vartheta_0,$$

$$(13.5) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0 y_{0,0} \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0) = \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{0,0}^2 \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0),$$

$$(13.6) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0^2 \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0) = \vartheta_0^2 + A_0^0(\mathcal{F}g)(\vartheta_0),$$

$$(13.7) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{0,0}^2 \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0) = \vartheta_0^2 + (A_0^0 - B_0)(\mathcal{F}g)(\vartheta_0),$$

$$(13.8) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} g(x_0) \Gamma_{(\vartheta_0, y_t)}^{g, E_0, c_0, K_0, e_0}(\mathrm{d}z_0) = (\mathcal{F}g)(\vartheta_0).$$

*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 (4.16).

1. We begin with the first moments of  $x$  and  $y_0$ . For  $k = 0$ , (4.8) becomes

$$(13.9) \quad \mathrm{d}x(t) = E_0 \left[ c_0 [\vartheta_0 - x(t)] \mathrm{d}t + \sqrt{g(x(t))} \mathrm{d}w_0(t) + K_0 e_0 [y_0(t) - x(t)] \mathrm{d}t \right],$$

$$(13.10) \quad \begin{aligned} \mathrm{d}y_0(t) &= e_0 [x(t) - y_0(t)] \mathrm{d}t, \\ \mathrm{d}y_m(t) &= 0. \end{aligned}$$

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.  $\Gamma_{\vartheta}^{g, c_0, E_0, K_0, e_0}$ . Integrating (13.9) and taking the expectation, we get

$$(13.11) \quad \begin{aligned} \mathbb{E}[x(t) - x(0)] &= E_0 \left[ \mathbb{E} \left[ \int_0^t \mathrm{d}s c_0 [\vartheta_0 - x(s)] + K_0 e_0 \int_0^t \mathrm{d}s [y_0(s) - x(s)] \right] \right], \\ &= E_0 \left\{ \int_0^t \mathrm{d}s \mathbb{E} \left[ c_0 [\vartheta_0 - x(s)] + K_0 e_0 [y_0(s) - x(s)] \right] \right\}, \end{aligned}$$

where in the second equation we use Fubini. Turning back to differential notation, we see from (13.11) that

$$(13.12) \quad \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}[x(t)] = 0 = E_0 \left\{ \mathbb{E} \left[ c_0 [\vartheta_0 - x(t)] + K_0 e_0 [y_0(t) - x(t)] \right] \right\},$$

and it follows that

$$(13.13) \quad \mathbb{E} \left[ c_0 [\vartheta_0 - x] + K_0 e_0 [y_0 - x] \right] = 0.$$

In the same way it follows from (13.10) that

$$(13.14) \quad \mathbb{E} \left[ e_0 [x - y_0] \right] = 0.$$

Therefore we obtain from (13.13)–(13.14) that

$$(13.15) \quad \mathbb{E}[x] = \mathbb{E}[y_0] = \vartheta_0.$$

2. We next compute the second moments. By Itô's formula,

$$(13.16) \quad \begin{aligned} \mathrm{d}(x(t))^2 &= 2x(t) \mathrm{d}x(t) + (\mathrm{d}x(t))^2 \\ &= 2c_0 x(t) E_0 [\vartheta_0 - x(t)] \mathrm{d}t + 2x(t) E_0 \sqrt{g(x(t))} \mathrm{d}w_0(t) \\ &\quad + E_0 [2K_0 e_0 x(t) y_0(t) - 2K_0 e_0 x(t)^2] \mathrm{d}t + E_0^2 g(x(t)) \mathrm{d}t. \end{aligned}$$

Taking expectations and using that we are in equilibrium, we get

$$(13.17) \quad 0 = 2c_0\vartheta_0^2 - 2c_0\mathbb{E}[x^2] + 2K_0e_0\mathbb{E}[xy_0] - 2K_0e_0\mathbb{E}[x^2] + E_0\mathbb{E}[g(x)].$$

Using  $\mathbb{E}[g(x)] = (\mathcal{F}g)(\vartheta_0)$ , we find

$$(13.18) \quad \mathbb{E}[x^2] = \frac{c_0}{(c_0 + K_0e_0)}\vartheta_0^2 + \frac{K_0e_0}{(c_0 + K_0e_0)}\mathbb{E}[xy_0] + \frac{E_0}{2(c_0 + K_0e_0)}(\mathcal{F}g)(\vartheta_0).$$

In the same way we find

$$(13.19) \quad \mathbb{E}[y_0^2] = \mathbb{E}[xy_0],$$

and for the mixed second moment

$$(13.20) \quad \mathbb{E}[xy_0] = \frac{E_0c_0}{(E_0c_0 + e_0)}\vartheta_0^2 + \frac{e_0}{(E_0c_0 + e_0)}\mathbb{E}[x^2].$$

Substituting (13.20) into (13.18), we find  $\mathbb{E}[x^2]$  and hence also  $\mathbb{E}[y_0^2]$  and  $\mathbb{E}[xy_0]$ . This finishes the proof of Proposition 13.1.  $\square$

Similar moment relations can be derived for the equilibrium measures of the block averages. Define

$$(13.21) \quad A_m^n = \frac{1}{2} \sum_{k=m}^n \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}, \quad m \in \mathbb{N}_0, n \in \mathbb{N},$$

and

$$(13.22) \quad B_m = \frac{1}{2} \frac{E_m^2}{(E_m c_m + e_m) + E_m K_m e_m}, \quad m \in \mathbb{N}_0.$$

Recall the definition of  $\mathcal{F}^{(n)}$  in (4.17).

**Proposition 13.2. [Moment relations: blocks]**

Let  $\vartheta_m$  be as defined in (4.3), and let  $\Gamma_{(\vartheta_m, y_m)}^{(m)} = \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m}$  be the equilibrium measure of (4.8) 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:

$$(13.23) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_m \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m) = \vartheta_m,$$

$$(13.24) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{m,m} \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m) = \vartheta_m,$$

$$(13.25) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_m y_{m,m} \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m) \\ = \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{m,m}^2 \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m),$$

$$(13.26) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_m^2 \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m) = \vartheta_m^2 + A_m^m (\mathcal{F}^{(m+1)}g)(\vartheta_m),$$

$$(13.27) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{m,m}^2 \Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m) = \vartheta_m^2 + (A_m^m - B_m) (\mathcal{F}^{(m+1)}g)(\vartheta_m),$$

$$(13.28) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} (\mathcal{F}^{(m)}g)(x_m) d\Gamma_{(\vartheta_m, y_m)}^{\mathcal{F}^{(m)}g, c_m, E_m, K_m, e_m} (dz_m) = (\mathcal{F}^{(m+1)}g)(\vartheta_m).$$

*Proof.* The proof follows the same line of argument as the proof of Proposition 13.1.  $\square$

### 13.2 Iterate moment relations

Recall the kernels defined in (4.20), the iterates of the kernels defined in (5.1) and  $\bar{\vartheta}^{(n)}$ . Recall that  $Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0)$  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 13.3.** [Iterated moment relations: single components] For  $n \in \mathbb{N}_0$ ,

$$(13.29) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \vartheta_n,$$

$$(13.30) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{0,0} Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \vartheta_n,$$

$$(13.31) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0^2 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \vartheta_n^2 + A_0^n (\mathcal{F}^{(n+1)}g)(\vartheta_n),$$

$$(13.32) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{0,0}^2 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \vartheta_n^2 + (A_0^n - B_0)(\mathcal{F}^{(n+1)}g)(\vartheta_n),$$

$$(13.33) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0 y_{0,0} Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{0,0}^2 Q^{(n)}(\vartheta_n, dz_0),$$

$$(13.34) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} g(x) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = (\mathcal{F}^{(n+1)}g)(\vartheta_n).$$

*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 13.1. Assume the result holds true for  $n = n$ , for  $n = n + 1$  we write

$$(13.35) \quad \begin{aligned} & \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0^2 Q^{(n+1)}(\bar{\vartheta}^{(n+1)}, dz_0) \\ &= \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0^2 (Q^{[n+1]} \circ Q^{(n)})(\bar{\vartheta}^{(n+1)}, dz_0) \\ &= \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0^2 Q^{[n+1]}(\bar{\vartheta}^{(n+1)}, dz_{n+1}) Q^{(n)}(z_{n+1}, dz_0) \\ &= \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} \left[ \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0^2 Q^{(n)}(z_{n+1}, dz_0) \right] Q^{[n+1]}(\bar{\vartheta}^{(n+1)}, dz_{n+1}) \\ &= \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} \left[ x_{n+1}^2 + A_0^n (\mathcal{F}^{(n+1)}g)(x_{n+1}) \right] \Gamma_{\bar{\vartheta}^{(n+1)}}(dz_{n+1}) \\ &= \vartheta_{n+1}^2 + A_0^n (\mathcal{F}^{(n+2)}g)(\vartheta_{n+1}) + A_{n+1}^{n+1} (\mathcal{F}^{(n+2)}g)(\vartheta_{n+1}) \\ &= \vartheta_{n+1}^2 + A_0^{n+1} (\mathcal{F}^{(n+2)}g)(\vartheta_{n+1}). \end{aligned}$$

The first and second equality use the definition in (5.1), the third equality uses Fubini, the fourth equality is the induction step, the fifth equality uses Proposition 13.2, in particular, (13.26) and (13.28).  $\square$

Similar iterate moment relations hold for blocks. Define, for  $m, n \in \mathbb{N}_0$  with  $n \geq m$ ,

$$(13.36) \quad Q_m^{(n)} = Q^{[n]} \circ \dots \circ Q^{[m]}.$$

**Proposition 13.4.** [Iterated moment relations: blocks of components] For  $n, m \in \mathbb{N}_0$  with

$n \geq m$ ,

$$(13.37) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_m Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m) = \vartheta_n,$$

$$(13.38) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{m,m} Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m) = \vartheta_n,$$

$$(13.39) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_m^2 Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m) = \vartheta_n^2 + A_m^n (\mathcal{F}^{(n+1)}g)(\vartheta_n),$$

$$(13.40) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{m,m}^2 Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m) = \vartheta_n^2 + (A_m^n - B_m)(\mathcal{F}^{(n+1)}g)(\vartheta_n),$$

$$(13.41) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_m y_{m,m} Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m) = \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} y_{m,m}^2 Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m),$$

$$(13.42) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} (\mathcal{F}^{(m)}g)(x_m) Q_m^{(n)}(\bar{\vartheta}^{(n)}, dz_m) = (\mathcal{F}^{(n+1)}g)(\vartheta_n).$$

*Proof.* Follow a similar induction argument as in the proof of Proposition 13.3.  $\square$

### 13.3 Clustering

To prove Theorem 5.1, we proceed as in [BCGdH95]. The following clustering property holds for the kernels associated with single colonies.

**Proposition 13.5.** [Clustering: single colonies] *Assume*

$$(13.43) \quad \lim_{n \rightarrow \infty} \vartheta_n = \theta.$$

Then

$$(13.44) \quad \begin{aligned} \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{(x_0, y_{(0,0)}) = (0, 0)\}) &= 1 - \theta, \\ \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{(x_0, y_{(0,0)}) = (1, 1)\}) &= \theta, \end{aligned}$$

if and only if

$$(13.45) \quad \lim_{n \rightarrow \infty} A_0^n = \infty.$$

Consequently,

$$(13.46) \quad \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{(x_0, y_{0,0})\} \notin \{(0, 0), (1, 1)\}) = 0.$$

*Proof.* The proof exploits the iterated moment relations. First assume (13.45)

1. By Proposition 13.3

$$(13.47) \quad \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} x_0(1-x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \vartheta_n(1-\vartheta_n) - A_0^n (\mathcal{F}^{(n+1)}g)(\vartheta_n).$$

Because  $x_0(1-x_0) \geq 0$  for  $x \in [0, 1]$ , we have

$$(13.48) \quad \vartheta_n(1-\vartheta_n) \geq A_0^n (\mathcal{F}^{(n+1)}g)(\vartheta_n) \quad \forall n \in \mathbb{N}.$$

Since  $\lim_{n \rightarrow \infty} A_0^n = \infty$ , it follows that  $\lim_{n \rightarrow \infty} (\mathcal{F}^{(n+1)}g)(\vartheta_n) = 0$ . On the other hand,

$$(13.49) \quad (\mathcal{F}^{(n+1)}g)(\vartheta_n) = \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} g(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0)$$

and, because  $g(x) > 0$  for  $x \in (0, 1)$ ,  $Q^{(n)}(\vartheta_n, dz_0)$  puts all its mass on  $x_0 = 0$  and  $x_0 = 1$  in the limit as  $n \rightarrow \infty$ . Let

$$(13.50) \quad Q^{(n)}(\bar{\vartheta}^{(n)}, \{x_0 = 0 \text{ or } x_0 = 1\}) = \int_{[0,1] \times [0,1]^{\mathbb{N}_0}} \mathbf{1}_{\{x_0=0 \text{ or } x_0=1\}}(z_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0),$$

then

$$(13.51) \quad \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{x_0 = 0 \text{ or } x_1 = 1\}) = 1.$$

The first moment of  $x_0$  converges to (recall (4.4))

$$(13.52) \quad \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x_0 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \lim_{n \rightarrow \infty} \vartheta_n = \theta.$$

Hence

$$(13.53) \quad \begin{aligned} \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{x_0 = 0\}) &= 1 - \theta, \\ \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{x_0 = 1\}) &= \theta, \end{aligned}$$

and

$$(13.54) \quad \begin{aligned} \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x_0^2 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\ = \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x_0^2 \left( \mathbf{1}_{\{1\}}(x_0) + \mathbf{1}_{\{0\}}(x_0) + \mathbf{1}_{\{(0,1)\}}(x_0) \right) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \theta. \end{aligned}$$

On the other hand,

$$(13.55) \quad \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x_0^2 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \theta^2 + \lim_{n \rightarrow \infty} A_0^n (\mathcal{F}^{(n+1)}g)(\theta),$$

and so, combining (13.54)–(13.55), we obtain

$$(13.56) \quad \lim_{n \rightarrow \infty} A_0^n (\mathcal{F}^{(n+1)}g)(\theta) = \theta(1 - \theta).$$

**2.** We know also that

$$(13.57) \quad \begin{aligned} \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x_0 y_{0,0} Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\ = \lim_{n \rightarrow \infty} \vartheta_n^2 + \left( A_0^n - \frac{E_0^2}{E_0 c_0 + e_0 + E_0 K_0 e_0} \right) (\mathcal{F}^{(n+1)}g)(\vartheta_n) \\ = \theta^2 + \theta(1 - \theta) = \theta \end{aligned}$$

and

$$(13.58) \quad \begin{aligned} \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x y_0 Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\ = \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} x_0 y_{0,0} \left( \mathbf{1}_{\{1\}}(x_0) + \mathbf{1}_{\{0\}}(x_0) + \mathbf{1}_{\{(0,1)\}}(x_0) \right) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\ = \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} y_{0,0} \mathbf{1}_{\{1\}}(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0). \end{aligned}$$

Therefore

$$(13.59) \quad \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} y_{0,0} \mathbf{1}_{\{1\}}(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \lim_{n \rightarrow \infty} \vartheta_n = \theta,$$

and hence

$$(13.60) \quad \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} (1 - y_{(0,0)}) \mathbf{1}_{\{1\}}(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = \theta - \theta = 0.$$

Since  $1 - y_{(0,0)} \geq 0$ , we conclude that if  $x_0 = 1$ , then  $Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0)$  puts all its mass on  $y_{0,0} = 1$  in the limit as  $n \rightarrow \infty$ . Hence

$$(13.61) \quad \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{(x_0, y_{0,0}) = (1, 1)\}) = \theta.$$



From Proposition 13.4 it also follows that

$$\begin{aligned}
 (13.62) \quad & \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} (1-x_0)(1-y_{0,0}) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\
 & = 1 - \theta - \theta + \theta^2 + \lim_{n \rightarrow \infty} \left( A_0^n - \frac{E_0^2}{E_0 c_0 + e_0 + E_0 K_0 e_0} \right) (\mathcal{F}^{(n+1)}g)(\vartheta_n) \\
 & = 1 - \theta.
 \end{aligned}$$

On the other hand,

$$\begin{aligned}
 (13.63) \quad & \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} (1-x_0)(1-y_{0,0}) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\
 & = \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} (1-x_0)(1-y_{0,0}) \left( \mathbf{1}_{\{1\}}(x_0) + \mathbf{1}_{\{0\}}(x_0) + \mathbf{1}_{\{(0,1)\}}(x_0) \right) \\
 & \quad \times Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\
 & = \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} (1-y_{0,0}) \mathbf{1}_{\{0\}}(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) \\
 & = 1 - \theta.
 \end{aligned}$$

Since  $y \in [0, 1]$  and

$$(13.64) \quad \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} \mathbf{1}_{\{0\}}(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = 1 - \theta,$$

it follows that

$$(13.65) \quad \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} y_{0,0} \mathbf{1}_{\{0\}}(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = 0.$$

This implies that if  $x_0 = 0$ , then  $Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0)$  puts all its mass on  $y_{0,0} = 0$  in the limit as  $n \rightarrow \infty$ . Hence

$$\begin{aligned}
 (13.66) \quad & \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{(x_0, y_{0,0}) = (0, 0)\}) = 1 - \theta, \\
 & \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, \{(x_0, y_{0,0}) = (1, 1)\}) = \theta.
 \end{aligned}$$

Now assume (13.44). Then (13.54) still holds. On the other hand, also (13.55) still holds by Proposition 13.3. Therefore we obtain (13.56). On the other hand, by (13.44)

$$(13.67) \quad \lim_{n \rightarrow \infty} \mathcal{F}^{(n+1)}g = \lim_{n \rightarrow \infty} \int_{[0,1] \times [0,1]^{N_0}} g(x_0) Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = 0.$$

Hence (13.45) holds. □

A similar clustering property holds for the kernels associated with blocks.

**Proposition 13.6. [Clustering: blocks]** *Assume*

$$(13.68) \quad \lim_{n \rightarrow \infty} \vartheta_n = \theta.$$

Then

$$\begin{aligned}
 (13.69) \quad & \lim_{n \rightarrow \infty} Q_m^{(n)}(\bar{\vartheta}^{(n)}, \{(x_m, y_{(m,m)}) = (0, 0)\}) = 1 - \theta, \\
 & \lim_{n \rightarrow \infty} Q_m^{(n)}(\bar{\vartheta}^{(n)}, \{(x_m, y_{(m,m)}) = (1, 1)\}) = \theta,
 \end{aligned}$$

if and only if

$$(13.70) \quad \lim_{n \rightarrow \infty} A_m^n = \infty.$$

Consequently,

$$(13.71) \quad \lim_{n \rightarrow \infty} Q_m^{(n)}(\bar{\vartheta}^{(n)}, \{(x_m, y_{(m,m)}) \notin \{(0, 0), (1, 1)\}\}) = 0.$$

*Proof.* We can proceed exactly as in the proof of Proposition 13.5.  $\square$

Finally, we can prove Theorem 5.1.

*Proof.* Note that (13.45) implies (13.70). 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). It follows that, for  $m \in \mathbb{N}_0$ ,

$$(13.72) \quad \begin{aligned} \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, (x_0, y_{0,m}) = (1, 1)) &= \theta, \\ \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, (x_0, y_{0,m}) = (0, 0)) &= 1 - \theta. \end{aligned}$$

Therefore, for  $N \in \mathbb{N}_0$ ,

$$(13.73) \quad \begin{aligned} &\lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \bigcap_{m=0}^N \{(x_0, y_{0,m}) = (1, 1) \text{ or } (x_0, y_{0,m}) = (0, 0)\} \right) \\ &= 1 - \lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \bigcup_{m=0}^N \{(x_0, y_{0,m}) \in [0, 1]^2 \setminus \{(0, 0), (1, 1)\}\} \right) \\ &\geq 1 - \lim_{n \rightarrow \infty} \sum_{m=0}^N Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, y_{0,m}) \in [0, 1]^2 \setminus \{(0, 0), (1, 1)\}\} \right) = 1 - 0 = 1. \end{aligned}$$

Note that

$$(13.74) \quad \begin{aligned} &\lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \bigcap_{m=0}^N \{(x_0, y_{0,m}) = (1, 1) \text{ or } (x_0, y_{0,m}) = (0, 0)\} \right) \\ &= \lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, (y_{0,m})_{0 \leq m \leq N}) = (0, 0^{N+1}) \right. \\ &\quad \left. \text{or } (x_0, (y_{0,m})_{0 \leq m \leq N}) = (1, 1^{N+1})\} \right) = 1. \end{aligned}$$

On the other hand,

$$(13.75) \quad \begin{aligned} &\lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, (y_{0,m})_{0 \leq m \leq N}) = (1, 1^{N+1})\} \right) \\ &\leq \lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, y_{0,0}) = (1, 1)\} \right) = \theta \end{aligned}$$

and

$$(13.76) \quad \begin{aligned} &\lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, (y_{0,m})_{0 \leq m \leq N}) = (0, 0^{N+1})\} \right) \\ &\leq \lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, y_{0,0}) = (0, 0)\} \right) = 1 - \theta. \end{aligned}$$

Hence we conclude that

$$(13.77) \quad \lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, (y_{0,m})_{0 \leq m \leq N}) = (1, 1^{N+1})\} \right) = \theta$$

and

$$(13.78) \quad \lim_{n \rightarrow \infty} Q^{(n)} \left( \bar{\vartheta}^{(n)}, \{(x_0, (y_{0,m})_{0 \leq m \leq N}) = (0, 0^{N+1})\} \right) = 1 - \theta.$$

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 (13.77)–(13.78) we find for every converging subsequence

$$(13.79) \quad \lim_{k \rightarrow \infty} Q^{(n_k)} \left( \bar{\vartheta}^{(n_k)}, \cdot \right) = (1 - \theta) \delta_{(0, 0^{\mathbb{N}_0})} + \theta \delta_{(1, 1^{\mathbb{N}_0})}.$$

We conclude that

$$(13.80) \quad \lim_{n \rightarrow \infty} Q^{(n)}(\bar{\vartheta}^{(n)}, dz_0) = (1 - \theta) \delta_{(0, 0^{\mathbb{N}_0})} + \theta \delta_{(1, 1^{\mathbb{N}_0})},$$

which is the claim in (5.4).  $\square$

### 13.4 Dichotomy finite versus infinite seed-bank

In this section we prove Theorem 5.3.

*Proof.* We investigate for what choices of the sequences  $c, K, e$  defined in (2.5) and (2.10) we meet the *clustering criterion*  $\lim_{n \rightarrow \infty} A_n \rightarrow \infty$  in (5.5). Recall from (4.5) and (5.3) that

$$(13.81) \quad A_n = \frac{1}{2} \sum_{k=0}^{n-1} \frac{E_k}{c_k} \frac{(E_k c_k + e_k)}{(E_k c_k + e_k) + E_k K_k e_k}, \quad E_k = \frac{1}{1 + \sum_{m=0}^{k-1} K_m}.$$

We distinguish between three regimes as  $k \rightarrow \infty$ :

1.  $E_k c_k + e_k \gg E_k K_k e_k$ .
2.  $E_k c_k + e_k \asymp E_k K_k e_k$ .
3.  $E_k c_k + e_k \ll E_k K_k e_k$ .

These regimes correspond to the following scaling for  $A_n$  as  $n \rightarrow \infty$ :

1.  $A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{E_k}{c_k}$ .
2.  $A_n \asymp \sum_{k=0}^{n-1} \frac{E_k}{c_k}$ .
3.  $A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{E_k c_k + e_k}{c_k K_k e_k}$ .

Recall from that (2.14) that

$$(13.82) \quad \rho = \sum_{m \in \mathbb{N}_0} K_m.$$

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  $\lim_{k \rightarrow \infty} E_k K_k = 0$  and hence we are in regime 1. Therefore

$$(13.83) \quad A_n \sim \frac{1}{2(1 + \rho)} \sum_{k=0}^{n-1} \frac{1}{c_k}, \quad n \rightarrow \infty,$$

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 (3.10) and (3.11), which fall in regimes 1 and 2.

**Asymptotically polynomial.** Suppose that

$$(13.84) \quad K_k \sim A k^{-\alpha}, \quad k \rightarrow \infty, \quad A \in (0, \infty), \quad \alpha \in (-\infty, 1).$$

Then

$$(13.85) \quad E_k \sim \frac{1 - \alpha}{A} k^{-(1-\alpha)}, \quad E_k K_k \sim \frac{1 - \alpha}{A} k^{-1}, \quad k \rightarrow \infty.$$

Hence we are in regime 1. Suppose that

$$(13.86) \quad c_k \sim F k^{-\phi}, \quad k \rightarrow \infty, \quad F \in (0, \infty), \quad \phi \in \mathbb{R}.$$

Then

$$(13.87) \quad A_n \sim \frac{1 - \alpha}{2AF} \sum_{k=1}^{n-1} k^{-1+\alpha+\phi}, \quad n \rightarrow \infty,$$

and clustering occurs if and only if  $-\phi \leq \alpha < 1$ . In this case

$$(13.88) \quad \begin{aligned} -\phi < \alpha: \quad A_n &\sim \frac{1-\alpha}{2AF(\alpha+\phi)} n^{\alpha+\phi}, \\ -\phi = \alpha: \quad A_n &\sim \frac{1-\alpha}{2AF} \log n. \end{aligned}$$

The case  $\alpha = 1$  can be included. Then (13.85) becomes

$$(13.89) \quad E_k \sim \frac{1}{A \log k}, \quad E_k K_k \sim \frac{1}{k \log k}, \quad k \rightarrow \infty,$$

so that we are again in regime 1. Now (13.87) becomes

$$(13.90) \quad A_n \sim \frac{1}{2AF} \sum_{k=1}^{n-1} \frac{k^\phi}{\log k}, \quad n \rightarrow \infty,$$

and clustering occurs if and only if  $-\phi \leq 1$ . In this case

$$(13.91) \quad \begin{aligned} -\phi < 1: \quad A_n &\sim \frac{1}{2AF(1+\phi)} \frac{n^{1+\phi}}{\log n}, \\ -\phi = 1: \quad A_n &\sim \frac{1}{2AF} \log \log n. \end{aligned}$$

**Pure exponential.** Suppose that

$$(13.92) \quad K_k = K^k, \quad k \in \mathbb{N}_0, \quad K \in (1, \infty).$$

Then

$$(13.93) \quad E_k = \frac{1}{1 + \sum_{m=0}^{k-1} K^m} = \frac{1}{1 + \frac{K^k - 1}{K - 1}} = \frac{K - 1}{K^k + K - 2}.$$

Suppose that

$$(13.94) \quad e_k = e^k, \quad c_k = c^k, \quad k \in \mathbb{N}_0, \quad e, c \in (0, \infty).$$

Then

$$(13.95) \quad E_k c_k + e_k = \frac{K - 1}{K^k + K - 2} c^k + e^k, \quad E_k K_k e_k = \frac{K - 1}{K^k + K - 2} K^k e^k,$$

and so

$$(13.96) \quad E_k c_k + e_k \sim (K - 1) \left(\frac{c}{K}\right)^k + e^k, \quad E_k K_k e_k \sim (K - 1) e^k, \quad k \rightarrow \infty.$$

For  $c \leq Ke$  we are in regime 2, and hence

$$(13.97) \quad A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{K - 1}{(Kc)^k} \frac{(K - 1) \left(\frac{c}{K}\right)^k + e^k}{(K - 1) \left(\frac{c}{K}\right)^k + Ke^k}, \quad n \rightarrow \infty,$$

which simplifies to

$$(13.98) \quad \begin{aligned} c < Ke: \quad A_n &\sim \frac{1}{2K} \sum_{k=0}^{n-1} \frac{K - 1}{(Kc)^k}, \\ c = Ke: \quad A_n &\sim \frac{K - 1}{2(2K - 1)} \sum_{k=0}^{n-1} \frac{K - 1}{(Kc)^k}. \end{aligned}$$

Clustering occurs if and only if  $Kc \leq 1$ . In this case

$$(13.99) \quad \begin{aligned} Kc < 1: & \quad \sum_{k=0}^{n-1} \frac{1}{(Kc)^k} \sim \frac{1}{1-Kc} (Kc)^{-(n-1)}, \\ Kc = 1: & \quad \sum_{k=0}^{n-1} \frac{1}{(Kc)^k} \sim n. \end{aligned}$$

For  $c > Ke$ , on the other hand, we are in regime 1, and hence

$$(13.100) \quad A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{K-1}{(Kc)^k}, \quad n \rightarrow \infty,$$

for which we can again use (13.99).

The case  $K = 1$  can be included. Then  $E_k = 1/k$  and (13.96) becomes

$$(13.101) \quad E_k c_k + e_k \sim \frac{1}{k} c^k + e^k, \quad E_k K_k e_k \sim \frac{1}{k} e^k, \quad k \rightarrow \infty,$$

we are again in regime 1. Hence

$$(13.102) \quad A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{kc^k}, \quad n \rightarrow \infty,$$

and clustering occurs if and only if  $c \leq 1$ . In that case

$$(13.103) \quad \begin{aligned} c < 1: & \quad A_n \sim \frac{1}{2(1-c)} \frac{1}{n} c^{-(n-1)}, \\ c = 1: & \quad A_n \sim \frac{1}{2} \log n. \end{aligned}$$

□

In the above computations, only regimes 1 and 2 arise. Regime 3 arises, for instance, when

$$(13.104) \quad \lim_{k \rightarrow \infty} \frac{1}{k} \log K_k = \infty, \quad \lim_{k \rightarrow \infty} K_k e_k / c_k = \infty.$$

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:

$$(13.105) \quad \begin{aligned} K_{k-1} e_k \ll c_k: & \quad A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{K_k K_{k-1} e_k}, \\ K_{k-1} e_k \gg c_k: & \quad A_n \sim \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{K_k c_k}. \end{aligned}$$

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.

## A Computation of scaling coefficients

In Appendices A.1–A.2 we spell out a technical computation for the tail of the wake-up time defined in (2.39)–(2.40) in the two parameter regimes given by (3.10)–(3.11). In Appendix A.3 we carry out a computation that is needed in Section 6.1.

### A.1 Regularly varying coefficients

In (2.39), 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

$$(A.1) \quad P(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m} e^{-(e_m/N^m)t} \sim \frac{1}{\chi} \int_c^\infty dm K_m \frac{e_m}{N^m} e^{-(e_m/N^m)t},$$

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  $m \rightarrow \infty$ , we have

$$(A.2) \quad s \sim Bm^{-\beta}N^{-m}$$

and hence

$$(A.3) \quad \begin{aligned} \log s &\sim \log B - \beta \log m - m \log N, \\ \log \frac{1}{s} &= m \log N \left( -\frac{B}{m \log N} + \frac{\beta \log m}{m \log N} + 1 \right) = [1 + o(1)] m \log N, \end{aligned}$$

which gives

$$(A.4) \quad m = [1 + o(1)] \frac{\log(\frac{1}{s})}{\log N}.$$

Thus,

$$(A.5) \quad \frac{1}{s} \frac{ds}{dm} = -\log N - \frac{\beta}{m} = -[1 + o(1)] \log N,$$

which implies

$$(A.6) \quad \frac{ds}{dm} = -[1 + o(1)] s \log N,$$

so that  $s(m)$  is asymptotically decreasing in  $m$ , and

$$(A.7) \quad \frac{dm}{ds} = -[1 + o(1)] (s \log N)^{-1}.$$

Note that if  $c \leq m < \infty$ , then asymptotically  $0 < m^{-\beta}N^{-m} < c^{-\beta}N^{-c} = C_2$ . Doing the substitution, we get

$$(A.8) \quad \begin{aligned} \mathbb{P}(\tau > t) &\sim \frac{1}{\chi} \int_0^{C_2} ds K_m s (s \log N)^{-1} e^{-st} \sim \frac{1}{\chi} \int_0^{C_2} ds A m^{-\alpha} (\log N)^{-1} e^{-st} \\ &\sim \frac{1}{\chi} \int_0^{C_2} ds A \left( \frac{\log(\frac{1}{s})}{\log N} \right)^{-\alpha} (\log N)^{-1} e^{-st} \sim \frac{A}{\chi} \left( \frac{1}{\log N} \right)^{-\alpha+1} \int_0^{C_2} ds \log\left(\frac{1}{s}\right)^{-\alpha} e^{-st}. \end{aligned}$$

Next, put  $st = u$ , so  $s = \frac{u}{t}$  and  $\frac{ds}{du} = \frac{1}{t}$  and  $0 < u < tC_2$ . Then

$$(A.9) \quad \mathbb{P}(\tau > t) \sim \frac{A}{\chi} \left( \frac{1}{\log N} \right)^{-\alpha+1} \frac{1}{t} \int_0^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u}.$$

We will show that

$$(A.10) \quad \frac{A}{\chi} \left( \frac{1}{\log N} \right)^{-\alpha+1} \frac{1}{t} \int_0^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} \asymp \frac{A}{\chi} \left( \frac{1}{\log N} \right)^{-\alpha+1} \frac{1}{t} \int_0^{C_2 t} du \log t^{-\alpha} e^{-u}.$$

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:

$$\begin{aligned}
 \int_0^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} &= \int_0^1 du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} + \int_1^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} \\
 &\leq \int_0^1 du \log\left(\frac{t}{u}\right)^{-\alpha} + \int_1^{C_2 t} du \log t^{-\alpha} e^{-u} \\
 (A.11) \quad &\leq 2^{-\alpha} \int_0^{\frac{1}{t}} du \log\left(\frac{1}{u}\right)^{-\alpha} + 2^{-\alpha} \int_{\frac{1}{t}}^1 du \log t^{-\alpha} + \log t^{-\alpha} [1 - e^{-1}] \\
 &\leq 2^{-\alpha} \Gamma(-\alpha + 1) + 2^{-\alpha} \log t^{-\alpha} \left[1 - \frac{1}{t}\right] + \log t^{-\alpha} [1 - e^{-1}] \\
 &= \log t^{-\alpha} \left[2^{-\alpha} \frac{\Gamma(-\alpha + 1)}{\log t^{-\alpha}} + 2^{-\alpha} \left[1 - \frac{1}{t}\right] + [1 - e^{-1}]\right] \asymp \log t^{-\alpha}.
 \end{aligned}$$

For the lower bound, note that

$$\begin{aligned}
 \int_0^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} &\geq \log(t)^{-\alpha} \int_0^1 du e^{-u} + \log\left(\frac{1}{C_2}\right)^{-\alpha} \int_1^{C_2 t} du e^{-u} \\
 (A.12) \quad &= \log t^{-\alpha} \left[1 - e^{-1} + \frac{\log\left(\frac{1}{C_2}\right)^{-\alpha}}{\log t^{-\alpha}}\right] \asymp \log t^{-\alpha}.
 \end{aligned}$$

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

$$\begin{aligned}
 \int_0^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} &\geq \lim_{u \rightarrow 0} \log\left(\frac{t}{u}\right)^{-\alpha} [1 - e^{-1}] + \log t^{-\alpha} [e^{-1} - e^{-C_2 t}] \\
 (A.13) \quad &= \log t^{-\alpha} [0 + e^{-1} - e^{-C_2 t}] \asymp \log t^{-\alpha}.
 \end{aligned}$$

For the upper bound estimate

$$\begin{aligned}
 &\int_0^{C_2 t} du \log\left(\frac{t}{u}\right)^{-\alpha} e^{-u} \\
 &\leq \log t^{-\alpha} [1 - e^{-1}] + \log\left(\frac{t}{\sqrt{C_2 t}}\right)^{-\alpha} \int_1^{\sqrt{C_2 t}} du e^{-u} + \log\left(\frac{1}{C_2}\right)^{-\alpha} \int_{\sqrt{C_2 t}}^{C_2 t} du e^{-u} \\
 (A.14) \quad &= \log t^{-\alpha} [1 - e^{-1}] + \left(\frac{1}{2}\right)^{-\alpha} \log\left(\frac{t}{C_2}\right)^{-\alpha} [e^{-1} - e^{-\sqrt{C_2 t}}] + \log\left(\frac{1}{C_2}\right)^{-\alpha} [e^{-\sqrt{C_2 t}} - e^{-C_2 t}] \\
 &= \log t^{-\alpha} \left[1 - e^{-1} + \left(\frac{1}{2}\right)^{-\alpha} \left(\frac{\log t - \log C_2}{\log t}\right)^{-\alpha} [e^{-1} - e^{-\sqrt{C_2 t}}] \right. \\
 &\quad \left. + \log\left(\frac{1}{C_2}\right)^{-\alpha} \frac{[e^{-\sqrt{C_2 t}} - e^{-C_2 t}]}{\log t^{-\alpha}}\right] \asymp \log t^{-\alpha}.
 \end{aligned}$$

## A.2 Pure exponential coefficients

In order to satisfy condition in (2.12), 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

$$(A.15) \quad P(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m} e^{-(e_m/N^m)t} \sim \int_M^\infty dm K_m \frac{e_m}{N^m} e^{-(e_m/N^m)t}.$$

Again we put  $s = \frac{e^m}{N^m}$ . Hence

$$(A.16) \quad \log s = m \log\left(\frac{e}{N}\right), \quad m = \frac{\log s}{\log \frac{e}{N}}, \quad \frac{dm}{ds} = \frac{1}{s \log \frac{e}{N}},$$

and

$$(A.17) \quad K_m \sim K^m \sim e^{\log s \frac{\log K}{\log \frac{e}{N}}} \sim s^{\frac{\log K}{\log \frac{e}{N}}}.$$

Since  $s(m)$  is decreasing in  $m$ , putting  $C = (\frac{e}{N})^M$  we obtain

$$(A.18) \quad \mathbb{P}(\tau > t) \sim \int_0^C ds K_m \frac{s}{s \log \frac{e}{N}} e^{-st} \sim \frac{1}{\log \frac{e}{N}} \int_0^C ds s^{\frac{\log K}{\log \frac{e}{N}}} e^{-st}.$$

Substitute  $u = st$ , i.e.,  $\frac{u}{t} = s$ , to get

$$(A.19) \quad \begin{aligned} \mathbb{P}(\tau > t) &\sim \frac{1}{\log \frac{e}{N}} t^{-1 - \frac{\log K}{\log \frac{e}{N}}} \int_0^{Ct} du u^{\frac{\log K}{\log \frac{e}{N}}} e^{-u} \\ &\sim \frac{1}{\log \frac{e}{N}} t^{-\frac{\log(\frac{e}{N}) - \log K}{\log \frac{e}{N}}} \int_0^{Ct} du u^{\frac{\log K}{\log \frac{e}{N}}} e^{-u} \sim \frac{1}{\log \frac{e}{N}} t^{-\frac{\log(\frac{N}{Ke})}{\log \frac{N}{e}}} \int_0^{Ct} du u^{\frac{\log K}{\log \frac{e}{N}}} e^{-u}. \end{aligned}$$

The last integral converges because  $\frac{\log K}{\log(\frac{e}{N})} > -1$ , and

$$(A.20) \quad \int_0^{Ct} du u^{\frac{\log K}{\log \frac{e}{N}}} e^{-u} \leq \int_0^\infty du u^{\frac{\log K}{\log \frac{e}{N}}} e^{-u} = \Gamma\left(\frac{\log K}{\log(\frac{e}{N})} + 1\right).$$

### A.3 Slowly varying functions

Return to Section 6.1. 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 (3.7), we have

$$(A.21) \quad \frac{\varphi(t)}{\varphi(s)} = \exp\left[-\int_{t(s)}^s \frac{du}{u} \psi(u)\right].$$

Since we are interested in  $s \rightarrow \infty$ , we may assume that  $s \gg 1$  and  $t(s) > 1$ , and estimate

$$(A.22) \quad \begin{aligned} \frac{\varphi(t)}{\varphi(s)} &\leq \exp\left[\int_{t(s)}^s \frac{du}{u} \frac{C}{\log u}\right] = \exp[C(\log \log s - \log \log t(s))] \\ &= \exp\left[C \log\left(\frac{\log s}{\log(\varphi(s)^{-1} s^\gamma)}\right)\right] = \exp\left[-C \log\left(\frac{\gamma \log s - \log \varphi(s)}{\log s}\right)\right]. \end{aligned}$$

A similar lower bound holds with the sign reversed. Using that  $\lim_{s \rightarrow \infty} \frac{\log \varphi(s)}{\log s} = 0$ , we get

$$(A.23) \quad \gamma^C \leq \liminf_{s \rightarrow \infty} \frac{\varphi(t)}{\varphi(s)} \leq \limsup_{s \rightarrow \infty} \frac{\varphi(t)}{\varphi(s)} \leq \gamma^{-C}.$$

Both bounds above are positive, so indeed  $\frac{\varphi(t)}{\varphi(s)} \asymp 1$ .

## B Meyer-Zheng topology

### B.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  $\psi_w$ . Note that the Borel- $\sigma$  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

$$(B.1) \quad \psi_w(A) = \int 1_A(t, w(t)) e^{-t} dt = \int_a^b 1_B(w(t)) e^{-t} dt,$$

i.e.,  $\psi_w$  is the image measure of the mapping  $t \rightarrow (t, w(t))$  under the measure  $\lambda(dt) = e^{-t} dt$ . The set of all pseudopaths is denoted by  $\Psi$ . Note that a pseudopath corresponding to  $(w(t))_{t > 0}$  is simply its occupation measure. The following important facts are stated in [MZ84]:



- 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  $\psi: \mathbf{D} \rightarrow \Psi$ ,  $w \mapsto \psi_w$  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  $[0, \infty) \times \mathbb{R}$ .
- Note if  $f$  is a function on  $[0, \infty) \times \mathbb{R}$  and  $w \in \mathbf{D}$ , then

$$(B.2) \quad \psi_w(f) = \int_0^\infty f(t, w(t)) e^{-t} dt.$$

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  $[0, \infty) \times \bar{\mathbb{R}}$ ,

$$(B.3) \quad \lim_{n \rightarrow \infty} \int_0^\infty f(t, w_n(t)) e^{-t} dt = \int_0^\infty f(t, w(t)) e^{-t} dt.$$

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.  $\Psi$  endowed with the pseudopath topology is a Polish space.
- According to [MZ84][Lemma 1], the pseudopath topology on  $\Psi$  is convergence in Lebesgue measure on  $\mathbf{D}$ .

## B.2 Pseudopaths of stochastic processes on a general metric separable space

In [Kur91] the results of [MZ84] 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  $A = [a, b] \times B$ ,  $0 \leq a < b$  and  $B \in \mathcal{B}(E)$ ,

$$(B.4) \quad \psi_{(Z(t, \omega))_{t \geq 0}}(A) = \int_a^b 1_B(Z(t, \omega)) e^{-t} dt.$$

Hence  $\psi_{(Z(t))_{t \geq 0}}$  is a random variable with state space  $\Psi$ , i.e.,  $\psi_{(Z(t))_{t \geq 0}} \in \mathcal{M}(\Psi)$ , the set of probability measures on pseudopaths. Note that

$$(B.5) \quad \mathbb{E}[\psi_{(Z(t))_{t \geq 0}} f] = \mathbb{E}\left[\int_0^\infty f(t, Z(t, \omega)) e^{-t} dt\right] = \mathbb{E}\left[\int_0^\infty f(t, Z(t)) e^{-t} dt\right].$$

**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

$$(B.6) \quad \mathcal{L}[(Z_n(t))_{t \geq 0}] = \mathcal{L}[(Z(t))_{t \geq 0}] \text{ in the Meyer-Zheng topology}$$

if, for all  $f \in \mathcal{C}_b(\Psi)$ ,

$$(B.7) \quad \lim_{n \rightarrow \infty} \mathbb{E}[f(\psi_{(Z_n(t))_{t \geq 0}})] = \mathbb{E}[f(\psi_{(Z(t))_{t \geq 0}})].$$

Let  $\mathcal{C}_m([0, \infty) \times E) \subset \mathcal{C}_b([0, \infty) \times E)$  be the set of functions of the form

$$(B.8) \quad \mathcal{C}_m([0, \infty) \times E) = \left\{ F \in \mathcal{C}_b([0, \infty) \times E) : F(t, x(t)) = \prod_{i=1}^m \int_0^{T_i} f_i(t, x(t)) dt, \right. \\ \left. m \in \mathbb{N}, \forall 1 \leq i \leq m, f_i \in \mathcal{C}_b([0, \infty) \times E), T_i > 0 \right\}.$$

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 [Kur91][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

$$(B.9) \quad \mathbb{E}[F(Z_1)] = \mathbb{E}[F(Z_2)] \quad \forall F \in \mathcal{C}_m,$$

then  $\mathcal{L}[Z_1] = \mathcal{L}[Z_2]$ .

Define

$$(B.10) \quad F(\psi) = \int d\psi \prod_{i=1}^m \int_0^{T_i} f_i(t, x(t)) dt.$$

Recall that a pseudopath  $\psi$  is associated with a path  $w \in M_E[0, \infty)$ . Hence this becomes

$$(B.11) \quad F(\psi_w) = \prod_{i=1}^m \int_0^{T_i} f_i(t, w(t)) dt.$$

Since each pseudopath  $\psi \in \Psi$  is associated with a path in  $M_E[0, \infty)$ ,  $\mathcal{C}_m$  also separates points on  $\Psi$  and hence  $\mathcal{C}_m$  separates measures on  $\Psi$ . This implies that if

$$(B.12) \quad \mathbb{E}[F(\psi_{Z_1})] = \mathbb{E}[F(\psi_{Z_2})] \quad \forall F \in \mathcal{C}_m,$$

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 weaker than the Skorohod topology.

**Lemma B.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*

$$(B.13) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(Z_n(t))_{t\geq 0}] = \mathcal{L}[(Z(t))_{t\geq 0}] \text{ in the Skorohod topology,}$$

then

$$(B.14) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(Z_n(t))_{t\geq 0}] = \mathcal{L}[(Z(t))_{t\geq 0}] \text{ in the Meyer-Zheng topology.}$$

*Proof.* Since we do not know whether  $\Psi$  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

$$(B.15) \quad \lim_{n \rightarrow \infty} \tilde{Z}^n = \tilde{Z} \quad a.s.$$

This implies

$$(B.16) \quad \lim_{n \rightarrow \infty} \psi_{\tilde{Z}^n} = \psi_{\tilde{Z}} \quad a.s.$$

Consequently, for all  $f \in \mathcal{C}_b(\Psi)$ ,

$$(B.17) \quad \lim_{n \rightarrow \infty} \mathbb{E}[f(\psi_{\tilde{Z}^n})] = \mathbb{E}[f(\psi_{\tilde{Z}})].$$

Note that, since  $\mathcal{L}[\tilde{Z}^n] = \mathcal{L}[Z^n]$  and  $\mathcal{L}[\tilde{Z}] = \mathcal{L}[Z]$ , we can use (B.9) and (B.12) 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.17) indeed implies that

$$(B.18) \quad \lim_{n \rightarrow \infty} \mathcal{L}[\psi_{Z^n}] = \mathcal{L}[\psi_Z].$$

□

**Convergence in probability in the Meyer-Zheng topology.** Let  $(S, d)$  be a metric space,  $\mathcal{B}(S)$  denote the Borel- $\sigma$  algebra on  $S$ , and  $\mathcal{P}(S)$  the set of probability measures on  $\mathcal{B}(S)$ . Recall (see e.g. [EK86, Chapter 3]) that the Prohorov metric  $d_P$  on the space  $\mathcal{P}(S)$  is given by

$$(B.19) \quad d_P(\mathbb{P}, \mathbb{Q}) = \inf \{ \epsilon > 0: \mathbb{P}(A) \leq \mathbb{Q}(A^\epsilon) + \epsilon \forall A \in \mathcal{C} \},$$

where  $\mathcal{C} \subset \mathcal{B}(S)$  is the set of all closed sets in  $S$  and  $A^\epsilon = \{x \in S: \inf_{y \in A} d(x, y) < \epsilon\}$ .

**Theorem B.2.** [EK86, Theorem 3.1.2] *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*

$$(B.20) \quad d_P(\mathbb{P}, \mathbb{Q}) = \inf_{\mu \in \mathcal{M}(\mathbb{P}, \mathbb{Q})} \inf\{\epsilon > 0: \mu(\{(x, y): d(x, y) \geq \epsilon\}) \leq \epsilon\}.$$

Moreover, [EK86, Theorem 3.3.1] states that convergence of measures in the Prohorov distance,  $\lim_{n \rightarrow \infty} d_P(\mathbb{P}_n, \mathbb{P}) = 0$ , 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  $\Psi$  with the metric  $d_P$ .

Let  $(\Psi, d_P)$  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

$$(B.21) \quad \forall \delta > 0: \lim_{n \rightarrow \infty} \mathbb{P}[d_P(\psi_{Z^n}, \psi_Z) > \delta] = 0.$$

**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  $\tau: 0 = t_0 < t_1 < \dots < t_n = \infty$ , set

$$(B.22) \quad V_\tau(U) = \sum_{0 \leq i < n} \mathbb{E}\left[\left|\mathbb{E}[U(t_{i+1}) - U(t_i) \mid \mathcal{F}(t_i)]\right|\right]$$

(with  $U(\infty) = 0$ ) and

$$(B.23) \quad V(U) = \sup_{\tau} V_\tau(U).$$

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.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 [MZ84, Theorem 7] for the identification of the limiting semi-martingale.)

**B.3 Proof of key lemmas**

• **Proof of Lemma 8.19.**

*Proof.* Fix  $\delta > 0$ . Then

$$(B.24) \quad \begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{P}[d_P(\psi_{Z_n}, \psi_Z) > \delta] \\ &= \lim_{n \rightarrow \infty} \mathbb{P}\left[\inf_{\mu \in \mathcal{M}(\psi_{Z_n}, \psi_Z)} \inf\{\epsilon > 0: \mu(\{(x, y): d(x, y) \geq \epsilon\}) \leq \epsilon\} > \delta\right] \\ &= \lim_{n \rightarrow \infty} \mathbb{P}[\forall \mu \in \mathcal{M}(\psi_{Z_n}, \psi_Z), \inf\{\epsilon > 0: \mu(\{(x, y): d(x, y) \geq \epsilon\}) \leq \epsilon\} > \delta] \\ &= \lim_{n \rightarrow \infty} \mathbb{P}[\forall \mu \in \mathcal{M}(\psi_{Z_n}, \psi_Z), \mu(\{(x, y): d(x, y) \geq \delta\}) > \delta]. \end{aligned}$$

Let  $\mu_n \in \mathcal{P}([0, \infty] \times E)^2$  be the measure defined by

$$(B.25) \quad \mu_n(A) = \int_0^\infty 1_A((t, Z_n(t)), (t, Z(t))) e^{-t} dt, \quad A \in \mathcal{B}([0, \infty] \times E)^2,$$

such that, for  $B \in \mathcal{B}([0, \infty] \times E)$ ,

$$(B.26) \quad \mu_n(B \times S) = \int_0^\infty 1_B(t, Z_n(t)) 1_S((t, Z(t))) e^{-t} dt = \psi_{Z_n}(B),$$

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.24) that

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} \mathbb{P} [d_P(\psi_{Z_n}, \psi_Z) > \delta] \\
 & \leq \lim_{n \rightarrow \infty} \mathbb{P} [\mu_n(\{(x, y) : d(x, y) \geq \delta\}) > \delta] \\
 & \leq \lim_{n \rightarrow \infty} \mathbb{P} \left[ \int_0^\infty 1_{\{(x, y) : d(x, y) \geq \delta\}}((t, Z_n(t)), (t, Z(t))) e^{-t} dt > \delta \right] \\
 \text{(B.27)} \quad & \leq \lim_{n \rightarrow \infty} \mathbb{P} \left[ \int_0^\infty 1_{\{d(Z_n(t), Z(t)) \geq \delta\}} e^{-t} dt > \delta \right] \\
 & \leq \lim_{n \rightarrow \infty} \frac{1}{\delta} \mathbb{E} \left[ \int_0^\infty d(Z_n(t), Z(t)) e^{-t} dt \right] \\
 & = \lim_{n \rightarrow \infty} \frac{1}{\delta} \int_0^\infty \mathbb{E} [d(Z_n(t), Z(t))] e^{-t} dt = 0.
 \end{aligned}$$

□

• **Proof of Lemma 8.20.**

*Proof.* We have to show that

$$\text{(B.28)} \quad \lim_{n \rightarrow \infty} \mathcal{L} [\psi_{(X_n, Y_n)}] = \mathcal{L} [\psi_{(X, c)}].$$

Hence we must show that, for all  $f \in \mathcal{C}_b(\Psi)$ ,

$$\text{(B.29)} \quad \lim_{n \rightarrow \infty} \mathbb{E}[f(\psi_{(X_n, Y_n)})] = \mathbb{E}[f(\psi_{(X, c)})].$$

We can write

$$\text{(B.30)} \quad |\mathbb{E}[f(\psi_{(X_n, Y_n)}) - f(\psi_{(X, c)})]| \leq |\mathbb{E}[f(\psi_{(X_n, Y_n)}) - f(\psi_{(X_n, c)})]| + |\mathbb{E}[f(\psi_{(X_n, c)}) - f(\psi_{(X, c)})]|.$$

Since  $\lim_{n \rightarrow \infty} \mathbb{E}[d(Y_n(t), c)] = 0$  implies  $\lim_{n \rightarrow \infty} \mathbb{E}[d((X_n(t), Y_n(t)), (X_n(t), c))] = 0$ , it follows from Lemma 8.19 that, for all  $\delta > 0$ ,

$$\text{(B.31)} \quad \lim_{n \rightarrow \infty} \mathbb{P} [d_P(\psi_{(X_n, Y_n)}, \psi_{(X_n, c)})] = 0.$$

Hence, for all  $f \in \mathcal{C}_b(\Psi)$ ,

$$\text{(B.32)} \quad \lim_{n \rightarrow \infty} |\mathbb{E}[f(\psi_{(X_n, Y_n)}) - f(\psi_{(X_n, c)})]| = 0.$$

To see that the second term in the right-hand side of (B.30) tends to zero, note that we can define

$$\text{(B.33)} \quad \tilde{f}(\psi_x) = f(\psi_{x, c}).$$

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  $\psi_{x_n} \rightarrow \psi_x$  if and only if, for all  $\delta > 0$ ,

$$\text{(B.34)} \quad \lim_{n \rightarrow \infty} \int_0^\infty 1_{\{d(x_n(t), x(t)) > \delta\}} e^{-t} dt = 0.$$

Therefore  $\psi_{x_n} \rightarrow \psi_x$  implies that, for all  $\delta > 0$ ,

$$\text{(B.35)} \quad \lim_{n \rightarrow \infty} \int_0^\infty 1_{\{d((x_n(t), c), (x(t), c)) > \delta\}} e^{-t} dt = 0,$$

and hence  $\psi_{x_n, c} \rightarrow \psi_{x, c}$ . Therefore

$$(B.36) \quad \lim_{n \rightarrow \infty} \tilde{f}(\psi_{x_n}) = \lim_{n \rightarrow \infty} f(\psi_{(x_n, c)}) = f(\psi_{(x, c)}) = \tilde{f}(\psi_x)$$

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  $f \in \mathcal{C}_b(\Psi)$ ,

$$(B.37) \quad \lim_{n \rightarrow \infty} |\mathbb{E}[f(\psi_{(X_n, c)}) - f(\psi_{(X, c)})]| = \lim_{n \rightarrow \infty} |\mathbb{E}[\tilde{f}(\psi_{X_n}) - \tilde{f}(\psi_X)]| = 0.$$

Therefore also the second term on the right-hand side of (B.30) tends to 0.  $\square$

• **Proof of Lemma 8.21.**

*Proof.* For part (a), suppose that  $\lim_{n \rightarrow \infty} \psi_{x_n} = \psi_x$ . Then, since convergence in pseudopath space is convergence in measure, we have, for all  $\delta > 0$ ,

$$(B.38) \quad \lim_{n \rightarrow \infty} \int_0^\infty 1_{\{d(x_n(t), x(t)) > \delta\}} e^{-t} dt = 0.$$

Since  $f$  is a continuous function, this implies that, for all  $\epsilon > 0$ ,

$$(B.39) \quad \lim_{n \rightarrow \infty} \int_0^\infty 1_{\{d(f(x_n(t)), f(x(t))) > \epsilon\}} e^{-t} dt = 0.$$

and we conclude that  $\lim_{n \rightarrow \infty} \psi_{f(x_n)} = \psi_{f(x)}$ . Hence  $h$  is indeed continuous.

For part (b), recall that

$$(B.40) \quad \lim_{n \rightarrow \infty} \mathcal{L}[X_n] = \mathcal{L}[X] \text{ in the Meyer-Zheng topology}$$

implies that, for all  $g \in \mathcal{C}_b(\Psi)$ ,

$$(B.41) \quad \lim_{n \rightarrow \infty} \mathbb{E}[g(\psi_{X_n})] = \mathbb{E}[g(\psi_X)].$$

Since  $h: \Psi \rightarrow \Psi$  is continuous, we have for all  $g \in \mathcal{C}_b(\Psi)$  that  $g \circ h \in \mathcal{C}_b(\Psi)$ . Hence

$$(B.42) \quad \lim_{n \rightarrow \infty} \mathbb{E}[g(\psi_{f(X_n)})] = \lim_{n \rightarrow \infty} \mathbb{E}[g \circ h(\psi_{X_n})] = \mathbb{E}[g \circ h(\psi_X)] = \mathbb{E}[g(\psi_{f(X)})].$$

We conclude that

$$(B.43) \quad \lim_{n \rightarrow \infty} \mathcal{L}[f(X_n)] = \mathcal{L}[f(X)] \text{ in the Meyer-Zheng topology.}$$

$\square$

• **Proof of Lemma 10.17.** Suppose that  $\lim_{n \rightarrow \infty} \psi_{(x_n, y_n)} = \psi_{(x, y)}$ . Since convergence of pseudopaths is convergence in Lebesgue measure, we have

$$(B.44) \quad \lim_{n \rightarrow \infty} \int_0^\infty 1_{\{d[(x_n, y_n), (x, y)] > \delta\}} e^{-t} dt = 0$$

and, consequently,

$$(B.45) \quad \lim_{n \rightarrow \infty} \int_0^\infty 1_{\{d[x_n, x] > \delta\}} e^{-t} dt = 0.$$

Therefore  $\lim_{n \rightarrow \infty} \psi_{x_n} = \psi_x$ . 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

$$(B.46) \quad \tilde{f}(\psi_{(x, y)}) = f(\psi_x).$$

Then  $\tilde{f} \in \mathcal{C}_b(\Psi(E^2))$  and

$$(B.47) \quad \lim_{n \rightarrow \infty} \tilde{f}(\psi_{(x_n, y_n)}) = \lim_{n \rightarrow \infty} f(\psi_{x_n}) = f(\psi_x) = \tilde{f}(\psi_{(x_n, y_n)}).$$

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

$$(B.48) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(X_n(s), Y_n(s))_{s>0}] = \mathcal{L}[(X(s), Y(s))_{s>0}] \text{ in Meyer-Zheng topology,}$$

then for all  $f \in \mathcal{C}_b(\Psi(E^2))$  we have

$$(B.49) \quad \lim_{n \rightarrow \infty} \mathbb{E}[f(\psi_{(X_n, Y_n)})] = \mathbb{E}[f(\psi_{(X, Y)})].$$

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.46), we obtain for all  $f \in \mathcal{C}_b(\Psi(E))$  that

$$(B.50) \quad \lim_{n \rightarrow \infty} \mathbb{E}[f(\psi_{(X_n)})] = \lim_{n \rightarrow \infty} \mathbb{E}[\tilde{f}(\psi_{(X_n, Y_n)})] = \mathbb{E}[\tilde{f}(\psi_{(X, Y)})] = \mathbb{E}[\tilde{f}(\psi_X)].$$

We conclude that

$$(B.51) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(X_n(s))] = \mathcal{L}[(X(s))_{s>0}] \text{ in Meyer-Zheng topology}$$

and, similarly,

$$(B.52) \quad \lim_{n \rightarrow \infty} \mathcal{L}[(Y_n(s))] = \mathcal{L}[(Y(s))_{s>0}] \text{ in Meyer-Zheng topology.}$$

## References

- [AB16] 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.
- [Arr79] R.A. Arratia. *Coalescing Brownian Motions on the Line*. PhD thesis, University of Wisconsin-Madison, USA, 1979.
- [BCGdH95] J.-B. Baillon, Ph. 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.
- [BCGdH97] J.-B. Baillon, Ph. 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.
- [BGCEK15] 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.
- [BGCKWB16] 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.
- [BGT87] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1987.
- [Bil99] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.
- [CG86] J.T. Cox and D. Griffeath. Diffusive clustering in the two-dimensional voter model. *Ann. Probab.*, 14:347–370, 1986.

- [DG93a] D.A. 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.
- [DG93b] D.A. Dawson and A. Greven. Multiple scale analysis of interacting diffusions. *Probab. Theory Relat. Fields*, 95:467–508, 1993.
- [DG96] D.A. Dawson and A. Greven. Multiple space-time scale analysis for interacting branching models. *Electron. J. Probab.*, 1:paper no. 14, 1–84, 1996.
- [DG14] D.A. Dawson and A. Greven. *Spatial Fleming-Viot Models with Selection and Mutation*, volume 2092 of *Springer Lecture Notes in Math*. Springer Verlag, 2014.
- [DGV95] D.A. 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.
- [DGW05] D.A. Dawson, L.G. Gorostiza, and A. Wakolbinger. Degrees of transience and recurrence and hierarchical random walks. *Potential Anal.*, 22:305–350, 2005.
- [dH01] 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.
- [EF96] S.N. Evans and K. Fleischmann. Cluster formation in a stepping stone model with continuous, hierarchically structured sites. *Ann. Probab.*, 24:1926–1952, 1996.
- [EK86] S.N. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. John Wiley, New York, 1986.
- [FG94] K. Fleischmann and A. Greven. Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Rel. Fields*, 98:517–566, 1994.
- [FG96] K. Fleischmann and A. Greven. Time-space analysis of the cluster formation in interacting diffusions. *Electr. J. Probab.*, 1:1–46, 1996.
- [Gar88] J. Gärtner. On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, 137:197–248, 1988.
- [GdHK18] 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.
- [GdHKK14] 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.
- [GHO20] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: well-posedness, duality and equilibrium. *Preprint*, 2020.
- [GHO21] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: finite-systems scheme. *Preprint*, 2021.
- [JM86] 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.
- [Kle96] A. Klenke. Different clustering regimes in hierarchically interacting diffusions. *Ann. Probab.*, 24:660–697, 1996.
- [Kle97] A. Klenke. Multiscale analysis of clusters in spatial branching models. *Ann. Probab.*, 25:1670–1711, 1997.

- [Kur91] Thomas G. Kurtz. Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Ann. Probab.*, 19(3):1010–1034, 1991.
- [LJ11] J.T. Lennon and S.E. Jones. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9:119–130, 2011.
- [MZ84] P.A. Meyer and W.A. Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré, Prob. Stat.*, 20:353–372, 1984.
- [RW00] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge University Press, 2000. 494 pages.
- [RY99] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, New York, 1999.
- [SF83] S. Sawyer and J. Felsenstein. Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.*, 20:1–10, 1983.
- [SS80] T. Shiga and A. Shimizu. Infinite-dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.*, 20:395–416, 1980.
- [Win02] A. Winter. Multiple scale analysis of branching processes under the palm distribution. *Electr. J. Probab.*, 7:paper no. 13, 72 pp, 2002.
- [YW71] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971.