Spatial populations with seed-bank:
finite-systems scheme

Andreas Greven[1], Frank den Hollander[2]

19 September 2022

Abstract

This is the third in a series of four papers in which we consider a system of interacting Fisher-Wright diffusions with seed-bank. Individuals carry type $\heartsuit$ or $\diamond$, live in colonies, and are subject to resampling and migration as long as they are *active*. Each colony has a structured seed-bank into which individuals can retreat to become *dormant*, suspending their resampling and migration until they become active again. As geographic space labelling the colonies we consider a countable Abelian group $\mathbb{G}$ endowed with the discrete topology. Our goal is to understand in what way the seed-bank enhances genetic diversity and causes new phenomena.

In [GHO22b] we showed that the system of continuum stochastic differential equations, describing the population in the large-colony-size limit, has a unique strong solution. We further showed that if the system starts from an initial law that is invariant and ergodic under translations with a density of $\heartsuit$ that is equal to θ, then it converges to an equilibrium ν_θ whose density of $\heartsuit$ also is equal to θ. Moreover, ν_θ exhibits a dichotomy of *coexistence* (= locally multi-type equilibrium) versus *clustering* (= locally mono-type equilibrium). We identified the parameter regimes for which these two phases occur, and found that these regimes are different when the mean wake-up time of a dormant individual is finite or infinite.

The goal of the present paper is to establish the *finite-systems scheme*, i.e., identify 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. Since the finite system exhibits clustering, we focus on the regime where the infinite system exhibits coexistence, which consists of two sub-regimes. If the wake-up time has finite mean, then the scaling time turns out to be proportional to the volume of the truncated geographic space, 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 of $\heartsuit$ that evolves according to a *renormalised Fisher-Wright diffusion* and ultimately gets trapped in either 0 or 1. On the other hand, if the wake-up time has infinite mean, then the scaling time turns out to grow faster than the volume of the truncated geographic space, 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. For slow growth the scaling limit is the same as when the wake-up time has finite mean, while for fast growth the scaling limit is different, namely, the density of $\heartsuit$ initially remains fixed at θ, afterwards makes *random switches* between 0 and 1 on a range of different time scales, driven by individuals in deep seed-banks that wake up, until it finally gets *trapped* in either 0 or 1 on the time scale where the individuals in the deepest seed-banks wake up. Thus, the system evolves through a sequence of *partial clusterings* (or partial fixations) before it reaches *complete clustering* (or complete fixation).

Keywords: Resampling, migration, dormancy, seed-bank, duality, coexistence versus clustering, finite-systems scheme.

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

[1] Department Mathematik, Universität Erlangen-Nürnberg, Cauerstrasse 11, D-91058 Erlangen, Germany
greven@mi.uni-erlangen.de

[2] Mathematisch Instituut, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, NL
denholla@math.leidenuniv.nl

Contents

1   Introduction

Section 1.1 outlines the goal of the paper. Section 1.2 recalls the three models introduced in [GHO22b, Section 2]. Section 1.3 contains some key observations made in [GHO22b, Section 2] concerning the choice of model parameters and initial laws, and the role of duality and diffusion function. Section 1.4 summarises the core results in [GHO22b, Section 3] and provides a brief outline of the remainder of the paper. Section 2 describes our main results for the finite-system scheme. Section 3 contains preparations and Sections 4–6 provide proofs. For background and motivation we refer the reader to [GHO22b, Section 1].

Sections 1.2–1.4 are largely copied from [GHO22b], but are needed to set the stage.

1.1   Goal

In [GHO22b] we considered a system of interacting Fisher-Wright diffusions with seed-bank. Individuals carry type $\heartsuit$ or $\diamond$, live in colonies, and are subject to resampling and migration as long as they are *active*. Each colony has a structured seed-bank into which individuals can retreat to become *dormant*, suspending their resampling and migration until they become active again. As geographic space labelling the colonies we considered a countable Abelian group $\mathbb{G}$ endowed with the discrete topology. We showed that the system of continuum stochastic differential equations, describing the population in the large-colony-size limit, has a unique strong solution. We further showed that if the system starts from an initial law that is invariant and ergodic under translations with a density of $\heartsuit$ that is equal to θ, then the system converges to an equilibrium ν_θ whose density of $\heartsuit$ also is equal to θ. Moreover, ν_θ exhibits a dichotomy of *coexistence* (= locally multi-type equilibrium) versus *clustering* (= locally mono-type equilibrium). We identified the parameter regimes for which these two phases occur, and found that these regimes are qualitatively different when the mean wake-up time of an individual is finite or infinite.

The goal of the present paper is to establish the *finite-systems scheme*, i.e., identify 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. Since the finite system exhibits clustering, we focus on the regime where the infinite system exhibits coexistence. To allow for a proper truncation limit, we additionally assume that $\mathbb{G}$ is *profinite*, i.e., $\mathbb{G}$ is (isomorphic to) the limit of a projective system of finite groups endowed with the discrete topology. Our main findings are the following:

- If the wake-up time has finite mean, then the scaling time is proportional to the volume of the truncated geographic space 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 of $\heartsuit$ that evolves according to a *renormalised Fisher-Wright diffusion* and ultimately gets trapped in either 0 or 1 (i.e., the system ultimately clusters).

- If the wake-up time has infinite mean, then the scaling time grows faster than the volume of the truncated geographic space, 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. For slow growth the scaling limit is the same as when the wake-up time has finite mean, while for fast growth the scaling limit is different, namely, the density of $\heartsuit$ initially remains fixed at θ, afterwards makes *random switches* between 0 and 1 on a range of different time scales, driven by individuals in deeper seed-banks that wake up, until it finally gets *trapped* in either 0 or 1 on the time scale where the individuals in the deepest seed-banks wake up. Thus, the system evolves through a sequence of *partial clusterings* (or partial fixations) before it reaches *complete clustering* (or complete fixation).

The finite-systems scheme underpins the relevance of systems with an infinite geographic space and an infinite seed-bank for the description of systems with a large but finite geographic space and a large but finite seed-bank per colony.

1.2   Migration, resampling and seed-bank: three models

In this section, which is largely copied from [GHO22b, Section 2], we restrict ourselves to recalling the definition of the three models introduced in [GHO22b]. In [GHO22b, Appendix A] we explained how the system of continuum stochastic differential equations that is our object of study arises as the large-colony-size limit of a sequence of discrete individual-based systems. In what follows we will interpret properties of the continuum system in terms of the underlying discrete systems in order to provide the proper intuition.

We consider populations of individuals that can be of two types: ♡ and ◇. These populations are located in a *geographic space* $\mathbb{G}$ with a group structure, namely, $(\mathbb{G}, +)$ is a countable Abelian group with $+$ a group action. We are interested in the frequencies of types in the various locations. To describe the migration we use a *migration kernel* $a(\cdot, \cdot)$, which is a $\mathbb{G} \times \mathbb{G}$ matrix of transition rates that determines a continuous-time random walk and satisfies:

Assumption 1.1. [Migration kernel]

$$(1.1) \qquad a(i, j) = a(0, j - i) \quad \forall\, i, j \in \mathbb{G}, \qquad \sum_{i \in \mathbb{G}} a(0, i) < \infty, \qquad a(\cdot, \cdot) \text{ is irreducible.} \qquad \square$$

In each of the three models to be described below, the population at a location consist of an *active part* and a *dormant part*. The seed-bank at a given location is the repository of the dormant population at that location (and for two of the models has an internal structure that regulates the wake-up time). For each active individual that becomes dormant a randomly chosen dormant individual becomes active, i.e., the active and the dormant population *exchange* individuals (see Fig. 1). This guarantees that the sizes of the active and the dormant population stay fixed over time.

Model 1: single-layer seed-bank. For $i \in \mathbb{G}$ and $t \geq 0$, let $x_i(t)$ denote the fraction of individuals in colony i of type ♡ that are active at time t, and $y_i(t)$ the fraction of individuals in colony i of type ♡ that are dormant at time t. These fractions evolve according to the SSDE

$$(1.2) \qquad \mathrm{d}x_i(t) = \sum_{j \in \mathbb{G}} a(i, j)\,[x_j(t) - x_i(t)]\,\mathrm{d}t + \sqrt{d x_i(t)[1 - x_i(t)]}\,\mathrm{d}w_i(t)$$
$$+ Ke\,[y_i(t) - x_i(t)]\,\mathrm{d}t,$$

$$(1.3) \qquad \mathrm{d}y_i(t) = e\,[x_i(t) - y_i(t)]\,\mathrm{d}t,$$

where $(w_i(t))_{t \geq 0}$, $i \in \mathbb{G}$, are independent standard Brownian motions. The first term in (1.2) describes the *migration* of individuals (at rate $a(i, j)$ from j to i), the second term in (1.2) describes the *resampling* of individuals (at rate $d \in (0, \infty)$ for all i). The third term in (1.2) together with the term in (1.3) describe the *exchange* of active and dormant individuals (at rate $e \in (0, \infty)$ for all i). See Fig. 1 for an illustration.

Figure 1: The evolution in a single colony in Model 1. Individuals are subject to migration (see Fig. 3), resampling and exchange with the seed-bank.

The factor $K \in (0, \infty)$ allows for an *asymmetry* between the sizes of the active and the dormant population. Indeed, because we are tracking *fractions* of individuals of type ♡, we have

$$(1.4) \qquad K = \frac{\text{size dormant population}}{\text{size active population}},$$

and this ratio is the same for all colonies.

The *state space* of the system is

$$(1.5) \qquad E = [0, 1]^{\mathbb{S}}, \qquad \mathbb{S} = \mathbb{G} \times \{A, D\},$$

endowed with the product topology, where A denotes the residence of the active population and D the repository of the dormant population. Alternatively, we may also think of two populations, one active and one dormant, and write $E = ([0, 1] \times [0, 1])^{\mathbb{G}}$. Accordingly, the configuration of the system at time t is written as

(1.6) $Z(t) = (z_u(t))_{u \in \mathbb{S}}$

with $z_u(t) = x_i(t)$ if $u = (i, A)$ and $z_u(t) = y_i(t)$ if $u = (i, D)$, respectively,

(1.7) $Z(t) = (z_i(t))_{i \in \mathbb{G}}$

with $z_i(t) = (x_i(t), y_i(t))$. Via duality, our SSDE can be understood in terms of tuples of random walks on $\mathbb{G}$ with internal states A and D [GHO22b, Section 2].
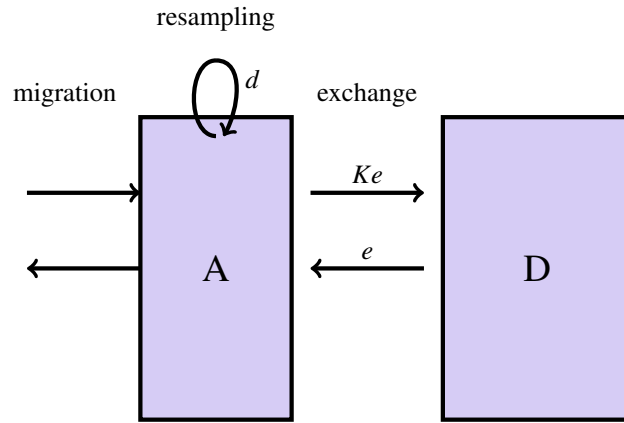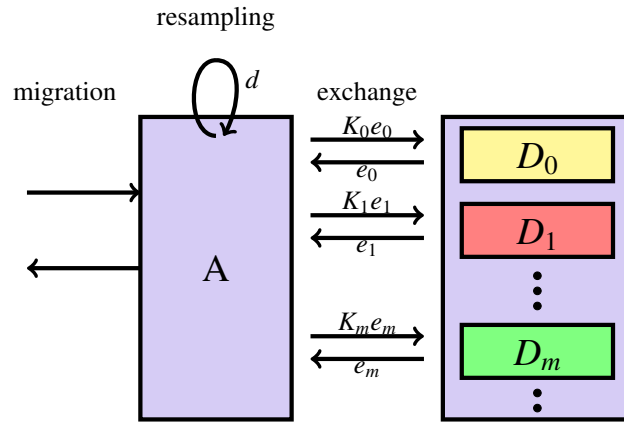

Figure 2: The evolution in a single colony in Model 2. Individuals are subject to migration (see Fig. 3), resampling and exchange with the seed-bank, as in Model 1. Additionally, when individuals become dormant they get a colour and when they become active they loose their colour.

Model 2: multi-layer seed-bank. In order to allow for fat tails in the wake-up times of individuals and still preserve the Markov property, we enrich the state space. Namely, we allow individuals to become dormant with a colour that is drawn randomly from an infinite sequence of colours, labelled by $\mathbb{N}_0$. See Figs. 2 and 3 for an illustration.

As before, let $x_i(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are active at time t, but now let $y_{i,m}(t)$ denote the fraction of individuals in colony i of type $\heartsuit$ that are dormant with colour m at time t. Suppose that active individuals exchange with dormant individuals with colour m at rate $e_m \in (0, \infty)$. Then the SSDE in (1.2)–(1.3) is replaced by

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

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

where the factor $K_m \in (0, \infty)$ captures the asymmetry between the size of the active population and the m-dormant population, i.e., similarly as in (1.4),

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

where $K_m \in (0, \infty)$ is the same for all colonies. The state space is

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

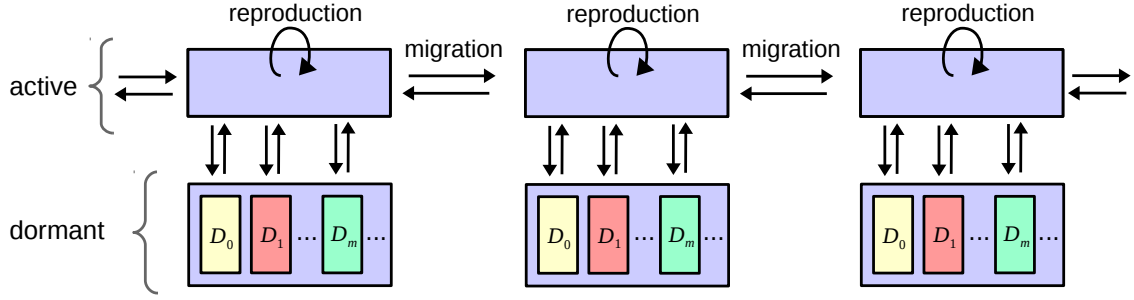Figure 3: The spatial evolution in Model 2. Each colony has its own seed-bank. Resampling occurs within the active population in each colony. Migration occurs between active populations in different colonies (top layer). Exchange between the active and the dormant population occurs in each colony (transitions between the top layer and the bottom layer).

endowed with the product topology, where A denotes the residence of the active population and D_m the repository of the m-dormant population. Alternatively, we may think of infinitely many populations, one active and all the others dormant, and write $E = ([0, 1] \times [0, 1]^{\mathbb{N}_0})^{\mathbb{G}}$. Accordingly, the configuration of the system at time t is written as

$$(1.12) \qquad Z(t) = (z_u(t))_{u \in \mathbb{S}}$$

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

$$(1.13) \qquad Z(t) = (z_i(t))_{i \in \mathbb{G}}$$

with $z_i(t) = (x_i(t), (y_{i,m}(t))_{m \in \mathbb{N}_0})$.

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

$$(1.14) \qquad a_m(i, j) = a_m(0, j - i) \quad \forall\, i, j \in \mathbb{G},\, m \in \mathbb{N}_0, \qquad \sup_{m \in \mathbb{N}_0} \sum_{i \in \mathbb{G}} a_m(0, i) < \infty,$$

and replacing (1.8)–(1.9) by

$$(1.15) \qquad \mathrm{d}x_i(t) \;=\; \sum_{j \in \mathbb{G}} a(i, j)\,[x_j(t) - x_i(t)]\,\mathrm{d}t + \sqrt{dx_i(t)[1 - x_i(t)]}\,\mathrm{d}w_i(t)$$
$$+ \sum_{j \in \mathbb{G}} \sum_{m \in \mathbb{N}_0} K_m e_m\, a_m(j, i)\,[y_{j,m}(t) - x_i(t)]\,\mathrm{d}t,$$

$$(1.16) \qquad \mathrm{d}y_{i,m}(t) \;=\; \sum_{j \in \mathbb{G}} e_m\, a_m(i, j)\,[x_j(t) - y_{i,m}(t)]\,\mathrm{d}t, \qquad m \in \mathbb{N}_0.$$

Here, the third term in (1.15) together with the term in (1.16) describe the *migration* of individuals and the *switch of colony* of individuals during the exchange between active to dormant (at rate $a_m(i, j)$ between i and j, m). The state space and the configuration at time t are the same as in (1.11). Also (1.10) remains the same.

Duality. As shown in [GHO22b, Section 2], the three models have a tractable dual, in which paths of individuals (see Fig. 3) are *reversed in time* to become ancestral lineages of individuals (see Fig. 4). In the dual, which plays a crucial role in the analysis, lineages evolve as independent continuous-time Markov processes with state space

$$(1.17) \qquad \mathbb{S} = \mathbb{G} \times \{A, D\} \text{ (Model 1)}, \qquad \mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\} \text{ (Models 2–3)},$$

and with transition kernel

(1.18) $b^{(1)}(\cdot, \cdot)$ in (A.1), $b^{(2)}(\cdot, \cdot)$ in (A.4), $b^{(3)}(\cdot, \cdot)$ in (A.5),

and coalesce at rate d when they are at the same site in $\mathbb{G}$ and are both active. Unlike in the original system, there is *no exchange* between active and dormant lineages in the dual, which makes the dual easier to work with. The duality relations are *mixed-moment relations* linking the original process to the dual process and vice versa. The dual is therefore called a *moment dual*.

For $g \neq dg_{\text{FW}}$ no tractable dual is available and we need to argue by comparison with models that have a tractable dual.
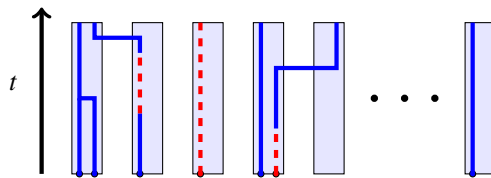

Figure 4: Picture of the evolution of lineages in the spatial coalescent. The purple blocks depict the colonies, the blue lines the active lineages, and the red dashed lines the dormant lineages. Blue lineages can migrate and become dormant, (i.e., become red dashed lineages). Two blue lineages can coalesce when they are at the same colony. Red dashed dormant lineages first have to become active (blue) before they can coalesce with other blue and active lineages or migrate. Note that the dual runs *backwards in time*. The collection of all lineages determines the genealogy of the system.

1.3 Observations: $\rho < \infty$ versus $\rho = \infty$

In this section, which is also largely copied from [GHO22b, Section 2], we recall a number of important observations regarding the models introduced in Section 1.2.

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

(1.19) $$\chi = \sum_{m \in \mathbb{N}_0} K_m e_m < \infty$$

in order for active *lineages* in the dual not to become dormant instantly. The advantage of having infinitely many colours is that it allows us to have *wake-up times with fat tails* and at the same time *preserve the Markov property for the evolution* of the system. Indeed, in Model 2 the rate for an active lineage to become dormant is χ, the go-to-sleep time σ of an active lineage has law

(1.20) $$\mathbb{P}(\sigma > t) = e^{-\chi t}, \qquad t \geq 0,$$

while the wake-up time τ of a dormant lineage has law

(1.21) $$\mathbb{P}(\tau > t) = \sum_{m \in \mathbb{N}_0} \frac{K_m e_m}{\chi} e^{-e_m t}, \qquad t \geq 0,$$

where $K_m e_m / \chi$ is the probability that a dormant lineage has colour m (when the system is in equilibrium). It is possible to talk about paths in the forward time direction rather than about lineages in the backward time direction, but this requires an enrichment of the state spaces and the introduction of historical processes. We will not elaborate on this extension.

Recall that in the continuum limit the size of the active population is scaled to 1. Note that

(1.22) $$\mathbb{E}[\sigma] = \frac{1}{\chi}, \qquad \mathbb{E}[\tau] = \frac{\rho}{\chi}$$

with

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

We saw in [GHO22b] that $\rho < \infty$ ('finite-size seed-bank') and $\rho = \infty$ ('infinite-size seed-bank') represent *different regimes* for the long-time behaviour of the system (see Section 1.4 below). In particular, the criterion whether coexistence or clustering prevails is different. In the present paper we will see that also the scaling in the finite-systems scheme is different.

Initial laws. We need to specify the law $\mu(0)$ from which the *initial configuration* is drawn. Let $\mathcal{P}(E)$ denote the set of probability measures on E. Define

$$(1.24) \quad \begin{aligned} \mathcal{T} &= \{\mu \in \mathcal{P}(E): \ \mu \text{ is invariant under translations in } \mathbb{G}\}, \\ \mathcal{T}^{\text{erg}} &= \{\mu \in \mathcal{T}: \ \mu \text{ is ergodic under translations in } \mathbb{G}\}, \end{aligned}$$

where translation stands for group action (recall (1.7) and (1.13)). For Model 1, as well as for Models 2 and 3 with $\rho < \infty$, $\mu(0)$ can be any element of $\mathcal{T}^{\text{erg}}$. However, for Models 2–3 with $\rho = \infty$ an extra restriction is needed, namely, we additionally require that $\mu(0)$ is *colour regular*:

$$(1.25) \quad \theta^{\bullet}_{\mu(0)} = \lim_{m \to \infty} \mathbb{E}_{\mu(0)}[y_{0,m}] \text{ exists.}$$

This condition gives us control over the *deep seed-banks*, which play a dominant role when $\rho = \infty$. (For $\mu(0) \in \mathcal{T} \setminus \mathcal{T}^{\text{erg}}$ (1.25) is replaced by the requirement that $\lim_{m \to \infty} \mathbb{E}_{\mu(0)}[y_{0,m} \mid \mathcal{B}]$ exists $\mu(0)$-a.s. with $\mathcal{B}$ the translation invariant sigma-field.) Accordingly, we define

$$(1.26) \quad \begin{aligned} \mathcal{T}^{\bullet} &= \{\mu \in \mathcal{T}: \ \mu \text{ is colour regular}\}, \\ \mathcal{T}^{\text{erg},\bullet} &= \{\mu \in \mathcal{T}^{\text{erg}}: \ \mu \text{ is colour regular}\}. \end{aligned}$$

Remark 1.2. Note that the set $\mathcal{T}^{\bullet}$ is not closed in the weak topology as a subset of $\mathcal{T}$. To remedy this problem, consider the set $\mathcal{T}^{\bullet} \times [0,1]$ endowed with the metric that is the sum of the metric for weak convergence on $\mathcal{T}$ and the Euclidean metric on $[0,1]$. Consider the subset $D = \{(\mu, \theta^{\bullet}_{\mu}): \ \mu \in \mathcal{T}^{\bullet}\}$. Clearly, D has a countable dense subset, consisting of sequences of truncated seed-banks that are extended to infinite seed-banks by repeating the value $\theta^{\bullet}_{\mu}$ in the seed-bank direction. Since D is also complete and metric, it is a Polish space in the stronger topology. □

We write $\mu(t)$ to denote the law evolved from $\mu(0)$ at time t, and further define

$$(1.27) \quad \mathcal{I} = \{\mu \in \mathcal{T}: \ \mu \text{ is invariant under the evolution}\}.$$

Topologies. On the set $\mathcal{T}$ we use the topology of weak convergence. On the set $\mathcal{T}^{\bullet}$, however, we need a stronger topology, which we call the *topology uniform weak convergence* and which is defined on the subset

$$(1.28) \quad \mathcal{T}^{\bullet,*} = \Big\{\mu \in \mathcal{T}^{\bullet}: \ \lim_{|m-m'| \to \infty} \text{Cov}_{\mu}(y_{0,m}, y_{0,m'}) = 0\Big\}.$$

In [GHO22b] we showed that

$$(1.29) \quad \hat{\theta} = \lim_{M \to \infty} \frac{x_0 + \sum_{m=0}^{M} K_m y_{0,m}}{1 + \sum_{m=0}^{M} K_m}$$

exists in $L_2(\mu)$ for every $\mu \in \mathcal{T}^{\bullet,*}$, which is the limiting density of $\heartsuit$.

Examples. Three natural examples of Abelian groups and migration kernels are:

- $\mathbb{G} = \mathbb{Z}^d$ the Euclidean lattice of dimension d, and

$$(1.30) \quad a(\cdot, \cdot) \text{ irreducible,} \quad \sum_{i \in \mathbb{Z}^d} i\, a(0, i) = 0, \quad \sum_{i \in \mathbb{Z}^d} |i|^2 a(0, i) < \infty.$$

  Transient migration corresponds to $d \geq 3$ (see [Spi64, Section II.8]).

- $\mathbb{G} = \mathbb{Z}$, and

$$(1.31) \quad a(\cdot, \cdot) \text{ symmetric,} \quad a(0, i) \sim Q\, |i|^{-1-q}, |i| \to \infty, \quad Q \in (0, \infty),\ q \in (0, 2).$$

  Transient migration corresponds to $q \in (0, 1)$ (see [Spi64, Section II.8]).

- $\mathbb{G} = \Omega_N$ the hierarchical group of order N given by

$$(1.32) \qquad \Omega_N = \left\{ i = (i_k)_{k \in \mathbb{N}_0} \in \{0, \dots, N-1\}^{\mathbb{N}_0} : \sum_{k \in \mathbb{N}_0} i_k < \infty \right\},$$

endowed with the hierarchical distance $d(i, j) = \inf\{k \in \mathbb{N}_0 : i_l = j_l \ \forall\, l \geq k\}$, and

$$(1.33) \qquad a(i, j) = \sum_{k \geq d(i,j)} \frac{c_{k-1}}{N^{k-1}} \frac{1}{N^k}, \quad i, j \in \Omega_N, i \neq j, \qquad a(i, i) = 0, \quad i \in \Omega_N,$$

where $(c_k)_{k \in \mathbb{N}_0}$ are coefficients satisfying $\limsup_{k \to \infty} \frac{1}{k} \log c_k < \log N$. The latter is the kernel of transition rates for the *hierarchical random walk* that, for each $k \in \mathbb{N}$, at rate $c_{k-1} N^{-(k-1)}$ chooses the k-ball around it current location and moves to a uniformly random location in this ball. Transient migration corresponds to $\sum_{k \in \mathbb{N}_0} (1/c_k) < \infty$ (see [DGW04], [DGW05]).

General diffusion function. In (1.2), (1.8) and (1.15) we can replace the diffusion functions dg_{FW}, $d \in (0, \infty)$, with $g_{\mathrm{FW}}(x) = x(1-x)$, $x \in [0, 1]$, the Fisher-Wright diffusion function, by a general diffusion function in the class $\mathcal{G}$ defined by

$$(1.34) \qquad \mathcal{G} = \left\{ g \colon [0, 1] \to [0, \infty) \colon g(0) = g(1) = 0, \ g(x) > 0 \ \forall\, x \in (0, 1), \ g \text{ Lipschitz} \right\}.$$

This class is appropriate because the diffusion stays confined to $[0, 1]$, yet can go everywhere in $[0,1]$ ([Bre68, Section 16.7]). Picking $g \neq dg_{\mathrm{FW}}$ amounts to allowing the resampling rate to be *state-dependent*, which is an important extension from a biological perspective. The resampling rate in state x equals $g(x)/x(1-x)$, $x \in (0, 1)$. An example is the Ohta-Kimura diffusion function $g(x) = [x(1-x)]^2$, $x \in [0, 1]$, for which the resampling rate is equal to the genetic diversity of the colony [OK73].

Trapping time for finite geographic space. Let

$$(1.35) \qquad \begin{aligned} \text{Model 1:} \qquad & H = \inf\left\{ t \geq 0 \colon (x(t), y(t)) \in \{(0, 0)^{\mathbb{G}}, (1, 1)^{\mathbb{G}}\} \right\}, \\ \text{Models 2–3:} \qquad & H = \inf\left\{ t \geq 0 \colon (x(t), y(t)) \in \{(0, 0^{\mathbb{N}_0})^{\mathbb{G}}, (1, 1^{\mathbb{N}_0})^{\mathbb{G}}\} \right\}, \end{aligned}$$

be the time until trapping in one of the mono-type states. We note that if the geographic space $\mathbb{G}$ would be *finite* and the seed-bank would be labeled by a *finite* set rather than $\mathbb{N}_0$, then

$$(1.36) \qquad \int_{[0,1]} \mathrm{d}x \, \frac{x(1-x)}{g(x)} < \infty \quad \Longrightarrow \quad \mathbb{P}(H < \infty) = 1.$$

The integral criterion makes $\{0, 1\}$ accessible for the Fisher-Wright diffusion g_{FW} with $g_{\mathrm{FW}}(x) = x(1-x)$, $x \in [0, 1]$, but not for the Ohta-Kimura diffusion g_{OK} with $g_{\mathrm{OK}}(x) = (x(1-x))^2$, $x \in [0, 1]$ (see [Bre68, Chapter 16, Section 7]).

1.4 Core results for the infinite system

In this section, which is largely copied from [GHO22b, Section 3], we summarise the core results for the infinite system in order to set the stage for the results for the finite system that will be presented in Section 2.

In [GHO22b] we showed that, for each of the three models introduced in Section 1.2, the system converges to a *unique equilibrium* ν_θ that depends on a single parameter θ, namely, the density of type $\heartsuit$ in the population (active or dormant) under the initial law $\mu(0)$ (see (2.29) and (2.55)–(2.58) below). This initial law is assumed to be an element of $\mathcal{T}^{\mathrm{erg}}$ (recall (1.24)) for $\rho < \infty$ and an element of $\mathcal{T}^{\mathrm{erg},\bullet}$ (recall (1.26)) for $\rho = \infty$. Moreover, we showed that ν_θ exhibits a dichotomy of *coexistence* (= locally multi-type equilibrium) versus *clustering* (= locally mono-type equilibrium), and identified the *parameter regimes* for which coexistence, respectively, clustering occurs.

Regularity conditions. For $\rho = \infty$ we required certain *regularity conditions* on top of Assumption 1.1, which we list next.

Assumption 1.3. [Regularity conditions for $\rho = \infty$]
(i) The migration is *symmetric*, i.e.,

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

and the time-t transition kernel satisfies

(1.38) $\qquad t \mapsto a_t(0, 0)$ is regularly varying at infinity.

(ii) The seed-bank coeffcients are *polynomial*, i.e.,

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

Consequently,

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

with $\gamma = \frac{\alpha+\beta-1}{\beta}$ and $C = \frac{A}{\beta\chi} B^{1-\gamma}\, \Gamma(\gamma)$, where Γ is the Gamma-function. $\qquad\qquad\square$

Examples satisfying (1.38) can be found in [Hug95, Chapter 3]. The conditions on α, β in (1.39) guarantee that $\rho = \infty$, $\chi < \infty$.

Dichotomy. In [GHO22b, Sections 5–7] we showed that

- coexistence occurs if and only if the *average total joint activity time* of two lineages in the dual *without coalescence*, both starting from $(0, A)$, is finite,

where joint activity means that the two lineages are active and are at the same site. (see also Appendix A). We further identified the *parameter regime* for coexistence, namely, for $\rho < \infty$ we found that, subject to Assumption 1.1, coexistence occurs if and only if

(1.41) $\qquad I_{\hat{a}} = \displaystyle\int_1^\infty \hat{a}_t(0, 0)\, \mathrm{d}t < \infty,$

while for $\rho = \infty$, subject to Assumptions 1.1–1.3, coexistence occurs if and only if

(1.42) $\qquad I_{\hat{a}, \gamma} = \displaystyle\int_1^\infty t^{-(1-\gamma)/\gamma} \hat{a}_t(0, 0)\, \mathrm{d}t < \infty.$

Here, $\hat{a}_t(0, 0)$ is the probability that the migration with symmetrised kernel $\hat{a}(i, j) = \frac{1}{2}[a(i, j) + a(j, i)]$, $i, j \in \mathbb{G}$, starting from 0 is back at 0 at time t, and γ is the exponent in (1.40). For $\rho = \infty$ we showed that the claim may fail without the symmetry assumption $a(i, j) = a(j, i)$, $i, j \in \mathbb{G}$. Remarkably, (1.41) depends on the migration kernel only, while (1.42) depends on the migration kernel and on the asymptotics of the seed-banks coefficients K_m and e_m. The diffusion function g plays no role: either there is coexistence for all $g \in \mathcal{G}$ or for no $g \in \mathcal{G}$.

Clearly, (1.42) shows that there is an interesting competition between migration and seed-bank, while a comparison of (1.41)–(1.42) shows that the seed-bank *enhances genetic diversity*. The criterion in (1.41) corresponds to the symmetrised migration being *transient*. The criterion in (1.42) (for symmetric migration) is less stringent and may even be met when the migration is *recurrent*. For instance, it holds as soon as $\gamma \in (0, \frac{1}{2})$, irrespective of the migration (because $\hat{a}_t(0, 0) \le 1$ for all t), while for $\gamma \in [\frac{1}{2}, 1]$ it holds for certain classes of recurrent migration (see the examples below).

Examples. To illustrate (1.42) we consider three examples.

- On $\mathbb{G} = \mathbb{Z}^d$, if $\hat{a}(\cdot, \cdot)$ satisfies (1.30), then $\hat{a}_t(0, 0) \sim C t^{-d/2}$, $t \to \infty$, for some $C \in (0, \infty)$ [Spi64, Chapter II, Section 7] and so (1.42) amounts to the requirement that

  (1.43) $\qquad \dfrac{1 - \gamma}{\gamma} + \dfrac{d}{2} > 1,$

  i.e, $\gamma \in (0, 1]$ when $d \ge 3$ (transient migration), $\gamma \in (0, 1)$ when $d = 2$ (critically recurrent migration), and $\gamma \in (0, \frac{2}{3})$ when $d = 1$ (strongly recurrent migration).

- On $\mathbb{G} = \mathbb{Z}$, if $\hat{a}(\cdot, \cdot)$ satisfies (1.31), then $\hat{a}_t(0,0) \asymp t^{-1/q}$ [Hug95, Section 3.3.5], and so (1.42) amounts to the requirement that

$$(1.44) \qquad \frac{1-\gamma}{\gamma} + \frac{1}{q} > 1,$$

  i.e., $\gamma \in (0,1]$ when $q \in (0,1)$ (transient migration), $\gamma \in (0,1)$ when $q = 1$ (critically recurrent migration) and $\gamma \in (0, \frac{q}{2q-1})$ when $q \in (1,2)$ (strongly recurrent migration).

- On $\mathbb{G} = \Omega_N$, if $\hat{a}(\cdot, \cdot)$ satisfies (1.33) with $c_k = c^k$, then $\hat{a}_t(0,0) \asymp t^{-1-\delta_N}$ [GHO22a, Section 3.2], and so (1.42) amounts to the requirement that

$$(1.45) \qquad \frac{1-\gamma_N}{\gamma_N} + \delta_N > 0,$$

  i.e., $\gamma_N \in (0,1]$ when $\delta_N \in (0,\infty)$ (transient migration), $\gamma_N \in (0,1)$ when $\delta_N = 0$ (critically recurrent migration), and $\gamma_N \in (0, \frac{1}{1-\delta_N})$ when $\delta_N \in (-\infty, 0)$ (strongly recurrent migration). The inequality in (1.45) holds if and only if $\log N \times \log(Kc) > \log c \times \log(K^2 e)$ (recall that $c < N$ and $Ke < N$). When $Kc = 1$, the latter condition holds for all N if and only if $K > 1$ and $Ke^2 > 1$ (recall that $K \geq 1$ is needed to ensure that $\rho = \infty$). For N large enough it holds when $Kc > 1$ and fails when $Kc < 1$.

Remark 1.4. [Dichotomy with modulation] As shown in [GHO22b, Sections 3.2 and 6.6], it is possible to modulate (1.40)–(1.39) by slowly varying functions. Namely, if (1.40) is replaced by $\mathbb{P}(\tau > t) \sim C\,\phi(t)\,t^{-\gamma}$, $t \to \infty$, with ϕ slowly varying at infinity, then $I_{\hat{a},\gamma}$ is replaced by

$$(1.46) \qquad I_{\hat{a},\gamma,\phi} = \int_1^\infty \hat{\phi}(t)^{-1/\gamma}\, t^{-(1-\gamma)/\gamma} \hat{a}_t(0,0)\, \mathrm{d}t,$$

where $\hat{\phi}(t) = \phi(t)$ when $\gamma \in (0,1)$ and $\hat{\phi}(t) = \int_1^t \phi(s)\, s^{-1}\, \mathrm{d}s$ when $\gamma = 1$. □

Remark 1.5. [Dichotomy for hierarchical seed-bank] As argued in [GHO22a, Section 2], on $\mathbb{G} = \Omega_N$ it is natural to consider a *hierarchical seed-bank*, which amounts to replacing e_m by e_m/N^m in (1.8)–(1.9), and to assume that, instead of (1.39),

$$(1.47) \qquad K_m = K^m, \quad e_m = e^m, \qquad m \in \mathbb{N}_0,\ K, e \in (0, \infty).$$

(To guarantee that $\chi < \infty$ in (1.19) and $\rho = \infty$ in (1.23), we need $Ke < N$ and $K \geq 1$.) As shown in [GHO22a, Section 3.2], the parameters in Remark 1.4 become

$$(1.48) \qquad \gamma_N = \frac{\log(N/Ke)}{\log(N/e)}, \quad \varphi(t) \asymp 1, \quad \hat{\varphi}(t) \asymp \begin{cases} 1, & K \in (1,\infty), \\ \log t, & K = 1, \end{cases}$$

while, if $c_k = c^k$, $k \in \mathbb{N}_0$, $c \in (0,\infty)$, then

$$(1.49) \qquad a_t(0,0) \asymp t^{-1-\delta_N},$$

where

$$(1.50) \qquad \delta_N = \frac{\log c}{\log(N/c)}.$$

(To guarantee that $\sum_{i \in \Omega_N} a(0,i) < \infty$ in (1.1), we need $c < N$.) Note that $\gamma_N = 1$ for all N when $K = 1$, while $\gamma_N < 1$ for all N when $K \in (1,\infty)$, but with $\gamma_N \to 1$ as $N \to \infty$. Further note that $\delta_N \to 0$ as $N \to \infty$ for all $c \in (0,\infty)$. Thus, the hierarchical mean-field limit $N \to \infty$ corresponds to a *critically infinite mean wake-up time for the seed-bank* and a *critically recurrent migration*. □

Equilibria. In [GHO22b] we also analysed the *equilibria* of the infinite system, i.e., the family of extremal invariant measures

$$(1.51) \qquad (\nu_\theta)_{\theta \in [0,1]}$$

parametrised by the density of $\heartsuit$. This family depends on all the parameters of the model, i.e., the migration kernel, the seed-bank coefficients and the diffusion function. We showed that ν_θ is associated and mixing for all $\theta \in [0, 1]$, and that $\mathbb{E}_{\nu_\theta}[x_0] = \theta$ and $\mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta$ for all $m \in \mathbb{N}_0$ (so that ν_θ is colour regular). We showed that for $\rho < \infty$ the deep seed-banks are *random* under ν_θ with a strictly positive variance, while (subject to (1.39) and (1.25)) for $\rho = \infty$ the deep seed-banks are asymptotically *deterministic* under ν_θ, i.e., converge in law to θ in every colony.

We found that the seed-bank *reduces volatility*, i.e., the active components have a smaller variance than in the model without seed-bank. We further showed that $\theta \mapsto \nu_\theta$ is continuous in the weak topology for Model 1 and for Model 2 with $\rho < \infty$, and is continuous in the uniform weak topology for Model 2 with $\rho = \infty$.

Domains of attraction. In [GHO22b, Sections 5 and 6] we showed that the sets $\mathcal{R}_\theta^{(1)}$, $\mathcal{R}_\theta^{(2)}$ and $\mathcal{R}_\theta^{(2),\bullet}$ defined below are the *domains of attraction* of ν_θ, $\theta \in [0, 1]$, in Model 1, Model 2 with $\rho < \infty$ and Model 2 with $\rho = \infty$, respectively. See [GHO22b, Definitions 5.6 and 6.5]. We recall the notation

$$
(1.52) \qquad
\begin{aligned}
&\text{Model 1:} \quad \mathbb{S} = \mathbb{G} \times \{A, D\}, \\
&\text{Model 2:} \quad \mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\},
\end{aligned}
$$

and, for $u \in \mathbb{S}$,

$$
(1.53) \qquad z_u = \begin{cases} x_i, & u = (i, A), \\ y_i, & u = (i, D), \end{cases}, \qquad z_u = \begin{cases} x_i, & u = (i, A), \\ y_{i,m}, & u = (i, D_m). \end{cases}
$$

We write $b_t^{(1)}(\cdot, \cdot)$ and $b_t^{(2)}(\cdot, \cdot)$ to denote the time-t transition kernels of the dual (whose transition kernel is defined in (A.1) and (A.4)).

Definition 1.6. [Liggett conditions]
(A) Model 1: $\mathcal{R}_\theta^{(1)}$ is the set of measures $\mu \in \mathcal{T}^{\mathrm{erg}}$ satisfying:

$$
(1.54) \qquad (*)_\theta \qquad
\begin{aligned}
&(1) \; \forall \, u_1 \in \mathbb{S}: \\
&\lim_{t \to \infty} \sum_{u_2 \in \mathbb{S}} b_t^{(1)}(u_1, u_2) \, \mathbb{E}_\mu[z_{u_2}] = \theta, \\
&(2) \; \forall \, u_1, u_2 \in \mathbb{S}: \\
&\lim_{t \to \infty} \sum_{u_3, u_4 \in \mathbb{S}} b_t^{(1)}(u_1, u_3) \, b_t^{(1)}(u_2, u_4) \, \mathbb{E}_\mu[z_{u_3} z_{u_4}] = \theta^2.
\end{aligned}
$$

(B) Model 2 with $\rho < \infty$: $\mathcal{R}_\theta^{(2)}$ is the set of measures $\mu \in \mathcal{T}^{\mathrm{erg}}$ satisfying:

$$
(1.55) \qquad (*)_\theta \qquad
\begin{aligned}
&(1) \; \forall \, u_1 \in \mathbb{S}: \\
&\lim_{t \to \infty} \sum_{u_2 \in \mathbb{S}} b_t^{(2)}(u_1, u_2) \, \mathbb{E}_\mu[z_{u_2}] = \theta, \\
&(2) \; \forall \, u_1, u_2 \in \mathbb{S}: \\
&\lim_{t \to \infty} \sum_{u_3, u_4 \in \mathbb{S}} b_t^{(2)}(u_1, u_3) \, b_t^{(2)}(u_2, u_4) \, \mathbb{E}_\mu[z_{u_3} z_{u_4}] = \theta^2.
\end{aligned}
$$

(C) Model 2 with $\rho = \infty$: $\mathcal{R}_\theta^{(2),\bullet}$ is the set

$$
(1.56) \qquad \mathcal{R}_\theta^{(2),\bullet} = \{\mu \in \mathcal{R}_\theta^{(2)}: \; \mu \text{ is colour regular}\}. \qquad \square
$$

The Liggett conditions are equivalent to saying that, for $k = 1, 2$ and $u_1 \in \mathbb{S}$, $\sum_{u_2 \in \mathbb{S}} b_t^{(k)}(u_1, u_2) z_{u_2}$ converges in $L_2(\mu)$ to θ as $t \to \infty$.

Outline. Section 2 states our main theorems: a detailed description of the finite-systems scheme for Model 1, Model 2 with $\rho < \infty$ and Model 2 with $\rho = \infty$. Section 3 contains preparatory lemmas and observations that play a central role in the paper. Sections 4–6 are devoted to the proofs of the main theorems. In Appendix A we show that the seed-bank reduces the volatility of the active components. In Appendix B we recall an *abstract scheme* from [CG94b] that lists general conditions for the existence of a finite-systems scheme, as well as a method from [DGV95] to check some of these conditions in concrete settings. The proofs in Sections 4–5 rely on this abstract scheme. The proofs in Section 6 follow a separate route. Appendix C computes the asymptotic fraction of time spent in the active state by a lineage in the dual, which controls the various scaling regimes. In Appendix D we speculate about how in Model 2 with $\rho = \infty$ and fast-growing seed-bank the crossover from partial clustering to complete clustering may take place.

2 Results: Finite-systems scheme

This section contains our main results for the finite-system scheme. In Section 2.1 we provide the general setting. In Section 2.2 we focus on Model 1 (Theorems 2.6 and 2.9). In Section 2.3 we prepare for Model 2 by introducing two regimes: slow growing seed-bank and fast-growing seed-bank. In Sections 2.4–2.5 we analyse Model 2 for $\rho < \infty$ (Theorem 2.10), respectively, $\rho = \infty$ (Theorems 2.11–2.12). We will not consider Model 3 because it behaves similarly as Model 2. In Section 2.6 we list a few open problems.

2.1 General setting

What does the dichotomy of the infinite system imply for *large but finite systems*, for which clustering always prevails? In the *coexistence regime* this question can be analysed by establishing what is called the *finite-systems scheme* [CG90], [CGS95], i.e., by identifying 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. Our target will be to show that the finite system adopts the equilibrium of the infinite system but with a *random density* of $\heartsuit$ (recall (1.29)), and that the law of the latter evolves, on the proper time scale, as a diffusion on $[0, 1]$ with a *renormalised diffusion function*.

Projective system.   Throughout the sequel the following assumption is in force on the geographic space:

Assumption 2.1. [Profinite geographic space] $\mathbb{G}$ is *profinite*, i.e.,

(2.1) $\mathbb{G}$ is (isomorphic to) the limit of a projective system $(\mathbb{G}_n)_{n\in\mathbb{N}}$
      of finite groups endowed with the discrete toplogy.   □

Key examples are

- $\mathbb{G} = \mathbb{Z}^d$ and $\mathbb{G}_n = [-n, n]^d \cap \mathbb{Z}^d \pmod{2n}$ the n-torus (viewed as a quotient group).

- $\mathbb{G} = \Omega_N$ and $\mathbb{G}_n = (\Omega_N)_n$ the n-ball (viewed as a subgroup).

We restrict the $\mathbb{G}$ to $\mathbb{G}_n$, i.e., we keep (1.2)–(1.3) and (1.8)–(1.9) but replace the migration kernel by

$$(2.2) \qquad a^n(i, j) = \sum_{\substack{k\in\mathbb{G} \\ k\downarrow j}} a(i, k), \qquad i, j \in \mathbb{G}_n,$$

where $k \downarrow j$ means that k is projected onto j.

Definition 2.2. [Mixing time] The *mixing time* of the truncated migration is defined as the minimal sequence $(\psi_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \psi_n = \infty$ such that

$$(2.3) \qquad \lim_{s\to\infty} \limsup_{n\to\infty} \sup_{i\in\mathbb{G}_n} \left| |\mathbb{G}_n|\, a^n_{s\psi_n}(0, i) - 1 \right| = 0,$$

where $a^n_t(\cdot, \cdot)$ is the time-t transition kernel associated with (2.2).   □

In words, the migration on $\mathbb{G}_n$ mixes on time scale ψ_n. Note that

$$(2.4) \qquad a^n_t(0, i) \geq a_t(0, i) \qquad \forall\, n \in \mathbb{N},\, i \in \mathbb{G}_n,\, t \geq 0.$$

Also note that, for symmetric random walk (recall (1.37)),

$$(2.5) \qquad a^n_t(0, i) \leq a^n_t(0, 0), \quad a_t(0, i) \leq a_t(0, 0), \qquad \forall\, n \in \mathbb{N},\, i \in \mathbb{G}_n,\, t \geq 0.$$

Assumption 2.3. [Relations between the transitions kernels]
(i) Prior to the mixing time the random walk on $\mathbb{G}_n$ is close to the random walk on $\mathbb{G}$:

$$(2.6) \qquad \lim_{n\to\infty} \sup_{t=o(\psi_n)} \sup_{i\in\mathbb{G}_n} \frac{a^n_t(0, i) - a_t(0, i)}{a_t(0, i)} = 0.$$

(ii) Until the mixing time the random walk on $\mathbb{G}_n$ is comparable to the random walk on $\mathbb{G}$:

$$(2.7) \qquad \exists\, C < \infty\colon\ C^{-1} a_t(0, i) \leq a^n_t(0, i) \leq C a_t(0, i) \quad \forall\, i \in \mathbb{G}_n\ \forall\, 0 \leq t \leq \psi_n.$$ □

Remark 2.4. [Three examples of mixing time]

(1) On $\mathbb{G} = \mathbb{Z}^d$, for $\hat{a}(\cdot, \cdot)$ irreducible with $\sum_{i \in \mathbb{Z}^d} i\, \hat{a}(0, i) = 0$ and $\sum_{i \in \mathbb{Z}^d} |i|^2 \hat{a}(0, i) < \infty$, it is known that $\psi_n \asymp n^2$ ('diffusive mixing'; see [BR76, Corollary 22.3], [Cox89, Proposition 2.8], [CGS95, Lemma 2.1]). In particular, $\psi_n \ll |\mathbb{G}_n|$ when $d \geq 3$, $\psi_n \asymp |\mathbb{G}_n|$ when $d = 2$, and $\psi_n \gg |\mathbb{G}_n|$ when $d = 1$. The crossover occurs at $d = 2$ (critically recurrent migration).

(2) On $\mathbb{G} = \mathbb{Z}$, for $\hat{a}(\cdot, \cdot)$ symmetric with $\hat{a}(0, i) \sim Q|i|^{-1-q}$, $|i| \to \infty$, $q \in (0, 2)$, $Q \in (0, \infty)$, it is known that $\psi_n \asymp n^\delta$ (the proof is an exercise in Fourier analysis). In particular, $\psi_n \ll |\mathbb{G}_n|$ when $q \in (0, 1)$, $\psi_n \asymp |\mathbb{G}_n|$ when $q = 1$, and $\psi_n \gg |\mathbb{G}_n|$ when $q \in (1, 2)$. The crossover occurs at $q = 1$ (critically recurrent migration).

(3) On $\mathbb{G} = \Omega_N$, for the hierarchical random walk with coefficients $(c_k)_{k \in \mathbb{N}_0}$, it is known that $\psi_n \asymp N^{n-1}/c_{n-1} = |\mathbb{G}_n|/Nc_{n-1}$ (the average time until the migration chooses horizon n). In particular, $\psi_n \ll |\mathbb{G}_n|$ when $\lim_{k \to \infty} c_k = \infty$, $\psi_n \asymp |\mathbb{G}_n|$ when $\lim_{k \to \infty} c_k \in (0, \infty)$, and $\psi_n \gg |\mathbb{G}_n|$ when $\lim_{k \to \infty} c_k = 0$. If $c_k = c^k$, then the crossover occur at $c = 1$ (critically recurrent migration). $\qquad\square$

Remark 2.5. [Open questions] Three interesting questions in the general setting are:

(I) Is it true that $\psi_n = o(|\mathbb{G}_n|)$ for transient random walk?

(II) Is it true that $\hat{a}_{\psi_n}(0, 0) \asymp |\mathbb{G}_n|^{-1}$?

(III) Do (2.6)–(2.7) hold always?

The answer is yes for the three examples in Remark 2.4, but appears not to be known in general. $\qquad\square$

Initial laws. Note that the state spaces for the finite system are

$$(2.8) \quad \begin{aligned} E_n &= ([0, 1] \times [0, 1])^{\mathbb{G}_n} && \text{Model 1,} \\ E_n &= E_{M_n, n} = ([0, 1] \times [0, 1])^{M_n+1})^{\mathbb{G}_n} && \text{Models 2–3,} \end{aligned}$$

when the seed-bank is truncated at colour M_n. We need to decide how to choose the initial law μ_n on $\mathbb{E}_n$ in such a way that it properly links up with the initial law μ on E (recall (1.5) and (1.11)), which is assumed to be translation invariant (= invariant under the group action; recall (1.24)). One way to go about is to let μ_n be the restriction of μ to E_n. This works perfectly well for $\mathbb{G} = \Omega_N$, because $\mathbb{G}_n = (\Omega_N)_n$ is a subgroup of $\mathbb{G}$, but does not work for $\mathbb{G} = \mathbb{Z}^d$, because μ_n is not translation invariant on $\mathbb{G}_n = \mathbb{Z}^d \cap [-n, n]^d \pmod{2n}$. A standard way out is to pick μ_n translation invariant such that μ_n converges to μ weakly as $n \to \infty$. A natural way to achieve this is by picking

$$(2.9) \quad \mu_n((x^n, y^n)) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_n} \mu(\tau_i(x^n, y^n)), \qquad (x^n, y^n) = (x_i(t), y_i(t))_{i \in \mathbb{G}_n},$$

where τ_i, $i \in \mathbb{G}_n$, is the *shift* on E_n defined by $(\tau_i(x^n, y^n))_j = (x^n, y^n)_{i+j}$, $i, j \in \mathbb{G}_n$. (We use the upper index n to indicate that (x^n, y^n) lives on E_n, but suppress this index from the single components.)

Note that (2.9) reads as

$$(2.10) \quad \mu_n = \mu \circ \Phi_n$$

with $\Phi_n \colon E \to E_n$ the translation invariant *restriction operator* defined by $\Phi_n(x, y) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_n} \tau_i(x^n, y^n)$. We will write (2.10) as $\mu_n = \Phi_n \mu$ by using the same symbol Φ_n for the restriction operator acting on $\mathcal{T}$, and put

$$(2.11) \quad \mathcal{T}_n = \{\Phi_n \mu \colon \mu \in \mathcal{T}\}.$$

Conversely, we write

$$(2.12) \quad \mu = \mu_n \circ \tilde{\Phi}_n$$

with $\tilde{\Phi}_n \colon E_n \to E$ the translation invariant *extension operator* defined by $\tilde{\Phi}_n(x^n, y^n) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_n} \tilde{\tau}_i(x^n, y^n)$, where $\tilde{\tau}_i$ is the periodic continuation of $\tau_i(x^n, y^n)$ from E_n to E in the geographic coordinate and the constant continuation of $\tau_i(x^n, y^n)$ from $\{0, \dots, M_n\}$ to $\mathbb{N}_0$ in the seed-bank coordinate by the value

$$(2.13) \quad \frac{x_i + \sum_{m=0}^{M_n} K_m y_{i,m}}{1 + \sum_{m=0}^{M_n} K_m}.$$

We will write (2.12) as $\mu = \tilde{\Phi}_n \mu$ by using the same symbol $\tilde{\Phi}_n$ for the extension operator acting on $\mathcal{T}_n$.

2.2 Model 1

Set-up. We denote the system restricted to E_n by

$$(2.14) \qquad (x^n(t), y^n(t))_{t\geq 0}.$$

The *empirical measure* of the system at time t is defined by

$$(2.15) \qquad \mathcal{E}^n(t) = \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} \delta_{\tau_i(x^n, y^n)}.$$

We consider the empirical process

$$(2.16) \qquad \left(\mathcal{E}^n(s\beta_n)\right)_{s\geq 0},$$

where β_n is a large time scale that needs to be properly chosen. We write $\mu_n(s\beta_n)$ to denote the law of $(x^n(s\beta_n), y^n(s\beta_n))$ when we choose $\mu_n(0)$ as in (2.9).

Macroscopic variable. For large n we expect that $\mathcal{E}^n(s\beta_n)$ and $\mu^n(s\beta_n)$ are controlled by a random process on the set $\{\nu_\theta \colon \theta \in [0,1]\}$ of *extremal invariant measures* of the infinite system that is determined by a *macroscopic variable* $\hat{\theta}^n(s\beta_n)$ of the configuration on E_n, whose limit as $n \to \infty$ is a *conserved quantity* of the infinite system, i.e., is itself a consistent estimator (recall (1.29)). This macroscopic variable is given by

$$(2.17) \qquad \hat{\theta}^n(t) = \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} \frac{x_i(t) + Ky_i(t)}{1 + K}, \qquad t \geq 0,$$

which evolves as

$$(2.18) \qquad \mathrm{d}\hat{\theta}^n(t) = \frac{1}{|\mathbb{G}_n|\,(1 + K)} \sum_{i\in\mathbb{G}_n} \sqrt{g(x_i(t))}\,\mathrm{d}w_i(t),$$

where the migration cancels out because of the averaging over the geographic space, and the exchange with the seed-bank cancels out because of the averaging over the seed-bank. Thus, in particular,

$$(2.19) \qquad (\hat{\theta}^n(t))_{t\geq 0} \text{ is a bounded and continuous martingale}$$

with increasing process

$$(2.20) \qquad \left\langle \hat{\theta}^n(t) \right\rangle_{t\geq 0} = \left(\int_0^t \mathrm{d}s \, \frac{1}{|\mathbb{G}_n|\,(1 + K)} \sum_{i\in\mathbb{G}_n} g(x_i(s))) \right)_{t\geq 0}.$$

Time scale. The goal is to identify β_n such that ($\mathcal{L}$ denotes probability law, convergence is always in the weak topology)

$$(2.21) \qquad \lim_{n\to\infty} \mathcal{L}\left[\mathcal{E}^n(s\beta_n)\right] = \lim_{n\to\infty} \mathcal{L}\left[\mu^n(s\beta_n)\right] = \mathcal{L}\left[\nu_{\Theta(s)}\right], \qquad s > 0,$$

and

$$(2.22) \qquad \lim_{n\to\infty} \mathcal{L}\left[\left(\hat{\theta}^n(s\beta_n)\right)_{s>0}\right] = \mathcal{L}\left[(\Theta(s))_{s>0}\right],$$

where along the way we also need to identify $(\Theta(s))_{s\geq 0}$ as a *Markov process* on $[0,1]$ (via an appropriate martingale problem or an associated SDE), and show that $\Theta(0) = \theta$ and $\Theta(\infty) \in \{0,1\}$. Accordingly, we want to show that

$$(2.23) \qquad \lim_{n\to\infty} \mathcal{L}\left[\mathcal{E}^n(\tilde{\beta}_n)\right] = \lim_{n\to\infty} \mathcal{L}\left[\mu^n(\tilde{\beta}_n)\right] = \nu_\theta$$

for any time scale $\tilde{\beta}_n \ll \beta_n$, and

$$(2.24) \qquad \lim_{n\to\infty} \mathcal{L}\left[\mathcal{E}^n(\tilde{\beta}_n)\right] = \lim_{n\to\infty} \mathcal{L}\left[\mu^n(\tilde{\beta}_n)\right] = \theta\,\delta_{(1,1)^{\mathbb{G}}} + (1-\theta)\,\delta_{(0,0)^{\mathbb{G}}}$$

for any time scale $\tilde{\beta}_n \gg \beta_n$. In the above scenario, the local configuration is controlled by the macroscopic variable in (2.17), in the sense that it approaches the corresponding equilibrium of the *infinite* system locally. The macroscopic variable itself follows an *autonomous* diffusion process, and *conditional* on $\hat{\theta}^n(s\beta_n)$ being close to some $\theta' \in [0,1]$ for some $s > 0$, the system converges in distribution to $\nu_{\theta'}$.

Scaling limit. In the model without seed-bank [CGS95], $(\Theta(s))_{s\geq 0}$ turned out to be the diffusion

$$(2.25) \qquad d\Theta(s) = \sqrt{(\mathcal{F}g)(\Theta(s)}\,dw(s), \qquad \Theta(0) = \theta,$$

with diffusion function $\mathcal{F}g$ given by

$$(2.26) \qquad (\mathcal{F}g)(\theta) = \mathbb{E}_{\nu_\theta}[g(x_0)] = \int_E g(x_0)\,\nu_\theta(dx), \qquad \theta \in [0,1],$$

and that if $g \in \mathcal{G}$, then also $\mathcal{F}g \in \mathcal{G}$. For the special case where $g = dg_{\mathrm{FW}}$, $d \in (0,\infty)$, it turned out that $\mathcal{F}g = d^* g_{\mathrm{FW}}$ with $d^* = d/(d + \hat{G}(0,0))$, where $\hat{G}(0,0)$ is the Green function at the origin of the symmetrised migration kernel $\hat{a}(\cdot,\cdot)$. In the presence of the seed-bank the role of $\hat{G}(0,0)$ is taken over by the quantity

$$(2.27) \qquad \begin{aligned} \hat{B}(0,0) \quad = \quad &\text{the mean total joint activity time of two independent Markov processes} \\ &\text{with transition kernel } b^{(1)}(\cdot,\cdot) \text{ in (A.1) both starting from } (0, A). \end{aligned}$$

The key observation is that $\hat{B}(0,0) < \infty$ if and only if the system is in the *coexistence regime*. In Appendix A we will see that $\mathcal{F}$ is non-trivial *only* in the coexistence regime, and that throughout this regime $\mathcal{F}g \in \mathcal{G}$ for all $g \in \mathcal{G}$. (In the clustering regime $\mathcal{F}g \equiv 0$ for all $g \in \mathcal{G}$, because the equilibrium lives on the states $x^n = y^n \equiv 1$ and $x^n = y^n \equiv 0$.)

Results. Abbreviate

$$(2.28) \qquad \kappa = (1 + K)^2.$$

Recall (1.24). For $\mu \in \mathcal{T}^{\mathrm{erg}}$, define

$$(2.29) \qquad \theta_\mu = \mathbb{E}_\mu\left[\frac{x_0 + Ky_0}{1 + K}\right]$$

and, for $\theta \in [0,1]$,

$$(2.30) \qquad \mathcal{T}_\theta^{\mathrm{erg}} = \left\{\mu \in \mathcal{T}^{\mathrm{erg}} \colon \theta_\mu = \theta\right\}.$$

Theorem 2.6. [Finite-systems scheme: Model 1] *Suppose that Assumptions 1.1, 2.1 and 2.3 are in force. Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$, and that $\mu_n(0)$ is given by (2.9). Then, in the coexistence regime, the following are true:*

(a) *(Convergence on macroscopic time scale) For every $s > 0$, (2.21)–(2.24) hold with time scale*

$$(2.31) \qquad \beta_n = \kappa|\mathbb{G}_n|$$

and $(\Theta(s))_{s\geq 0}$ the diffusion on $[0,1]$ given by

$$(2.32) \qquad d\Theta(s) = \sqrt{(\mathcal{F}g)(\Theta(s))}\,dw(s), \qquad \Theta(0) = \theta,$$

with $\mathcal{F}g$ given by (recall (1.51))

$$(2.33) \qquad (\mathcal{F}g)(\theta) = E_{\nu_\theta}[g(x_0)] = \int_E g(x_0)\,\nu_\theta(dx, dy), \qquad \theta \in [0,1].$$

(b) *(Fisher-Wright diffusion) If $g = dg_{FW}$, $d \in (0,\infty)$, then*

$$(2.34) \qquad \mathcal{F}g = d^* g_{FW}$$

with $d^* = d/(1 + d\hat{B}(0,0))$.

Thus, we find the same behaviour as in the system without seed-bank, except that the macroscopic time scale β_n runs slower by a factor $\kappa = (1 + K)^2$. One factor $1 + K$ arises from the time that is lost in the seed-bank, while the other factor $1 + K$ arises from the fact that the seed-bank brings in no volatility during the time that is lost (see (2.20), and Section 4 for further details). Fig. 5 draws a typical realisation of $(\Theta(s))_{s\geq 0}$.

Remark 2.7. [Clustering regime not universal] Theorem 2.6 *only holds in the coexistence regime*, because it captures on what time scale the finite system feels the boundary and begins to cluster. To derive an analogue of Theorem 2.6 in the clustering regime we would need to compare the different modes of clustering, which would be more difficult (see [CG91] for systems without seed-bank on $\mathbb{G} = \mathbb{Z}^2$). The outcome would *not be as universal* and would depend on the *degree* of recurrence of the migration kernel.
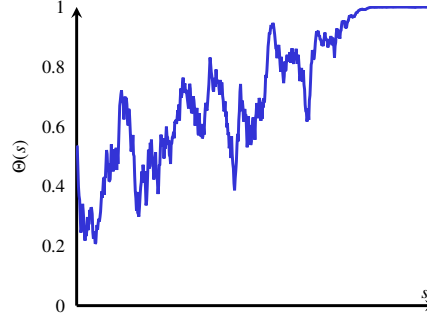□

Figure 5: Picture of the diffusion $(\Theta(s))_{s\geq 0}$. The boundary set $\{0, 1\}$ is hit in finite time when 0 and 1 are accessible, which is the case for instance when $\mathcal{F}g$ is a multiple of the Fisher-Wright diffusion function. The blue curve is the path of the diffusion.

Comments. The scaling of time by $|\mathbb{G}_n|$ is natural when we think of the dual. On $\mathbb{G}_n$, two random walks on average need $|\mathbb{G}_n|$ moves until they meet. Indeed, the coexistence regime corresponds to transient migration (recall (1.41)), for which the mixing time is $\psi_n = o(|\mathbb{G}_n|)$ (recall Remarks 2.4 and 2.5(1)). Hence, at time $s\beta_n = s\kappa|\mathbb{G}_n|$ the two random walks are more or less uniformly distributed on $\mathbb{G}_n$, and have probability $|\mathbb{G}_n|^{-1}$ to be at the same site. If we consider two Markov processes with transition kernel given by (A.1), then we get similar behaviour and the *slow-down factor* $1/\kappa$ in (2.28) is the *fraction of time that in the dual two lineages are jointly active*. This is natural because only active lineages in the dual can move and coalesce (i.e., forward in time only active individuals can move and exchange genetic information).

The reason why $\mathcal{F}g$ appears in (2.32) is that, given the value of the macroscopic variable $\hat{\theta}^n$, the $|\mathbb{G}_n|$ constituent components equilibrate on a time scale that is fast with respect to the time scale β_n on which $\hat{\theta}^n$ fluctuates. Consequently, the volatility of $\hat{\theta}^n$ is close to the expectation of the volatility of the constituent components in the quasi-equilibrium $\nu_{\hat{\theta}^n}$, as expressed by (2.33) (see Fig. 6).

We may think of $\mathcal{F}$ as a *renormalisation map* acting on the class $\mathcal{G}$ of diffusion functions defined in (1.34). According to (2.33), $\mathcal{F}$ is a *non-linear integral transform*, with the non-linearity arising from the fact that ν_θ depends on g (apart from e, K). In general no explicit formula is available for $\mathcal{F}$. However, the result in (2.34) says that if g is a multiple of the Fisher-Wright diffusion, then so is $\mathcal{F}g$, in which case $\{0, 1\}$ is accessible at both ends.
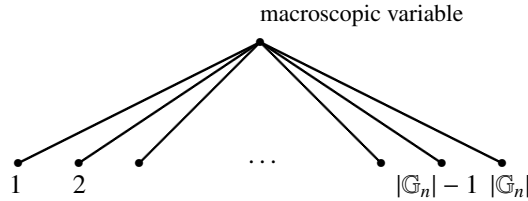
Figure 6: Pictorial representation of the averaging procedure behind the renormalisation map. The volatility of the empirical average of the $|\mathbb{G}_n|$ components converges as $n \to \infty$ to the average of the volatility of the single components with respect to their limiting equilibrium distribution.

Remark 2.8. [Hierarchical mean-field limit] In [GHO22a] we consider the case where the geographic space $\mathbb{G}$ is the hierarchical group Ω_N. In the so-called *hierarchical mean-field limit* $N \to \infty$ we are able to analyse $\mathcal{F}$ in detail. Without seed-bank the *mean-field finite-systems scheme* on Ω_N was established in [DG93], [DGV95]. □

Trapping time. Because $(\Theta(s))_{s\geq 0}$ is a bounded martingale, $\lim_{s\to\infty} \Theta(s)$ exists $\mathbb{P}$-a.s. Since $\mathcal{F}g \in \mathcal{G}$, the limit lies in $\{0, 1\}$. Let

(2.35) $\mathfrak{T} = \inf\{s \geq 0\colon \Theta(s) \in \{0, 1\}\}.$

If $\{0, 1\}$ is *accessible* at both ends, i.e.,

$$(2.36) \qquad \int_{[0,1]} dx \, \frac{x(1-x)}{(\mathcal{F}g)(x)} < \infty,$$

then $\mathbb{P}(\mathfrak{T} < \infty) = 1$ (see Fig. 5). We will show that the hitting time of the traps in the finite system

$$(2.37) \qquad H_n = \inf \left\{ t \geq 0 \colon \ (x^n(t), y^n(t)) \in \{(0,0)^{\mathbb{G}_n}, (1,1)^{\mathbb{G}_n}\} \right\}$$

(compare with (1.35)) satisfies the following.

Theorem 2.9. [Scaling of hitting time: Model 1] *Under the conditions in Theorem 2.6, for all diffusion functions $g \in \mathcal{G}$ satisfying (2.36),*

$$(2.38) \qquad \lim_{n \to \infty} \mathcal{L}[H_n/\beta_n] = \mathcal{L}[\mathfrak{T}].$$

Note that Theorem 2.9 does *not* follow directly from the path convergence in (2.22). The relation in (2.38) implies that at time $t = [1 + o(1)] \mathfrak{T}\beta_n$ the finite system *clusters*. If, on the other hand, $\{0, 1\}$ is *not accessible* at both ends, then $\mathfrak{T} = \infty$ and the finite system only approximately clusters on time scale β_n.

2.3 Model 2: Scaling

Truncation. In Model 2 the seed-bank is truncated from $\mathbb{N}_0$ to $\{0, \ldots, M_n\}$ with

$$(2.39) \qquad \lim_{n \to \infty} M_n = \infty.$$

Different choices for M_n make sense in biological applications, since they capture *how many seeds can accumulate in a single colony*. The truncated state space is $E_n = E_{M_n,n} = ([0,1] \times [0,1]^{M_n+1})^{\mathbb{G}_n}$, the truncated system

$$(2.40) \qquad (x^{M_n,n}(t), y^{M_n,n}(t)) = (x_i(t), (y_{i,m}(t))_{0 \leq m \leq M_n})_{i \in \mathbb{G}_n}$$

evolves as

$$(2.41) \qquad dx_i(t) \ = \ \sum_{j \in \mathbb{G}_n} a^n(i, j) \, [x_j(t) - x_i(t)] \, dt + \sqrt{g(x_i(t))} \, dw_i(t)$$

$$+ \sum_{m=0}^{M_n} K_m e_m \, [y_{i,m}(t) - x_i(t)] \, dt,$$

$$(2.42) \qquad dy_{i,m}(t) \ = \ e_m \, [x_i(t) - y_{i,m}(t)] \, dt, \qquad 0 \leq m \leq M_n,$$

where $a^n(i, j)$ is defined in (2.2).

Truncated macroscopic variable. The analogue of (2.17) is the *macroscopic variable*

$$(2.43) \qquad \hat{\theta}^{M_n,n}(t) = \frac{1}{|\mathbb{G}_n|} \sum_{i \in \mathbb{G}_n} \frac{x_i(t) + \sum_{m=0}^{M_n} K_m y_{i,m}(t)}{1 + \sum_{m=0}^{M_n} K_m},$$

the analogue of (2.15) is the *empirical measure*

$$(2.44) \qquad \mathcal{E}^{M_n,n}(t) = \frac{1}{|\mathbb{G}_n|} \sum_{i \in \mathbb{G}_n} \delta_{\tau_i(x^{M_n,n}(t), y^{M_n,n}(t))},$$

and we write $\mu^{M_n,n}(t)$ to denote the law of the truncated system at time t.

In order to understand the behaviour of the macroscopic variable $\hat{\theta}^{M_n,n}(t)$ defined in (2.43), we must keep track of the variance of the average of the active components, and look for a time scale in which this variance remains positive in the limit as $n \to \infty$. By (2.41) and (2.42), we find that $\hat{\theta}^{M,n}(t)$ evolves according to the SDE

$$(2.45) \qquad d\hat{\theta}^{M_n,n}(t) = \frac{1}{|\mathbb{G}_n| \, (1 + \sum_{m=0}^{M_n} K_m)} \sum_{i \in \mathbb{G}_n} \sqrt{g(x_i(t))} \, dw_i(t),$$

where the migration cancels out because of the averaging over the geographic space, and the exchange with the seed-bank cancels out because of the averaging over the seed-bank. Thus, in particular,

(2.46) $(\hat{\theta}^{M_n,n}(t))_{t\geq 0}$ is a bounded and continuous martingale

with increasing process

(2.47) $\left\langle \hat{\theta}^{M_n,n}(t) \right\rangle_{t\geq 0} = \left(\int_0^t \mathrm{d}s \, \frac{1}{|\mathbb{G}_n|(1 + \sum_{m=0}^{M_n} K_m)} \sum_{i\in\mathbb{G}_n} g(x_i(s))) \right)_{t\geq 0} .$

Time scale. To identify the proper time scale, we speed up time. Using the scaling properties $W(a^2 t) \cong aW(t)$ and $\sqrt{a}\,W(t) + \sqrt{b}\,W'(t) \cong \sqrt{a+b}\,W''(t)$, and putting

(2.48) $\beta_n = \kappa_{M_n} |\mathbb{G}_n|, \qquad \kappa_{M_n} = \left(1 + \sum_{m=0}^{M_n} K_m \right)^2 ,$

we see that (2.45) becomes

(2.49) $\mathrm{d}\hat{\theta}^{M_n,n}(s\beta_n) = \sqrt{\frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} g(x_i(s\beta_n))} \, \mathrm{d}w_i(s).$

Thus, important quantities for the evolution of (2.43) are the *active macroscopic variable*

(2.50) $\hat{\theta}_x^{M_n,n}(t) = \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} x_i(t)$

and the *active empirical measure*

(2.51) $\mathcal{E}_x^{M_n,n}(t) = \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} \delta_{\tau_i(x^{M_n,n})}.$

Recall (2.25)–(2.26). We might expect that again

(2.52) $\lim_{n\to\infty} \mathcal{L}\left[(\hat{\theta}^{M_n,n}(s\beta_n))_{s\geq 0} \right] = \mathcal{L}[(\Theta(s))_{s\geq 0}]$

and

(2.53) $\lim_{n\to\infty} \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} g(x_i(s\beta_n)) = (\mathcal{F}g)(\Theta(s)).$

Moreover, we might expect that again

(2.54) $\lim_{n\to\infty} \mathcal{L}\left[(\mu^{M_n,n}(s\beta_n))_{s\geq 0} \right] = \mathcal{L}[(\nu_{\Theta(s)})_{s\geq 0}],$

i.e., at time $s\beta_n$ the law of the finite system converges as $n \to \infty$ to a *random mixture* of equilibria for the infinite system, selected through the random variable $\Theta(s)$.

We will see in Sections 2.4–2.5 that the above heuristics is correct for $\rho < \infty$, and also for $\rho = \infty$ when the growth of M_n is slow, but fails for $\rho = \infty$ when the growth of M_n is fast.

2.4   Model 2: $\rho < \infty$

Results. For $\rho < \infty$ the behaviour is qualitatively similar to what we saw in Model 1. Let $\mu(0) \in \mathcal{T}^{\mathrm{erg}}$. For $\mu \in \mathcal{T}^{\mathrm{erg}}$, define

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

and, for $\theta \in [0, 1]$,

(2.56) $\mathcal{T}_\theta^{\mathrm{erg}} = \left\{ \mu \in \mathcal{T}^{\mathrm{erg}} : \theta_\mu = \theta \right\}.$

Theorem 2.10. [Finite-systems scheme: Model 2 with $\rho < \infty$] *Suppose that Assumptions 1.1, 2.1 and 2.3 are in force. Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$, and that $\mu_n(0)$ is given by (2.9). Put*

$$(2.57) \qquad \beta_n = \kappa |\mathbb{G}_n|, \qquad \kappa = (1 + \rho)^2.$$

Then, in the coexistence regime, for any M_n satisfying (2.39), the same formulas as in (2.21)–(2.24) and (2.32)–(2.33) apply, with $(0,0)^\mathbb{G}$ and $(1,1)^\mathbb{G}$ replaced by $(0, 0^{\mathbb{N}_0})^\mathbb{G}$ and $(1, 1^{\mathbb{N}_0})^\mathbb{G}$ in (2.24).

Comments. Theorem 2.10 shows that for $\rho < \infty$ the behaviour is similar as in Theorem 2.6. The macroscopic time scale β_n runs slower by a factor $\kappa = (1 + \rho)^2$, with one factor $1 + \rho$ coming from the time that is lost in the seed-bank and the other factor $1 + \rho$ coming from the fact that the seed-bank brings in no volatility during the time that is lost (again see Section 4 for further details). The scaling of time by $|\mathbb{G}_n|$ is again natural, because for $\rho < \infty$ the coexistence regime again corresponds to transient migration (recall (1.41)), for which the mixing time of the migration is $o(|\mathbb{G}_n|)$ (recall Remarks 2.4–2.5). The slow-down factor $1/\kappa$ in (2.28) is again the fraction of time the two lineages are jointly active.

Trapping time. Theorem 2.9 carries over verbatim.

2.5 Model 2: $\rho = \infty$

For $\rho = \infty$, the situation is much more delicate. It turns out that there are *two regimes*, which we define in Section 2.5.1 and refer to as *slow growing* seed-bank and *fast growing* seed-bank. Here, slow and fast refer to the time scale on which two active lineages in the dual coalesce, which depends on $|\mathbb{G}_n|$ and on γ, the exponent of the tail of the wake-up time defined in (1.40). We identify the finite-systems scheme in these regimes in Sections 2.5.2, respectively, 2.5.3, and find that they exhibit completely different behaviour. For slow growing seed-bank there is a *gradual approach* towards complete clustering as for $\rho < \infty$, while for fast growing seed-bank there is a *bursty approach* towards complete clustering, with *random switches* between partially clustered states, driven by dormant lineages that gradually wake up on time scales that exceed the time scale on which two active lineages in the dual coalesce.

For $\mu \in \mathcal{T}^{\mathrm{erg},\bullet}$ (recall (1.24)), define

$$(2.58) \qquad \theta_\mu = \lim_{M \to \infty} \mathbb{E}_\mu \left[\frac{x_0 + \sum_{m=0}^M K_m \, y_{0,m}}{1 + \sum_{m=0}^M K_m} \right].$$

and, for $\theta \in [0, 1]$,

$$(2.59) \qquad \mathcal{T}_\theta^{\mathrm{erg},\bullet} = \left\{ \mu \in \mathcal{T}^{\mathrm{erg},\bullet} : \theta_\mu = \theta \right\}.$$

The limit in (2.58) exists because $\mu(0)$ is colour regular (recall (1.25)).

2.5.1 Two growth regimes

Recall (1.40)–(1.39). We introduce *two time scales*:

$$(2.60) \qquad \beta_n^* = M_n^\beta, \qquad \beta_n^{**} = \begin{cases} |\mathbb{G}_n|^{1/(2\gamma-1)}, & \gamma \in (\frac{1}{2}, 1], \\ \mathrm{e}^{|\mathbb{G}_n|}, & \gamma = \frac{1}{2}, \\ \infty, & \gamma \in (0, \frac{1}{2}). \end{cases}$$

The interpretation of these time scales is as follows (recall (1.17)–(1.18)):

- β_n^* is the time scale on which a single dormant lineage in the dual starting from the deepest seed-bank (colour M_n) becomes active.

- β_n^{**} is the time scale on which two active lineages in the dual coalesce (on the active layer $\mathbb{G}_n$).

Note that β_n^* only depends on the size of the seed-bank, while β_n^{**} only depends on the size of the geographic space and the exponent of the tail of the wake-up time.

The two regimes of interest are

(2.61)
$$\begin{aligned} &\text{(I)} \quad \beta_n^* \ll \beta_n^{**} \quad \text{('slow growth')},\\ &\text{(II)} \quad \beta_n^{**} \ll \beta_n^* \quad \text{('fast growth')}. \end{aligned}$$

(For the critical value $\gamma = \frac{1}{2}$, the two regimes are actually $\log \beta_n^* \ll \log \beta_n^{**}$, respectively, $\log \beta_n^{**} \ll \log \beta_n^*$. Note that for $\gamma \in (0, \frac{1}{2})$ only regime (I) is possible.) In regime (I) the system behaves as if $\rho < \infty$ (see Section 5). In regime (II) the system essentially behaves like a *hybrid* system in which the geographic space is truncated but the seed-bank is not (see Section 6). In the crossover regime

(2.62)
$$\beta_n^* \asymp \beta_n^{**},$$

which we will not consider, the behaviour is more involved.

We will see that the two regimes exhibit different behaviour. In other words, the limits $|\mathbb{G}_n| \to \infty$ and $M_n \to \infty$ *cannot be interchanged*. Note that both in regime (I) and regime (II) the time scale $\kappa_{M_n}|\mathbb{G}_n|$ falls between the time scales β_n^* and β_n^{**}.

2.5.2 Slow growing seed-bank

Results. In regime (I) the behaviour is similar as for $\rho < \infty$ except for a change of time scale.

Theorem 2.11. [Finite-systems scheme: Model 2 with $\rho = \infty$ in regime (I)] *Suppose that Assumptions 1.1, 1.3, 2.1 and 2.3 are in force. Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$, and that $\mu_n(0)$ is given by (2.9). Put*

(2.63)
$$\beta_n = \kappa_{M_n}|\mathbb{G}_n|.$$

Then, in the coexistence regime, the same formulas as in (2.21)–(2.24) and (2.32)–(2.33) apply.

Comments. Theorem 2.11 shows that for $\rho = \infty$ the macroscopic time scale β_n needs to be speeded up in order to see fluctuations of the macroscopic variable. The fraction of time that two lineages in the dual are jointly active scales like $1/\kappa_{M_n}$. Hence the time scale must be multiplied by κ_{M_n}, but otherwise the behaviour in regime (I) is similar as for $\rho < \infty$. Note that $\kappa_{M_n} \sim [A/(1-\alpha)]^2 M_n^{2(1-\alpha)}$ for $\alpha \in (-\infty, 1)$ and $\kappa_{M_n} \sim [A \log M_n]^2$ for $\alpha = 1$ (recall (1.39)).

Trapping time. Theorem 2.9 again carries over verbatim.

2.5.3 Fast growing seed-bank

Results. In regime (II) the behaviour is different than in regime (I), both in terms of time scale and scaling limit. In the following we analyse what happens on time scale $\bar{\beta}_n$ in *three different ranges*:

(2.64)
$$\begin{aligned} &\text{(1)} \quad 1 \ll \bar{\beta}_n \ll \beta_n^{**},\\ &\text{(2)} \quad \beta_n^{**} \ll \bar{\beta}_n \ll \beta_n^*,\\ &\text{(3)} \quad \bar{\beta}_n \gg \beta_n^*. \end{aligned}$$

In order to state our result, we need to fix a *depth* L_n until which the seed-bank is being monitored. The proper choice turns out to be any L_n satisfying

(2.65)
$$1 \ll L_n \ll \bar{\beta}_n^{1/\beta},$$

where $\beta > 0$ is the exponent in (1.39).

Recall from (1.5) that $\mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$. Abbreviate $\mathbb{S}_n = \mathbb{G}_n \times \{A, (D_m)_{0 \le m \le M_n}\}$ and $\mathbb{S}_n^{L_n} = \mathbb{G}_n \times \{A, (D_m)_{0 \le m \le L_n}\}$, $0 \le L_n \le M_n$. Given a sequence of laws $(\mu_n)_{n \in \mathbb{N}}$ in $\mathcal{P}([0,1]^{\mathbb{S}})$ and a law $\mu \in \mathcal{P}([0,1]^{\mathbb{S}})$, we say that $\lim_{n \to \infty}^{L_n} \mu_n = \mu$ when

(2.66)
$$\lim_{n \to \infty} \sup_{\substack{A_n \subset [0,1]^{\mathbb{S}_n^{L_n}} \\ \text{measurable}}} \left| \int_{A_n} f \mathrm{d}\mu_n - \int_{A_n} f \mathrm{d}\mu \right| = 0 \qquad \forall\, f \in C_b([0,1]^{\mathbb{S}}; \mathbb{R}),$$

which we refer to as *weak convergence to depths* $(L_n)_{n \in \mathbb{N}_0}$.

Theorem 2.12. [Finite-systems scheme: Model 2 with $\rho = \infty$ in regime (II)] *Suppose that Assumptions 1.1, 1.3, 2.1 and 2.3 are in force, and that $\psi_n = o((\beta_n^{**})^\gamma)$. Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$, and that $\mu_n(0)$ is given by (2.9). Then, in the coexistence regime, for the three ranges in (2.64):*

(1) *Equilibrium:* (2.23) *holds with β_n replaced by $\bar\beta_n$ and $\lim_{n\to\infty}$ replaced by $\lim_{n\to\infty}^{L_n}$.*

(2) *Partial clustering:*

$$
\begin{aligned}
(2.67) \quad & \lim_{n\to\infty} \mathcal{L}\left[\left(\hat\theta^{M_n,n}(\bar\beta_n), \hat\theta_x^{M_n,n}(\bar\beta_n)\right)\right] = \delta_\theta \otimes \delta_{\tilde\Upsilon}, \\
& \lim_{n\to\infty}^{L_n} \mathcal{L}\left[\left(\mathcal{E}^{M_n,n}(\bar\beta_n)\right)\right] = \lim_{n\to\infty}^{L_n} \mathcal{L}\left[\left(\mu^{M_n,n}(\bar\beta_n)\right)\right] = \tilde\Upsilon\, \delta_{(1,1^{\mathbb{N}_0})^{\mathbb{G}}} + [1 - \tilde\Upsilon]\, \delta_{(0,0^{\mathbb{N}_0})^{\mathbb{G}}},
\end{aligned}
$$

where $\tilde\Upsilon \in \{0, 1\}$ is a random variable with law

$$(2.68) \qquad \mathcal{L}[\tilde\Upsilon] = \theta\, \delta_1 + (1 - \theta)\, \delta_0.$$

Moreover, for any two time scales $\beta_n^{**} \ll \bar\beta_n \ll \tilde\beta_n \ll \beta_n^*$, $\tilde\Upsilon$ and $\tilde\Upsilon$ are independent.

(3) *Complete clustering:* (2.24) *holds with β_n replaced by $\bar\beta_n$ and $\lim_{n\to\infty}$ replaced by $\lim_{n\to\infty}^{M_n}$.*

Comments. The behaviour in regime (II) is strikingly different from that in regime (I).

(1) Before time scale β_n^{**} *local convergence to equilibrium occurs*, as in the infinite system, up to seed-bank depth $L_n = o(\bar\beta_n^{1/\beta})$.

(1-2) On time scale β_n^{**}, *partial clustering* sets in that is *global in the geographic space but local in the seed-bank up to depth $L_n = o((\beta_n^{**})^{1/\beta})$*, i.e., the active population and the seed-banks up to depth L_n gradually move towards one of the clustered states in which they are either all 1 or all 0. (The latter corresponds to the geographic space partitioning into two parts, where all the active individuals and all the dormant individuals in the seed-banks up to depth L_n stem from a single ancestor.) There is no movement yet towards the equilibria of the finite system, i.e., towards the completely clustered states, because the deeper seed-banks have not yet made themselves felt.

(2) After time scale β_n^{**} but before time scale β_n^*, deeper seed-banks come into play, and *partial clustering* occurs up to depth $L_n = o((\bar\beta_n)^{1/\beta})$. Since the initial mean is θ, and the mean is preserved under the evolution, the value that is taken in the partial clustering is the *random variable $\tilde\Upsilon \in \{0, 1\}$* with mean θ, i.e., $\tilde\Upsilon = 0$ with probability θ and $\tilde\Upsilon = 1$ with probability $1 - \theta$. The deeper seed-banks make the active population and the seed-banks up to depth L_n undergo *random switches* between the two partially clustered states: on a larger time scale $\tilde\beta_n$ the density is equal to a freshly sampled random variable $\tilde\Upsilon \in \{0, 1\}$ with mean θ, i.e., *the deeper seed-banks overrule the shallower seed-banks*.

(2-3) Only when time scale β_n^* is reached do the deepest seed-banks come into play, partial clustering occurs up to depth $L_n = o((\beta_n^*)^{1/\beta})$, with $(\beta_n^*)^{1/\beta} = M_n$, after which *complete clustering* sets in that is *global in the geographic space and in the seed-bank*, i.e., the active population and all the seed-banks gradually move towards one of the clustered states in which they are either all 1 or all 0, and complete clustering occurs up to depth $L_n = M_n$. (The latter corresponds to the geographic space partitioning into two parts in which all the individuals, both active and dormant, stem from a single ancestor.)

(3) After time scale β_n^* *complete fixation* has been achieved.

Note that in regime (II) the time scale $\kappa_{M_n}|\mathbb{G}_n|$ of regime (I) is no longer relevant. In fact, *the role of β_n is taken over by β_n^*, the time scale at which complete clustering sets in*. The fact that the macroscopic variable tends to 0 or 1 on time scale β_n^* and not on time scale $\kappa_{M_n}|\mathbb{G}_n|$ is due to the partial clustering, which makes the right-hand side of (2.45) small.

It remains open to identify what exactly happens in the crossover regimes $\bar\beta_n \asymp \beta_n^{**}$ and $\bar\beta_n \asymp \beta_n^*$. We expect that on time scale β_n^{**} the macroscopic variable associated with the active population and the seed-banks up to depth $L_n = o((\beta_n^{**})^{1/\beta})$ move towards partial fixation according to a *jump process*, i.e., it follows a piecewise constant path that ends in 0 or 1. We expect that on time scale β_n^* the macroscopic variable associated with the full population, i.e., the active population and all the seed-banks up to depth M_n, moves towards fixation according to a jump process in the same manner, modulo a constant multiple of time. In both instances the behaviour is *different* from the diffusion in Fig. 5 with diffusion function $\mathcal{F}g$. Nonetheless, we expect that this diffusion still plays a role in the background. See Appendix D for speculations.

Explanation. In regime (II) we may pretend that the active and dormant time lapses of a lineage in the dual behave in the same way as when $M_n \equiv \infty$. Indeed, if τ_n denotes the wake-up time of a lineage in the dual for the finite system after it has become dormant, then (compare with (1.21))

$$(2.69) \qquad \mathbb{P}(\tau_n > t) = \frac{1}{\chi} \sum_{m=0}^{M_n} K_m e_m \, \mathrm{e}^{-e_m t}, \qquad t \geq 0.$$

Inserting (1.39), approximating the sum by an integral and passing to the new variable $x = e_m t$, we find that

$$(2.70) \qquad \mathbb{P}(\tau_n > t) \sim \frac{A}{\beta \chi} B^{1-\gamma} t^{-\gamma} \int_{Bt/\beta_n^*}^{Bt} \mathrm{d}x \, x^{\gamma-1} \mathrm{e}^{-x}, \qquad t \to \infty.$$

Thus, we see that the asymptotics found for the infinite system in (1.40) prevails in regime (II) because $\bar{\beta}_n/\beta_n^* \downarrow 0$ as $n \to \infty$.

On $\mathbb{G}_n$, the mean total joint activity time up to time T for two lineages in the dual *without coalescence*, starting anywhere in $\mathbb{G}_n$ and being both active, equals the *hazard*

$$(2.71) \qquad H_n(T) \asymp \int_1^T t^{-2(1-\gamma)} |\mathbb{G}_n|^{-1} \, \mathrm{d}t, \qquad T \to \infty.$$

Indeed, by (2.70), the probability for each lineage to be active at time t falls off like $t^{-(1-\gamma)}$, while the probability for the two lineages to be at the same site is $|\mathbb{G}_n|^{-1}$, provided these lineages are well mixed on $\mathbb{G}_n$ at times of order T. From (2.71) we see that

$$(2.72) \qquad H_n(T) \asymp |\mathbb{G}_n|^{-1} \times \begin{cases} T^{2\gamma-1}, & \gamma \in (\frac{1}{2}, 1], \\ \log T, & \gamma = \frac{1}{2}, \\ C, & \gamma \in (0, \frac{1}{2}), \end{cases} \qquad T \to \infty,$$

and so $H_n(T)$ starts to diverge on time scale β_n^{**} (recall (2.60)). Up to time T, each lineage takes $\asymp T^\gamma$ migration steps, and so we require that $\psi_n = o((\beta_n^{**})^\gamma)$ to get proper mixing on time scale β_n^{**}, which explains the assumption on ψ_n in Theorem 2.12. This assumption is met for all three examples in Remark 2.4 when $\gamma \in (\frac{1}{2}, 1]$. For the first example we have $|\mathbb{G}_n| = n^d$ and $\psi_n \asymp n^2$, and so we require that $2 < d\gamma/(2\gamma - 1)$, which is precisely the condition for coexistence in (1.43). For the second example we have $|\mathbb{G}_n| = n$ and $\psi_n \asymp n^\delta$, and so we require that $\delta < \gamma/(2\gamma - 1)$, which is precisely the condition for coexistence in (1.44). For the third example we have $|\mathbb{G}_n| = N^n$ and $\psi_n \asymp (N/c)^{n-1}$, and so we require that $N/c < N^{\gamma_N/(2\gamma_N-1)}$, which is precisely the condition for coexistence in (1.45). The assumption is trivially met for $\gamma \in (0, \frac{1}{2})$ (recall (2.60)) and $\gamma = \frac{1}{2}$ (recall Remark 2.5(1)).

Remark 2.13. [Crossover for the hazard] The integral in (1.42) arises as the mean total joint activity time on $\mathbb{G}$

$$(2.73) \qquad I_{\hat{a},\gamma} \asymp \int_1^\infty t^{-2(1-\gamma)} \hat{a}_{t^\gamma}(0,0) \, \mathrm{d}t.$$

In the coexistence regime we have $I_{\hat{a},\gamma} < \infty$, while $H_n(T)$ in (2.71) starts to diverges on scale $T = \beta_n^{**}$. The reason why this is possible is related to the observation made in Remark 2.5(2), namely, because $|\mathbb{G}_n|^{-1} \asymp \hat{a}_{\psi_n}(0,0)$ and $\psi_n = o(T^\gamma)$, we have $|\mathbb{G}_n|^{-1} \gg \hat{a}_{T^\gamma}(0,0)$. $\qquad \square$

2.6 Open problems

We close by listing a few open problems.

(A) How can we refine Theorem 2.12? In particular, what happens at times of order β_n^{**} and β_n^*? On the shorter time scale β_n^{**} the system undergoes clustering in the geographic space but not in the seed-bank, while on the longer time scale β_n^* the system undergoes clustering in the seed-bank (and hence complete clustering). In Section 6 we will argue that in regime (II), unlike in regime (I), clustering in the geographic space is *not diffusive*, but rather proceeds in *random bursts*, i.e., on time scale β_n^{**} the active macroscopic variable follows a random jump process that lives on $(0, 1)$ and ends at 0 or 1.

(B) What is the analogue of Theorem 2.9 in regime (II)? We expect the trapping time to be of order β_n^*, but not to be controlled by a diffusion because of the random bursts.

(C) What happens in the crossover regime in (2.62)? We expect diffusive scaling of the macroscopic variable, but driven by a diffusion function different from $\mathcal{F}g$.

3   Preparation: Preservation and convergence

The strategy of the proof of the theorems stated in Section 2 is to make the following simple ideas rigorous, where for notational convenience we focus on Model 1. Let $S = (S(t))_{t \geq 0}$ denote the semigroup associated with the Markov process $Z = (Z(t))_{t \geq 0}$ on $\mathbb{G}$ defined in (1.7), and let S^n, Z^n be their finite-system counterparts on $\mathbb{G}_n$. Suppose that we have identified the time scale β_n on which the estimator process $\hat{\theta}^n = (\hat{\theta}^n(s\beta_n))_{s>0}$ is tight in $D((0, \infty), [0, 1])$. Then, by the Markov property of Z^n, we can write

$$(3.1) \qquad \mathcal{L}[Z^n(s\beta_n)] = \int_{[0,1]} \mathbb{P}(\hat{\theta}^n(s\beta_n - L_n) \in d\theta') \, \mu_{\theta'}^n S^n(L_n), \qquad s > 0,$$

where L_n is chosen such that $\lim_{n \to \infty} L_n = \infty$ and $\lim_{n \to \infty} L_n/\beta_n = 0$, and $\mu_{\theta'}^n = \mathcal{L}[Z^n(s\beta_n - L_n) \mid \hat{\theta}^n = \theta']$ (regular version of the conditional probability). From the tightness assumption on $\hat{\theta}^n$ we know that the law in (3.1) barely changes during the time interval $[s\beta_n - L_n, s\beta_n]$. Suppose that, along some subsequence, $\hat{\theta}^n$ converges to some limit process $\Theta = (\Theta(s))_{s>0}$ in $D((0, \infty), [0, 1])$. Then we should have that

$$(3.2) \qquad \mu_{\theta'}^n S^n(L_n) \approx \mu_{\theta'}^{n,\uparrow}(L_n) \implies \nu_{\theta'}, \qquad n \to \infty,$$

where $\uparrow$ denotes the extension to a law on $\mathbb{G}$ instead of $\mathbb{G}_n$. Combining (3.1)–(3.2), we get

$$(3.3) \qquad \mathcal{L}[Z^n(s\beta_n)] \implies \int_{[0,1]} \mathbb{P}(\Theta(s) \in d\theta') \, \nu_{\theta'}, \qquad n \to \infty, \qquad s > 0.$$

At the same time, we should be able to identify Θ by taking the increasing process of $\hat{\theta}^n$ (recall (2.20)),

$$(3.4) \qquad \langle \hat{\theta}^n(s) \rangle = \frac{1}{1+K} \int_0^s du \, \frac{1}{|\mathbb{G}_n|} \sum_{i \in \mathbb{G}_n} g(x_i(u\beta_n)), \qquad s > 0,$$

to conclude that, by the law of large numbers, the volatility function of Θ is given by

$$(3.5) \qquad \theta \mapsto \frac{1}{1+K} \mathbb{E}_{\nu_\theta}[g(x_0(\cdot))].$$

The question is how to transform the *heuristic argument* in (3.1)–(3.5) into a *rigorous argument* (one of the difficulties being that limits are being exchanged at several places). It was shown in [CG94b] that, subject to a collection of assumptions listed as (A1)–(A10) in Appendix A, the above steps can indeed be made rigorous. Our task will therefore be to verify these assumptions. In the present section we collect some properties that are needed to complete this task. These properties concern approximations of the infinite system by finite systems on time scales of order 1, ergodic theorems for the infinite system, and regularity properties of the equilibria for the infinite system as a function of underlying parameters. The verification of the assumptions is carried out in Sections 4 and 5 for $\rho < \infty$ and $\rho = \infty$, respectively.

Before we proceed we need the following important observations concerning the set of initial laws. In Section 3.1 we introduce a *class of initial laws* $\mathcal{R}_\theta$ parametrised by the density $\theta \in [0, 1]$, for which the system decorrelates in space over time intervals of length $o(\beta_n)$, where β_n is the *macroscopic time scale*, which must be properly chosen (see Appendix B, part *III*). In fact, the proper choices are:

$$(3.6) \qquad \beta_n = \begin{cases} \kappa|\mathbb{G}_n| & \text{Model 1} & \kappa = (1+K)^2, \\ \kappa|\mathbb{G}_n| & \text{Model 2 with } \rho < \infty & \kappa = (1+\rho)^2, \\ \kappa_{M_n}|\mathbb{G}_n| & \text{Model 2 with } \rho = \infty, \text{ regime (I)} & \kappa_{M_n} = (1 + \sum_{m=0}^{M_n} K_m)^2, \\ \beta_n^* & \text{Model 2 with } \rho = \infty, \text{ regime (II)}. \end{cases}$$

The initial laws chosen for Model 1 and Model 2 all fall in $\mathcal{R}_\theta$. In Section 3.2 we show that the evolved law of the system stays inside the class $\mathcal{R}_\theta$ over time intervals of length $o(\beta_n)$. In Section 3.3 we show that

the macroscopic variable converges to θ over time intervals of length $o(\beta_n)$. In Section 3.4 we use this fact to show that on time scale β_n the law of the system conditional on the macroscopic variable being θ' falls in the class $\mathcal{R}_{\theta'}$. We will see in Sections 4–5 that the latter is the key to the proof that the finite system locally converges to an equilibrium of the infinite system, parametrised by the instantaneous value of the macroscopic variable.

In Sections 3.2–3.4 we *exclude* Model 2 with $\rho = \infty$ in regime (II). In Section 3.5 we show that the results do carry over to this case as well, but *only* for time intervals of length $o(\beta_n^{**})$ rather than $o(\beta_n)$.

Throughout this section, $b^{(1),n}(\cdot, \cdot)$ and $b^{(2),n}(\cdot, \cdot)$ denote the Markov process transition kernels defined in (A.1) and (A.4) in Appendix A, which describe the motion of the *lineages* of the individuals in the spatial population with seed-bank. Depending on the model under consideration, we write

$$(3.7) \qquad \mathcal{R}_\theta = \begin{cases} \mathcal{R}_\theta^{(1)}, & \text{Model 1,} \\[6pt] \mathcal{R}_\theta^{(2)}, & \text{Model 2 with } \rho < \infty, \\[6pt] \mathcal{R}_\theta^{(2),\bullet}, & \text{Model 2 with } \rho = \infty. \end{cases}$$

3.1 Classes of initial laws

For each model a specific class of initial laws on $\mathbb{G}$ and their restrictions to $\mathbb{G}_n$ defined by (2.9) play an important role. The following are adaptations of Definition 1.6, where we write $\mathbb{S}_n = \mathbb{G}_n \times \{A, D\}$ in Model 1 and $\mathbb{S}_n = \mathbb{G}_n \times \{A, (D_m)_{0 \le m \le M_n}\}$ in Model 2, and we recall the definition of z_u in (1.53).

Definition 3.1. [Model 1: Class of initial laws] *For $\theta \in [0, 1]$, let*

$$(3.8) \qquad \mathcal{R}_\theta^{(1)} = \{\mu \in \mathcal{T}^{\mathrm{erg}} : (*)_\theta \text{ holds}\},$$

with $(*)_\theta$ the requirement that, for every sequence of times $(t_n)_{n \in \mathbb{N}}$ satisfying $\lim_{n \to \infty} t_n = \infty$ and $\lim_{n \to \infty} t_n/\beta_n = 0$,

$$(3.9) \qquad (*)_\theta \qquad \begin{array}{l} (1) \ \forall\, u_1 \in \mathbb{S}_n : \\ \displaystyle \lim_{n \to \infty} \sum_{u_2 \in \mathbb{S}_n} b_{t_n}^{(1),n}(u_1, u_2)\, \mathbb{E}_{\Phi_n\mu}[z_{u_2}] = \theta, \\[10pt] (2) \ \forall\, u_1, u_2 \in \mathbb{S}_n : \\ \displaystyle \lim_{n \to \infty} \sum_{u_3, u_4 \in \mathbb{S}_n} b_{t_n}^{(1),n}(u_1, u_3)\, b_{t_n}^{(1),n}(u_2, u_4)\, \mathbb{E}_{\Phi_n\mu}[z_{u_3} z_{u_4}] = \theta^2, \end{array}$$

where $\Phi_n\mu$ is the restriction of μ defined in (2.9)–(2.10), $b_{t_n}^{(1),n}(\cdot, \cdot)$ is the time-t_n transition kernel of the Markov process $b^{(1)}(\cdot, \cdot)$ in (A.1) restricted to $\mathbb{G}_n$.

Definition 3.2. [Model 2: Class of initial laws] *For $\theta \in [0, 1]$, let*

$$(3.10) \qquad \begin{array}{ll} \rho < \infty : & \mathcal{R}_\theta^{(2)} = \{\mu \in \mathcal{T}^{\mathrm{erg}} : (*)_\theta \text{ holds}\}, \\[10pt] \rho = \infty : & \mathcal{R}_\theta^{(2),\bullet} = \{\mu \in \mathcal{T}^{\mathrm{erg}} : (*)_\theta \text{ holds and } \mu \text{ is colour regular}\}, \end{array}$$

where in $(*)_\theta$ in (3.9) $b_{t_n}^{(1),n}(\cdot, \cdot)$ is replaced by $b_{t_n}^{(2),n}(\cdot, \cdot)$, the time-$t_n$ transition kernel of the Markov process $b^{(2)}(\cdot, \cdot)$ in (A.4) restricted to $\mathbb{S}_n$.

The two properties in $(*)_\theta$ in (3.9), which we refer to as *Liggett conditions* (see [GHO22b, Remark 2.15]), say that over time intervals of length $o(\beta_n)$ the following are true: (1) the average of the component seen at time t_n by the Markov process with transition kernel $b^{(1),n}(\cdot, \cdot)$ and $b^{(2),n}(\cdot, \cdot)$ converges to the initial density θ defined in (2.55); (2) the covariance of the components seen at time t_n by two independent Markov processes with transition kernel $b^{(1),n}(\cdot, \cdot)$ and $b^{(2),n}(\cdot, \cdot)$ converges to zero. Consequently, the component seen by such a Markov process converges to θ in L_2.

The following lemma shows that all laws in $\mathcal{R}_\theta^{(1)}$, $\mathcal{R}_\theta^{(2)}$ and $\mathcal{R}_\theta^{(2),\bullet}$ are invariant and ergodic under translations, have density θ, and are colour regular (recall (2.30), (2.56) and (2.59)).

Lemma 3.3. [Ergodicity] *(a) For every $\theta \in [0, 1]$, $\mathcal{R}_\theta^{(1)} \subset \mathcal{T}_\theta^{\mathrm{erg}}$.*
(b) For every $\theta \in [0, 1]$, $\mathcal{R}_\theta^{(2)} \subset \mathcal{T}_\theta^{\mathrm{erg}}$, respectively, $\mathcal{R}_\theta^{(2),\bullet} \subset \mathcal{T}_\theta^{\mathrm{erg},\bullet}$.

Proof. The claim will be immediate from the proof of Lemma 3.4 below. For the finite-systems scheme, we will work with $\mathcal{T}^\bullet = \cup_{\theta \in [0,1]} \mathcal{T}_\theta^\bullet$ as the set of initial laws (recall (1.25)–(1.26)) and use that $\theta_\mu^\bullet = \lim_{m \to \infty} \mathbb{E}_\mu[y_{0,m}] = \theta$ for $\mu \in \mathcal{T}_\theta^\bullet$. $\qquad\square$

Note that (recall (1.24) and (1.26))

$$(3.11) \qquad \mathcal{T}^{\mathrm{erg}} = \cup_{\theta \in [0,1]} \mathcal{T}_\theta^{\mathrm{erg}}, \qquad \mathcal{T}^{\mathrm{erg},\bullet} = \cup_{\theta \in [0,1]} \mathcal{T}_\theta^{\mathrm{erg},\bullet}.$$

3.2 Preservation of the classes

The classes introduced in Definitions 3.1–3.2 are preserved over time intervals of length $o(\beta_n)$. In the following lemmas $\mathcal{R}_\theta$ stands for either $\mathcal{R}_\theta^{(1)}$, $\mathcal{R}_\theta^{(2)}$ and $\mathcal{R}_\theta^{(2),\bullet}$. Recall the restriction operator Φ_n defined in (2.10).

Lemma 3.4. [Preservation properties] *For any $\theta \in [0,1]$ and any sequence of times $(t_n)_{n \in \mathbb{N}}$ satisfying $\lim_{n \to \infty} t_n = \infty$ and $\lim_{n \to \infty} t_n/\beta_n = 0$, the following holds for, respectively, Model 1, Model 2 with $\rho < \infty$, and Model 2 with $\rho = \infty$ in regime (I):*

(a) *If $\mu(0) = \mu \in \mathcal{R}_\theta$, then $\mu(t_n) \in \mathcal{R}_\theta$.*

(b) *If $\mu(0) = \mu \in \mathcal{R}_\theta$, then $\nu \in \mathcal{R}_\theta$ for any weak limit point ν of $(\Phi_n \mu(t_n))_{n \in \mathbb{N}}$.*

Proof. (a) The claim is obvious after we replace t_n by $2t_n$ in $(*)_\theta$.

(b) The proof proceeds in 3 Steps, each based on a lemma. Before we start we note that the macroscopic variable $\hat{\theta}^n$ in Model 1 (defined in (2.17)) satisfies

$$(3.12) \qquad \mathbb{E}_{\Phi_n \mu(t_n)}[\hat{\theta}^n] = \mathbb{E}_{\Phi_n \mu}[\hat{\theta}^n] = \theta,$$

because $(\hat{\theta}^n(t))_{t \geq 0}$ is a martingale under the law $\Phi_n \mu$ (recall (2.46)). Therefore, also for any weak limit point ν of $(\Phi_n \mu(t_n))_{n \in \mathbb{N}}$,

$$(3.13) \qquad \mathbb{E}_{\Phi_n \nu(t_n)}[\hat{\theta}^n] = \mathbb{E}_{\Phi_n \nu}[\hat{\theta}^n] = \theta.$$

The same observation applies to the macroscopic variable $\hat{\theta}^{M_n,n}$ in Model 2 (defined in (2.43)).

Step 1. The following lemma says that under the law $\Phi_n \nu$ all components have mean θ. Via (3.13) this implies that property $(*)_\theta(1)$ holds and that ν is colour regular.

Lemma 3.5. [Constant means] *Let $\mu(0) = \mu \in \mathcal{R}_\theta$. Let $(t_n)_{n \in \mathbb{N}}$ satisfy $\lim_{n \to \infty} t_n = \infty$ and $\lim_{n \to \infty} t_n/\beta_n = 0$. Let ν be any weak limit point of $(\Phi_n \mu(t_n))_{n \in \mathbb{N}}$. Then $\mathbb{E}_{\Phi_n \nu}[z_{(i,R_i)}] = \theta$ for all $n \in \mathbb{N}$, all $i \in \mathbb{G}_n$, and all $R_i \in \{A, D\}$ or $\{A, (D_m)_{0 \leq m \leq M_n}\}$.*

Proof. Since ν is a weak limit point of $(\Phi_n \mu(t_n))_{n \in \mathbb{N}}$, there exists $(t_{n_k})_{k \in \mathbb{N}}$ satisfying $\lim_{k \to \infty} t_{n_k} = \infty$ and $\lim_{k \to \infty} t_{n_k}/\beta_{n_k} = 0$ such that $\nu = \lim_{k \to \infty} \Phi_{n_k} \mu(t_{n_k})$. Hence, for Model 1, for all $u_1 \in \mathbb{S}_n$,

$$(3.14) \qquad \mathbb{E}_\nu[z_{u_1}] = \lim_{k \to \infty} \mathbb{E}_{\Phi_{n_k} \mu(t_{n_k})}[z_{u_1}] = \lim_{k \to \infty} \sum_{u_2 \in \mathbb{S}_{n_k}} b_{t_{n_k}}^{(1),n}(u_1, u_2) \, \mathbb{E}_{\Phi_{n_k} \mu}[z_{u_2}] = \theta,$$

where the last equality holds because $\mu \in \mathcal{R}_\theta^{(1)}$. The same holds for Model 2 with $b_{t_{n_k}}^{(1),n}(\cdot, \cdot)$ replaced by $b_{t_{n_k}}^{(2),n}(\cdot, \cdot)$. Since $\mathbb{E}_\nu[z_{u_1}] = \theta$, it follows that $\mathbb{E}_{\Phi_n \nu}[z_{u_1}] = \theta$ for all $n \in \mathbb{N}$ (use (2.9) and the fact that ν is invariant under translations). $\qquad\square$

Step 2. To settle property $(*)_\theta(2)$, we follow the covariance computations for the infinite system carried out in [GHO22b, Sections 5–6]. Together with property $(*)_\theta(1)$, the following settles property $(*)_\theta(2)$.

Lemma 3.6. [Decaying correlations]
Let $\mu(0) = \mu \in \mathcal{R}_\theta$. Let $(t_n)_{n \in \mathbb{N}}$ satisfy $\lim_{n \to \infty} t_n = \infty$ and $\lim_{n \to \infty} t_n/\beta_n = 0$. Let ν be any weak limit point of $(\Phi_n \mu(t_n))_{n \in \mathbb{N}}$. Then

$$(3.15) \qquad \lim_{|i-j| \to \infty} \lim_{n \to \infty} \left| \mathrm{Cov}_{\Phi_n \nu}(z_{(i,R_i)}, z_{(j,R_j)}) \right| = 0 \quad \textit{uniformly in } R_i, R_j.$$

Proof. For Model 1, we use Itô's formula for the first and second moment to write (see [GHO22b, Lemma 5.1])

$$\text{Cov}_{\Phi_n \mu(t_n)}(z_{u_1}, z_{u_2}) = \int_0^{t_n} ds \sum_{k \in \mathbb{G}_n} b^{(1),n}_{t_n-s}(u_1, (k, A)) \, b^{(1),n}_{t_n-s}(u_2, (k, A)) \, \mathbb{E}_{\Phi_n \mu}[g(x_k(s))]$$

(3.16)

$$+ \sum_{u_3, u_4 \in \mathbb{S}_n} b^{(1),n}_{t_n}(u_1, u_3) \, b^{(1),n}_{t_n}(u_2, u_4) \, \text{Cov}_{\Phi_n \mu}(z_{u_3}, z_{u_4}).$$

The same expression holds for Model 2, with $b^{(1),n}$ replaced by $b^{(2),n}$ (see [GHO22b, Lemma 6.1]). In what follows we focus on Model 2. Model 1 has no truncation of the seed-bank: $M_n = 0$ for all $n \in \mathbb{N}$.

1. The second term in (3.16) tends to zero as $n \to \infty$ by (3.9), because $\nu \in \mathcal{R}^{(2)}_\theta$ by Lemma 3.4(b). To estimate the first term, which we abbreviate by

$$(3.17) \qquad I_{t_n}(u_1, u_2) = \int_0^{t_n} ds \sum_{k \in \mathbb{G}_n} b^{(2),n}_{t_n-s}(u_1, (k, A)) \, b^{(2),n}_{t_n-s}(u_2, (k, A)) \, \mathbb{E}_{\Phi_n \mu}[g(x_k(s))],$$

we proceed as follows (in which we no longer need that $\nu \in \mathcal{R}^{(2)}_\theta$). Let $\tau^\uparrow$ denote the first time the two Markov processes are *jointly active*, which is a random variable whose law depends on u_1, u_2. Because the building up of joint activity time can only start after time τ^*, we have

$$(3.18) \qquad I_{t_n}(u_1, u_2) \leq I_{t_n}((I, A), (J, A))$$

where $I, J \in \mathbb{G}_n$ are the random locations of the two Markov processess at time $\tau^\uparrow$. For $s \geq 0$, let $\mathcal{E}(t)$ and $\mathcal{E}'(t)$ be the events that the respective random walks are active at time t, and let

$$(3.19) \qquad T(t) = \int_0^t ds \, 1_{\mathcal{E}(s)}, \qquad T'(t) = \int_0^t ds \, 1_{\mathcal{E}'(s)},$$

be their total activity time up to time t. Write $\mathbb{P}_n$ to denote the law of $(\mathcal{E}(s), \mathcal{E}'(s))_{s \geq 0}$ given that the two random walks start in the active state. Note that this process is independent of the migration of the Markov processes. Estimate (henceforth $\|g\|$ denotes the supremum norm of g)

$$I_{t_n}((i, A), (j, A)) \leq \|g\| \int_0^{t_n} ds \sum_{k \in \mathbb{G}_n} b^{(2),n}_{t_n-s}((i, A), (k, A)) \, b^{(2),n}_{t_n-s}((j, A), (k, A))$$

(3.20)

$$= \|g\| \int_0^{t_n} ds \sum_{k \in \mathbb{G}_n} \mathbb{E}_n \left[a^n_{T(s)}(i, k) \, 1_{\mathcal{E}(s)} \, a^n_{T'(s)}(j, k) \, 1_{\mathcal{E}'(s)} \right]$$

$$= \|g\| \int_0^{t_n} ds \, \mathbb{E}_n \left[\hat{a}^n_{T(s)+T'(s)}(i - j, 0) \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right].$$

We rewrite and split the integral in the right-hand side into three parts:

$$J^1_{t_n} = |\mathbb{G}_n|^{-1} \int_0^{t_n} ds \, \mathbb{E}_n \left[1_{\{T(s)+T'(s) > \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right],$$

$$(3.21) \qquad J^2_{t_n}(i, j) = |\mathbb{G}_n|^{-1} \int_0^{t_n} ds \, \mathbb{E}_n \left[\left(|\mathbb{G}_n| \hat{a}^n_{T(s)+T'(s)}(i - j, 0) - 1 \right) 1_{\{T(s)+T'(s) > \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right],$$

$$J^3_{t_n}(i, j) = \int_0^{t_n} ds \, \mathbb{E}_n \left[\hat{a}^n_{T(s)+T'(s)}(i - j, 0) \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right],$$

where ψ_n is the mixing time defined in (2.3).

2. First we look at $J^3_{t_n}(i, j)$. Define the event

$$(3.22) \qquad \mathcal{A}_n = \{\text{until time } t_n \text{ neither of the two Markov processes visits a seed-bank with colour } m > M_n\}.$$

Since $\mathbb{P}_n(\mathcal{A}_n) = 1$, we have

$$(3.23) \qquad \begin{aligned} &\mathbb{E}_n \left[\hat{a}^n_{T(s)+T'(s)}(i - j, 0) \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right] \\ &\qquad = \mathbb{E}_n \left[\hat{a}^n_{T(s)+T'(s)}(i - j, 0) \, 1_{\mathcal{E}(s)} \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}'(s)} \, 1_{\mathcal{A}_n} \right]. \end{aligned}$$

On the event $\mathcal{A}_n$, the law of $\{\mathcal{E}(s), \mathcal{E}'(u)\}_{u \in [0,t_n]}$ is the same under $\mathbb{P}_n$ as under $\mathbb{P}$, where the former refers to the truncated system and the latter refers to the non-truncated system. Hence

$$
(3.24) \qquad
\begin{aligned}
&\mathbb{E}_n \left[\hat{a}^n_{T(s)+T'(s)}(i-j, 0) \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \, 1_{\mathcal{A}_n} \right] \\
&= \mathbb{E} \left[\hat{a}^n_{T(s)+T'(s)}(i-j, 0) \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \, 1_{\mathcal{A}_n} \right].
\end{aligned}
$$

On the event $\{T(s) + T'(s) \leq \psi_n\}$, we can estimate

$$
(3.25) \qquad
\begin{aligned}
&\mathbb{E} \left[\hat{a}^n_{T(s)+T'(s)}(i-j, 0) \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \, 1_{\mathcal{A}_n} \right] \\
&\leq C \, \mathbb{E} \left[\hat{a}_{T(s)+T'(s)}(i-j, 0) \, 1_{\{T(s)+T'(s) \leq \psi_n\}} \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right] \leq C \, \mathbb{E} \left[\hat{a}_{T(s)+T'(s)}(0, 0) \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right],
\end{aligned}
$$

where the first inequality uses (2.7) and the second inequality uses that $\hat{a}_t(k, 0) \leq \hat{a}_t(0, 0)$ for all $k \in \mathbb{G}$ and $t \geq 0$. The right-hand side integrated over $s \in [0, \infty)$ is the mean total joint activity time for two Markov processes on $\mathbb{G}$ both starting from $(0, A)$, which is finite if and only if the system is in the *coexistence regime* (see [GHO22b, Section 2.5]). Since $\lim_{|k| \to \infty} \hat{a}_t(k, 0) = 0$ uniformly in $t \geq 0$, it follows from the third line of (3.21) via (3.23)–(3.25) and dominated convergence that

$$
(3.26) \qquad \lim_{|i-j| \to \infty} J^3_{t_n}(i, j) = 0 \text{ uniformly in } (t_n)_{n \in \mathbb{N}}.
$$

Hence, to show that $J^3_{t_n}(I, J)$ gives a vanishing contribution to (3.18) as $|i - j| \to \infty$, we must show that the latter implies that $|I - J| \to \infty$ with $\mathbb{P}_n$-probability tending to 1 as $n \to \infty$. But this again trivially follows from the fact that $\lim_{|k| \to \infty} a_t(0, k) = 0$ uniformly in $t \geq 0$.

3. Next we look at $J^1_{t_n}$ and $J^2_{t_n}(i, j)$. Estimate, with the help of (2.3),

$$
(3.27) \qquad \forall \, i, j \in \mathbb{G}_n: \quad |J^2_{t_n}(i, j)| \leq C' J^1_{t_n}, \qquad C' = \sup_{n \in \mathbb{N}} \sup_{t \geq \psi_n} \sup_{k \in \mathbb{G}_n} \left| |\mathbb{G}_n| \, \hat{a}^n_t(0, k) - 1 \right| < \infty.
$$

Further estimate

$$
(3.28) \qquad J^1_{t_n} \leq \tilde{J}_{t_n} = |\mathbb{G}_n|^{-1} \int_0^{t_n} ds \, \mathbb{P}_n(\mathcal{E}(s)) \, \mathbb{P}_n(\mathcal{E}'(s)).
$$

We will show that

$$
(3.29) \qquad \lim_{n \to \infty} \tilde{J}_{t_n} = 0.
$$

Together with (3.26)–(3.28) this will imply (3.15).

4. For the proof of (3.29) we distinguish between Model 1 and 2.

Model 1. Since $\mathbb{P}_n = \mathbb{P}$ and $\lim_{s \to \infty} \mathbb{P}(\mathcal{E}(s)) = \lim_{s \to \infty} \mathbb{P}(\mathcal{E}'(s)) = (1 + K)^{-1}$, the integral in (3.28) scales like t_n/κ with $\kappa = (1 + K)^2$. Since $t_n = o(\beta_n)$ and $\beta_n = \kappa |\mathbb{G}_n|$, we get (3.29).

Model 2. For $\rho < \infty$, we have $\lim_{n \to \infty} \mathbb{P}_n = \mathbb{P}$, and $\lim_{s \to \infty} \mathbb{P}(\mathcal{E}(s)) = \lim_{s \to \infty} \mathbb{P}(\mathcal{E}'(s)) = (1 + \rho)^{-1}$, and so the argument for Model 1 carries over. For $\rho = \infty$, a bit more care is needed. In particular, we need to distinguish between regimes (I) and (II), which have different macroscopic time scales: $\beta_n = \kappa_{M_n} |\mathbb{G}_n|$, respectively, $\beta_n = \beta_n^*$.

As shown in Appendix C,

$$
(3.30) \qquad \mathbb{P}_n(\mathcal{E}(s)) \, \mathbb{P}_n(\mathcal{E}'(s)) \asymp
\begin{cases}
1/\kappa_{s^{1/\beta}}, & 0 \leq s \leq \beta_n^*, \\
1/\kappa_{M_n}, & \beta_n^* < s \leq t_n.
\end{cases}
$$

where we recall from (2.60) that $\beta_n^* = M_n^\beta$ and from (2.48) $\kappa_M = (1 + \sum_{m=0}^M K_m)^2$. Indeed, heuristically, at time $s = M^\beta$ the active component has communicated with the first M dormant components, and so joint activity occurs with probability $1/\kappa_M = 1/\kappa_{s^{1/\beta}}$. This holds all the way up to $s = \beta_n^*$, after which the active component has communicated with all M_n dormant components, and so joint activity occurs with probability $1/\kappa_{M_n}$. Now, inserting (3.30) into the right-hand side of (3.28), we get

$$
(3.31) \qquad \tilde{J}_{t_n} \asymp \tilde{J}^1_{t_n} + \tilde{J}^2_{t_n}
$$

with

$$(3.32) \qquad \tilde{J}^1_{t_n} = |\mathbb{G}_n|^{-1} \int_0^{\beta^*_n} \mathrm{d}s\,(1/\kappa_{s^{1/\beta}}), \qquad \tilde{J}^2_{t_n} = |\mathbb{G}_n|^{-1} \int_{\beta^*_n}^{t_n} \mathrm{d}s\,(1/\kappa_{M_n}).$$

▶ Regime (I). Since $\beta_n = \kappa_{M_n}|\mathbb{G}_n|$, we have $\tilde{J}^2_{t_n} \leq t_n/\beta_n = o(1)$ because $t_n = o(\beta_n)$, and so we need only worry about $\tilde{J}^1_{t_n}$. Recall from (1.39) that $\gamma = \frac{\alpha+\beta-1}{\beta}$ with $\alpha \leq 1 < \alpha + \beta$. Note that $\alpha = 1$ implies $\gamma = 1$.

- For $\gamma \in (\frac{1}{2}, 1]$ and $\alpha \in (-\infty, 1)$, we have $\tilde{J}^1_{t_n} \asymp |\mathbb{G}_n|^{-1}(\beta^*_n)^{2\gamma-1} = (\beta^*_n/\beta^{**}_n)^{2\gamma-1}$, where we use that $\kappa_M \asymp M^{2(1-\alpha)}$ as $M \to \infty$, $\gamma = \frac{\alpha+\beta-1}{\beta}$, and $\beta^{**}_n = |\mathbb{G}_n|^{1/(2\gamma-1)}$. Since $\beta^*_n = o(\beta^{**}_n)$, we get $\tilde{J}^1_{t_n} = o(1)$.

- For $\gamma = \alpha = 1$, we have $\tilde{J}^1_{t_n} \asymp |\mathbb{G}_n|^{-1}(\beta^*_n/\log\beta^*_n) = (\beta^*_n/\beta_n)(\kappa_{M_n}/\log\beta^*_n) \asymp \beta^*_n/\beta_n$, where we use that $\kappa_M \asymp \log M$ as $M \to \infty$ and $\beta^*_n = M^\beta_n$. Since $\beta^*_n = o(\beta_n)$, we again get $\tilde{J}^1_{t_n} = o(1)$.

- For $\gamma = \frac{1}{2}$ and $\alpha \in (-\infty, 1)$, we have $\tilde{J}^1_{t_n} \asymp |\mathbb{G}_n|^{-1}\log\beta^*_n$. Since $\log\beta^*_n = o(\log\beta^{**}_n) = o(|\mathbb{G}_n|)$, we again get $\tilde{J}^1_{t_n} = o(1)$.

- For $\gamma \in (0, \frac{1}{2})$, we use that $\lim_{n\to\infty}\mathbb{P}_n = \mathbb{P}$ and $\mathbb{P}(\mathcal{E}(s)) \asymp u^{-(1-\gamma)}$, $u \to \infty$, as shown in [GHO22b, Section 6.2]. Hence, by monotone convergence, $\tilde{J}_{t_n} \asymp C''|\mathbb{G}_n|^{-1} = o(1)$ with $C'' = \int_0^\infty \mathrm{d}s\,u^{-2(1-\gamma)} < \infty$.

▶ Regime (II). This regime is excluded because it exhibits different behaviour. See Section 3.5. □

Step 3. In [GHO22b, Lemma 6.10] we showed for Model 2 with $\rho = \infty$ that for the system on $\mathbb{G}$ in equilibrium the deep seed-banks are deterministic, i.e.,

$$(3.33) \qquad \lim_{m\to\infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) = 0.$$

The following lemma says that the same holds for the scaling limit of the system on $\mathbb{G}_n$.

Lemma 3.7. [Deterministic deep seed-banks: Model 2, $\rho = \infty$] *Let $\mu(0) = \mu \in \mathcal{R}_\theta$. Let $(t_n)_{n\in\mathbb{N}}$ satisfy $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} t_n/\beta_n = 0$. Let ν be any weak limit point of $(\Phi_n\mu(t_n))_{n\in\mathbb{N}}$. Then*

$$(3.34) \qquad \lim_{m\to\infty} \limsup_{n\to\infty} \mathrm{Var}_{\Phi_n\nu}(y_{0,m}) = 0.$$

Proof. (The proof does in fact not use that $\mu \in \mathcal{R}_\theta$.) Write (see [GHO22b, Lemma 6.1])

$$
\begin{aligned}
\mathrm{Var}_{\Phi_n\mu(t_n)}(y_{0,m}) &= \mathbb{E}_{\Phi_n\mu}\left[(y_{0,m}(t_n) - \mathbb{E}_{\Phi_n\mu}[y_{0,m}(t_n)])^2\right] \\
(3.35) \qquad &= \int_0^{t_n} \mathrm{d}s \sum_{k\in\mathbb{G}_n} b^{(2),n}_{t_n-s}((0, R_m), (k, A))\, b^{(2),n}_{t_n-s}((0, R_m), (k, A))\, \mathbb{E}_{\Phi_n\mu}[g(x_k(u))] \\
&\leq \|g\| \int_0^{t_n} \mathrm{d}s \sum_{k\in\mathbb{G}_n} b^{(2),n}_{t_n-s}((0, R_m), (k, A))\, b^{(2),n}_{t_n-s}((0, R_m), (k, A)),
\end{aligned}
$$

where the estimate is uniform in μ. The sum under the integral is the probability that two Markov processes, both starting from $(0, R_m)$ and moving according to $b^{(2),n}(\cdot, \cdot)$, at time $t_n - s$ are at the same site and both active. Define

$$(3.36) \qquad \tau = \{t \geq 0\colon RW(t) = RW'(t) = (i, A) \text{ for some } i \in \mathbb{G}_n\},$$

i.e., the first time when the two Markov processes are jointly active at the same site. Write $\mathbb{P}^{(2),n}_{(0,D_m),(0,D_m)}$ to denote the joint law of the two Markov processes. Then

$$
\begin{aligned}
\int_0^{t_n} &\mathrm{d}s \sum_{k\in\mathbb{G}_n} b^{(2),n}_{t_n-s}((0, R_m), (k, A))\, b^{(2),n}_{t_n-s}((0, R_m), (k, A)) \\
&= \int_0^{t_n} \mathrm{d}s\, \mathbb{E}^{(2),n}_{(0,D_m),(0,D_m)}\left[\sum_{k\in\mathbb{G}_n} \mathbb{1}_{\{RW(s)=RW'(s)=k\}}\, \mathbb{1}_{\mathcal{E}(s)}\, \mathbb{1}_{\mathcal{E}'(s)}\right] \\
(3.37) \qquad &= \mathbb{E}^{(2),n}_{(0,D_m),(0,D_m)}\left[\mathbb{1}_{\{\tau \leq t_n\}} \int_0^{t_n-\tau} \mathrm{d}s\, E^{(2),n}_{(0,A),(0,A)}\left[\sum_{k\in\mathbb{G}_n} \mathbb{1}_{\{RW(s)=RW'(s)=k\}}\, \mathbb{1}_{\mathcal{E}(s)}\, \mathbb{1}_{\mathcal{E}'(s)}\right]\right] \\
&\leq \mathbb{P}^{(2),n}_{(0,D_m),(0,D_m)}(\tau \leq t_n) \int_0^{t_n} \mathrm{d}s\, \mathbb{E}_n\left[\hat{a}^n_{T(s)+T'(s)}(0, 0)\, \mathbb{1}_{\mathcal{E}(s)}\, \mathbb{1}_{\mathcal{E}'(s)}\right],
\end{aligned}
$$

where for the second equality we use the strong Markov property at time τ, together with the fact that on the event $\{\tau > t_n\}$ the product of the indicators equals 0 for all $u \in [0, t_n]$, and for the inequality we recall (3.20). By (3.21), we have

$$(3.38) \qquad \int_0^{t_n} ds \, \mathbb{E}_n \left[\hat{a}^n_{T(s)+T'(s)}(0,0) \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right] = J^1_{t_n} + J^2_{t_n}(0,0) + J^3_{t_n}(0,0),$$

By (3.27)–(3.29), the first two terms tends to zero as $n \to \infty$. By (3.23)–(3.25),

$$(3.39) \qquad J^3_{t_n}(0,0) \le C \int_0^{t_n} ds \, \mathbb{E} \left[\hat{a}_{T(s)+T'(s)}(0,0) \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right],$$

which is finite in the coexistence regime, uniformly in $(t_n)_{n\in\mathbb{N}}$. Furthermore, recall (3.22) to estimate

$$(3.40) \qquad \mathbb{P}^{(2),n}_{(0,D_m),(0,D_m)} (\tau \le t_n) = \mathbb{P}^{(2),n}_{(0,D_m),(0,D_m)} (\tau \le t_n, \mathcal{A}_n) = \mathbb{P}^{(2)}_{(0,D_m),(0,D_m)} (\tau \le t_n, \mathcal{A}_n) \le \mathbb{P}^{(2)}_{(0,D_m),(0,D_m)} (\tau < \infty).$$

Hence, to get the claim in (3.34) it suffices to show that

$$(3.41) \qquad \lim_{m\to\infty} \mathbb{P}^{(2)}_{(0,D_m),(0,D_m)} (\tau < \infty) = 0.$$

But this was done in [GHO22b, Section 6.3, Step 3], the idea being that two Markov processes starting from deep seed-banks are far apart when they are jointly awake for the first time. □

Steps 1–3 complete the proof of Lemma 3.4. □

3.3 Law of large numbers for the macroscopic variable

The following lemma will play a crucial role in Sections 4–6.

Lemma 3.8. [L_2-convergence of the macroscopic variable] *For any $\theta \in [0,1]$, any $\mu(0) = \mu \in \mathcal{R}_\theta$ and any sequence of times $(t_n)_{n\in\mathbb{N}}$ satisfying $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} t_n/\beta_n = 0$,*

$$(3.42) \qquad \lim_{n\to\infty} \mathbb{E}_{\Phi_n\mu} \left[(\hat{\theta}^n(t_n) - \theta)^2 \right] = 0$$

in Model 1, and similarly for $\hat{\theta}^{M_n,n}$ in Model 2.

Proof. Recall (3.12). We again distinguish between Model 1 and 2.

Model 1. Use (2.17) to write

$$(3.43) \qquad \mathbb{E}_{\Phi_n\mu(t_n)} \left[(\hat{\theta}^n - \theta)^2 \right] = \frac{1}{(1+K)^2 |\mathbb{G}_n|^2} \sum_{u_1,u_2\in\mathbb{S}_n} K(u_1) \, K(u_2) \, \text{Cov}_{\Phi_n\mu(t_n)}(z_{u_1}, z_{u_2})$$

with $K(u) = 1$ for $u = (i, A)$ and $K(u) = K$ for $u = (i, D)$. Via (3.9) and (3.18) (with ν replaced by μ) it follows that

$$
\begin{aligned}
(3.44) \qquad \limsup_{n\to\infty} \mathbb{E}_{\Phi_n\mu(t_n)} \left[(\hat{\theta}^n - \theta)^2 \right] &= \limsup_{n\to\infty} \frac{1}{(1+K)^2 |\mathbb{G}_n|^2} \sum_{u_1,u_2\in\mathbb{S}_n} K(u_1) \, K(u_2) \, I_{t_n}(u_1, u_2) \\
&\le \limsup_{n\to\infty} \frac{1}{|\mathbb{G}_n|^2} \sum_{i,j\in\mathbb{G}_n} I_{t_n}((i,A),(j,A)).
\end{aligned}
$$

We have

$$
\begin{aligned}
(3.45) \qquad \frac{1}{|\mathbb{G}_n|^2} \sum_{i,j\in\mathbb{G}_n} I_{t_n}((i,A),(j,A)) &\le \|g\| \frac{1}{|\mathbb{G}_n|^2} \sum_{i,j\in\mathbb{G}_n} \int_0^{t_n} ds \, \mathbb{E} \left[\hat{a}^n_{T(s)+T'(s)}(i-j,0) \, 1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right] \\
&= \|g\| \frac{1}{|\mathbb{G}_n|} \int_0^{t_n} ds \, \mathbb{E} \left[1_{\mathcal{E}(s)} \, 1_{\mathcal{E}'(s)} \right] = \|g\| \tilde{J}_{t_n},
\end{aligned}
$$

where the inequality uses (3.20), the first equality uses that $\sum_{k\in\mathbb{G}_n} \hat{a}^n_t(k,0) = 1$ for all $t \ge 0$ and all $n \in \mathbb{N}$, and the second equality is (3.28). We already showed that $\tilde{J}_{t_n} = o(1)$, and so this settles (3.42).

Model 2. Use (2.43) to write, similarly as in (3.43),

$$
(3.46) \quad
\begin{aligned}
&\mathbb{E}_{\Phi_n \mu(t_n)} \left[(\hat{\theta}^{M_n,n} - \theta)^2 \right] \\
&= \frac{1}{(1 + \sum_{m=0}^{M_n} K_m)^2 |\mathbb{G}_n|^2} \sum_{u_1, u_2 \in \mathbb{S}_n} K(u_1) K(u_2) \operatorname{Cov}_{\Phi_n \mu(t_n)}(z_{u_1}, z_{u_2})
\end{aligned}
$$

with $K(s) = 1$ for $s = (i, A)$ and $K(s) = K_m$ for $s = (i, D_m)$ (recall (2.48)). Via (3.9) and (3.16)–(3.18) (with ν replaced by μ) it follows that, similarly as in (3.44),

$$
(3.47) \quad
\begin{aligned}
&\limsup_{n \to \infty} \mathbb{E}_{\Phi_n \mu(t_n)} \left[(\hat{\theta}^{M_n,n} - \theta)^2 \right] \\
&= \limsup_{n \to \infty} \frac{1}{(1 + \sum_{m=0}^{M_n} K_m)^2 |\mathbb{G}_n|^2} \sum_{u_1, u_2 \in \mathbb{S}_n} K(u_1) K(u_2) I_{t_n}(u_1, u_2) \\
&\leq \limsup_{n \to \infty} \frac{1}{|\mathbb{G}_n|^2} \sum_{i,j \in \mathbb{G}_n} I_{t_n}((i, A), (j, A)).
\end{aligned}
$$

The rest of the argument is the same. □

3.4 Convergence at macroscopic times conditional on the macroscopic variable

Recall (1.24), (1.26), (2.9), (2.12) and (2.30), (2.56) and (2.59). We begin with the observation that, for all $\theta \in [0, 1]$,

$$
(3.48) \quad
\begin{aligned}
\mathcal{T}_\theta^{\mathrm{erg}} &= \mathcal{R}_\theta^{(1)} && \text{(Model 1)}, \\
\mathcal{T}_\theta^{\mathrm{erg}} &= \mathcal{R}_\theta^{(2)} && \text{(Model 2 with } \rho < \infty), \\
\mathcal{T}_\theta^{\mathrm{erg}, \bullet} &= \mathcal{R}_\theta^{(2), \bullet} && \text{(Model 2 with } \rho = \infty).
\end{aligned}
$$

Indeed the inclusion $\subset$ was shown in [GHO22b, Lemma 5.7] and [GHO22b, Lemmas 6.6. and 6.9], while the inclusion $\supset$ was shown in Lemma 3.3.

For Model 2 with $\rho = \infty$ we need the topology of uniform weak convergence on the set $\mathcal{T}^{\bullet, *}$ defined in (1.28) (recall also Remark 1.2). We additionally define

$$
(3.49) \quad \mathcal{T}_\theta^{\mathrm{erg}, \blacksquare} = \left\{ \mu \in \mathcal{T}_\theta^{\mathrm{erg}, \bullet} : \lim_{m \to \infty} \operatorname{Var}_\mu(y_{0,m}) = 0 \right\},
$$

which is a subset of $\mathcal{T}^{\mathrm{erg}, \bullet, *}$.

Lemma 3.9. [Conditional convergence at macroscopic time] *Fix $\mu(0) = \mu \in \mathcal{T}$ (Model 1, Model 2 with $\rho < \infty$) and $\mu \in \mathcal{T}^\bullet$ (Model 2 with $\rho = \infty$). Fix $s > 0$, and let*

$$
(3.50) \quad \mu_n = (\Phi_n \mu)(s \beta_n).
$$

Then every weak limit point $\bar{\nu}$ of $(\mu_n)_{n \in \mathbb{N}}$ has the representation

$$
(3.51) \quad \bar{\nu} = \int_{[0,1]} Q(\mathrm{d}\theta) \, \bar{\nu}_\theta,
$$

where

$$
(3.52) \quad Q = \begin{cases} \lim_{n \to \infty} \mathcal{L}_{\mu_n}[\hat{\theta}^n] & \text{(Model 1)}, \\ \lim_{n \to \infty} \mathcal{L}_{\mu_n}[\hat{\theta}^{M_n,n}] & \text{(Model 2)}, \end{cases}
$$

is the associated limit law of the macroscopic variable, and

$$
(3.53) \quad \bar{\nu}_\theta = \begin{cases} \int_{\mathcal{T}_\theta^{\mathrm{erg}}} Q_\theta(\mathrm{d}\mu) \, \mu & \text{(Model 1)}, \\ \int_{\mathcal{T}_\theta^{\mathrm{erg}}} Q_\theta(\mathrm{d}\mu) \, \mu & \text{(Model 2 with } \rho < \infty), \\ \int_{\mathcal{T}_\theta^{\mathrm{erg}, \blacksquare}} Q_\theta(\mathrm{d}\mu) \, \mu & \text{(Model 2 with } \rho = \infty), \end{cases}
$$

for some Choquet measure Q_θ. Moreover, $\theta \mapsto \bar{\nu}_\theta$ is continuous on $\mathcal{T}$ in the weak topology, respectively, on $\mathcal{T}^\bullet$ in the uniform weak topology.

Proof. Fix $s > 0$, pick any $(t_n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} t_n = \infty$ and $\lim_{n \to \infty} t_n / \beta_n = 0$, and put

$$(3.54) \qquad \mu_n^- = (\Phi_n \mu)(s\beta_n - t_n).$$

Let $\bar{\nu}^-$ be any weak limit point of $(\mu_n^-)_{n \in \mathbb{N}}$. Let $Q^- = \lim_{n \to \infty} \mathcal{L}_{\mu_n^-}[\hat{\theta}^n]$ (Model 1), respectively, $Q^- = \lim_{n \to \infty} \mathcal{L}_{\mu_n^-}[\hat{\theta}^{M_n,n}]$ (Model 2) be the associated limit law of the macroscopic variable. Then

$$(3.55) \qquad \bar{\nu}^- = \int_{[0,1]} Q^-(\mathrm{d}\theta)\, \bar{\nu}_\theta^-,$$

where, by Choquet's theorem,

$$(3.56) \qquad \bar{\nu}_\theta^- = \begin{cases} \int_{\mathcal{T}_\theta^{\mathrm{erg}}} Q_\theta^-(\mathrm{d}\mu)\, \mu, & (\text{Model 1, Model 2 with } \rho < \infty), \\[2mm] \int_{\mathcal{T}_\theta^{\mathrm{erg},\bullet}} Q_\theta^-(\mathrm{d}\mu)\, \mu, & (\text{Model 2 with } \rho = \infty), \end{cases}$$

for some Choquet measure Q_θ^-.

Next we evolve the dynamics over time t_n to see what happens at time $s\beta_n$. To that end we consider the limit points $\bar{\nu}_\theta$ of $((\Phi_n \bar{\nu}_\theta^-) S_n(t_n))_{n \in \mathbb{N}}$, where $(S_n(s))_{s \geq 0}$ is the semigroup of the dynamics on $\mathbb{G}_n \times \{A, (D_m)_{0 \leq m \leq M_n}\}$. By (3.48) and Lemmas 3.5–3.7, these limit points are in $\mathcal{T}_\theta^{\mathrm{erg}}$, respectively, $\mathcal{T}_\theta^{\mathrm{erg},\bullet}$. Moreover, by Lemma 3.8, $\hat{\theta}^n$ and $\hat{\theta}^{M_n,n}$ have the same limit points under μ_n^- and μ_n. Hence (3.55) and (3.56) imply (3.51) and (3.53).

The continuity of $\theta \mapsto \bar{\nu}_\theta$ uses the same coupling argument that was used in [GHO22b] for the infinite system, and that was modified in Sections 3.1–3.2 to deal with the finite system. $\qquad \square$

The observation made in the proof of Lemma 3.3 guarantees that

$$(3.57) \qquad \mathcal{T}_\theta^{\mathrm{erg},\blacksquare} = \left\{ \mu \in \mathcal{T}_\theta^{\mathrm{erg},*} : \lim_{m \to \infty} y_{0,m} = \theta \ \mu\text{-a.s.} \right\},$$

i.e., in Model 2 with $\rho = \infty$ the deep seed-banks are deterministic in equilibrium.

In Sections 4–6 we will see that Lemma 3.9 is the key to showing that, for any $\theta \in [0,1]$, $\mu(0) = \mu \in \mathcal{T}$ and $s > 0$,

$$(3.58) \qquad \lim_{n \to \infty} \mathcal{L}[(\Phi_n \mu)(s\beta_n)(\cdot \mid \hat{\theta}^n = \theta)] = \nu_\theta(\cdot),$$

i.e., $\bar{\nu}_\theta = \nu_\theta$. The proof requires the use of an abstract scheme, which is outlined in Appendix B. The latter will allow us to identify the limit law of the macroscopic variable on time scale β_n. A similar statement holds for Model 2 conditional on $\hat{\theta}^{M_n,n} = \theta$, for $\mu(0) = \mu \in \mathcal{T}$, respectively, $\mu(0) = \mu \in \mathcal{T}^\bullet$.

3.5   Extension to fast growing seed-banks

For Model 2 with $\rho = \infty$ in regime (II), Lemmas 3.3–3.7 carry over provided we restrict to time intervals of length $o(\beta_n^{**})$ rather than $o(\beta_n^*)$ (recall that β_n^* is the macroscopic time scale for Model 2 with $\rho = \infty$ in regime (II)). Indeed, recall from (3.29) that

$$(3.59) \qquad \tilde{J}_{t_n} = |\mathbb{G}_n|^{-1} \int_0^{t_n} \mathrm{d}s\, (1/\kappa_{s^{1/\beta}}).$$

It suffices to check that $\lim_{n \to \infty} \tilde{J}_{t_n} = 0$ when $t_n = o(\beta_n^{**})$, because this allows us to carry through the estimates given in Part 4 of Step 2 in the proof of Lemma 3.4. For this we refer to the four bullets below (3.32).

The reason for the restriction is that *ergodicity breaks down on time scale* β_n^{**}. In Section 6 we will analyse Model 2 with $\rho = \infty$ in regime (II). Lemmas 3.8–3.9 will not be needed there (even though they are trivially true on times scales $o(\beta_n^{**})$).

4   Proofs: $\rho < \infty$

In this section we prove Theorems 2.6, 2.9 and 2.10. Section 4.1 focusses on Model 1, Section 4.2 on Model 2 with $\rho < \infty$.

4.1 Model 1

The proof of Theorem 2.6 will follow from the same argument as given below for Model 2 when $\rho < \infty$.

To prove Theorem 2.9 note that, by using the same Brownian motions for every $n \in \mathbb{N}$, we can realise weak convergence of the path in (2.14) as a.s. uniform convergence on a compact space via the Skorohod representation. Moreover, the increasing process of the macroscopic variable $\hat{\theta}^n$ in (2.43) is the increasing process of the active macroscopic variable $\hat{\theta}_x^n$ in (2.50) divided by $1 + K$, and both hit the traps 0 or 1 if and only if their increasing process hits 0 and remains 0. Since $(\hat{\theta}^n(t))_{t \geq 0}$ is a continuous martingale, it is a time-transformed Brownian motion, with the time transformation given by its increasing process. We conclude from the path convergence that the increasing processes converge to the increasing process of the $\mathcal{F}g$-diffusion. The latter has a derivative that converges to zero as the macroscopic tome s tends to infinity, and so the limit path becomes constant. For that reason we can conclude that on time scale β_n the hitting time of the traps by the path in (2.14) converges to the hitting time of the traps for the limit process, which is the $\mathcal{F}g$-diffusion.

4.2 Model 2: $\rho < \infty$

In this section we prove Theorem 2.10. The proof is built on an *abstract scheme* for deriving the finite-systems scheme, developed in [CG94b, Section 1] and [DGV95, Section 4] for general spatial systems and summarised in Appendix B. This abstract scheme has been applied to several classes of systems, but not yet to systems with seed-banks.

We will see that we can use this abstract scheme by *incorporating the seed-bank into the single-component state space*, namely, by considering the state space $I^{\mathbb{G}}$ with $I = [0, 1] \times [0, 1]^{\mathbb{N}_0}$, respectively, $I^{\mathbb{G}_n}$ with $I = [0, 1] \times [0, 1]^{M_n+1}$ (see Remark B.1). Along the way, various ingredients need to be specified, for which we use the lemmas derived in Section 3. In particular, we need to show that the macroscopic time scale chosen in (2.57), namely, $\beta_{M_n,n} = \kappa|\mathbb{G}_n|$ with $\kappa = (1 + \rho)^2$, emerges as the correct scaling time (which is related to the verification of Assumption (A10) in Corollary B.3 in Appendix B).

We will see that the abstract scheme allows for a *bootstrappig argument*. Namely, we first use tightness of the macroscopic variable process to establish that the finite system locally converges to the equilibrium of the infinite system with a density that is given by the limiting value of the macroscopic variable. Afterwards, we use this equilibrium to identify how to macroscopic variable process evolves on the macroscopic time scale.

The proof is organised into 5 Steps. In Steps 1–3 we first carry out the proof for $M_n \equiv M$ and $g = dg_{\mathrm{FW}}$, $d \in (0, \infty)$. In Steps 4–5 we let $M_n \to \infty$ and consider general $g \in \mathcal{G}$.

Step 1: Checking the assumptions.   We must check Assumptions (A1)–(A9) in Theorem B.2 and assumption (A10) in Corollary B.3. We proceed item per item, after first checking that our set-up fits into the abstract scheme.

• **Set-up.** Our geographic space is a countable Abelian group $\mathbb{G}$ endowed with the discrete topology, while our single-component state space $I = [0, 1] \times [0, 1]^{M+1}$ is a Polish space equipped with the product topology of $[0, 1]$. Our assumption that $\mathbb{G}$ is *profinite* yields a projective system $(\mathbb{G}_n)_{n \in \mathbb{N}}$ of finite groups endowed with the discrete topology. Our full state spaces $E, E_{M,n}$ are $I^{\mathbb{G}}$, $I_M^{\mathbb{G}_n}$. We can apply Theorem B.2, provided we choose the initial law appropriately, namely, from the class $\mathcal{R}_\theta^{(2)}$ introduced in Section 3.1, which is the domain of attraction of the equilibrium ν_θ and which, by Lemma 3.4, is preserved on time scale $o(\beta_{M,n})$.

(A1).   If $g(x) = dg_{\mathrm{FW}}$, $d \in (0, \infty)$, then the dual on $\mathbb{G}_n$ is the spatial coalescent with truncated transition kernel $a_n(\cdot, \cdot)$ for the random walk on $\mathbb{G}_n$. For this process it is straightforward to see that, as $n \to \infty$, the random walk on $\mathbb{G}_n$ converges to the random walk on $\mathbb{G}$ because $a_n(\cdot, \cdot)$ converges pointwise to $a(\cdot, \cdot)$. Therefore $b^n(\cdot, \cdot)$ converges to $b(\cdot, \cdot)$, and also the dual lineages converge. The seed-bank leads to different waiting times until the next jump occurs in the dual. Therefore, as $n \to \infty$, the spatial coalescent on $\mathbb{G}_n$ converges in law to the spatial coalescent on $\mathbb{G}$. Hence, by the duality relation and the fact that the moments are convergence determining, as $n \to \infty$ the forward process on $\mathbb{G}_n$ converges to the forward process on $\mathbb{G}$ in the sense of marginal distributions.

(A2).   Let

(4.1)

$$\mathcal{T} = \text{set of translation invariant probability measures on } I^{\mathbb{G}} \text{ under the group action on } \mathbb{G},$$

$$\mathcal{T}_n = \text{set of translation invariant probability measures on } I^{\mathbb{G}_n} \text{ under the group action on } \mathbb{G}_n.$$

In our context, $\mathcal{T}$ is defined in (1.24), while $\mathcal{T}_n$ is the image of $\mathcal{T}$ under the projection given by (2.9). Since the evolution mechanism is translation invariant, $\mathcal{T}_n$ and $\mathcal{T}$ are preserved under the semigroup of the evolution, and so (A2) holds.

(A3).   The Polish space $\mathcal{J}$ is the set $[0, 1]$ labelling the density, and $\mathcal{T}_\theta^{\mathrm{erg}}$ is the set of translation invariant probability measures that are *ergodic* under translation with density θ given in (2.56). Thus, (A3) is a consequence of [GHO22b, Theorem 3.2].

(A4).   Define the statistics

(4.2) $$\hat{\theta}^{M,n} = \frac{1}{|\mathbb{G}_n|} \sum_{i \in \mathbb{G}_n} \frac{x_i + \sum_{m=0}^{M} K_m y_{i,m}}{1 + \sum_{m=0}^{M} K_m},$$

and recall the *empirical measure* $\mathcal{E}_{M,n}$ defined in (2.16). Recall that $\hat{\theta}^{M,n}$ is the *average* w.r.t. $\mathcal{E}_{M,n}$ of

(4.3) $$\frac{x_0 + \sum_{m=0}^{M} K_m y_{0,m}}{1 + \sum_{m=0}^{M} K_m}.$$

We get the definition of $\mathcal{J}$ in (A3) from the general ergodic theorem in [Kre85, Section 6.4].

(A5).   The continuity property of the statistic $\hat{\theta}^{M,n}$ and the ergodic behaviour of the infinite system can be deduced with the help of L^2-theory and moment calculations, given by Lemma 3.6.

(A6).   For the time scale choose

(4.4) $$\beta_{M,n} = \kappa_M |\mathbb{G}_n|, \qquad \kappa_M = \left(1 + \sum_{m=0}^{M} K_m\right)^2.$$

We have to show that, with $z^{M,n} = (z_{i,m})_{i \in \mathbb{G}_n, 0 \le m \le M}$,

(4.5) $$\mathcal{L}\left[\left(\hat{\theta}^{M,n}(z^{M,n}(t\beta_{M,n}))\right)_{t \ge 0}\right], \qquad n \in \mathbb{N},$$

is tight in the space of paths $C([0, \infty), [0, 1])$. For that purpose we consider the *increasing process* on time scale $\beta_{M,n}$, and compute (recall (2.49))

(4.6) $$\langle \hat{\theta}^{M,n} \rangle = \left\langle \hat{\theta}(z^{M,n}(t\beta_{M,n})) \right\rangle_{t \ge 0} = \left(\int_0^t \mathrm{d}s \, \frac{1}{|\mathbb{G}_n|} \sum_{i \in \mathbb{G}_n} g(x_i(s\beta_{M,n}))\right)_{t \ge 0}.$$

Our task is to show that as $n \to \infty$ the right-hand side converges to

(4.7) $$\left(\int_0^t \mathrm{d}s \, (\mathcal{F}g)(\Theta(s))\right)_{t \ge 0},$$

with $(\Theta(u))_{s \ge 0}$ a non-trivial limit process that starts at θ and as $u \to \infty$ tends to 0 or 1, in other words, the choice of time scale $\beta_{M,n}$ in (4.4) is proper. The associated processes for finite n are *continuous martingales*, namely, time-changed Brownian motions, and the increasing processes are bounded and continuous with a bounded derivative on finite time intervals (uniformly in n). Hence the time-changed Brownian motions are tight in $C([0, \infty), \mathbb{R})$, and so are the martingales. Since $s \mapsto \Theta(s)$ is continuous, it follows that

(4.8) $$s \mapsto \mathbb{E}\left[f(\Theta(s))\right] \text{ is continuous for } f \in C([0, 1], \mathbb{R}),$$

for every weak limit point arising from (4.5), where $\mathbb{E}$ is with respect to the law of $\Theta(s)$.

(A7).   Translation invariance of laws is preserved under weak convergence. Tightness follows from compactness of the state space.

(A8).   We choose to work with coupling rather than with duality, because this will be convenient and instructive later, when we include general g. In [CGS95], assumptions (A8) and (A9) are verified for the model without seed-bank when $\mathcal{T}_n, \mathcal{T}$ are translation invariant laws, based on the construction of a coupling of the two finite or the two infinite processes, starting from different initial points, and a coupling of the finite and the infinite system, with the finite system starting in the translation invariant version of the restriction of the infinite system. The coupling is done by constructing the two processes as strong solutions of an SSDE with the same Brownian motions.

In order to prove (A8), it suffices to have a coupling and a Lyapunov function such that the distance between the two coupled processes can be controlled by a Lyapunov function with non-increasing expectation. For the model with seed-bank this coupling was constructed in [GHO22b, Section 5.3]. Hence, we indeed have (A8).

(A9).   In order to prove (A9) (i.e., to establish ergodicity), we proceed as was done for systems without seed-bank in [CGS95, Proposition 2.4]. In particular, we define the bivariate process

$$(4.9) \qquad (z^{M,n}, z)$$

as the strong solutions to the corresponding SSDEs in (2.41)–(2.42), respectively, (1.15)–(1.16) with the same collection $(w_i(t))_{t\geq 0}$, $i \in \mathbb{G}$, of standard Brownian motions, starting from initial laws in $\mathcal{T}_n, \mathcal{T}$ that are linked as in (2.9). As Lyapunov function we use the same quantity as in [GHO22b, Lemma 5.8]. Because of (A8), it suffices to prove Lemma 4.1 below.

(A10).   This assumption will be verified in Step 3.

Step 2: Coupling the finite system and the infinite system.   To compare the finite and the infinite system, we apply the coupling techniques that were developed in [GHO22b] to deal with different initial laws and in [GHO22a] to deal with different dynamics. Abbreviate $\mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m\in\mathbb{N}_0}\}$, $\mathbb{S}_M = \mathbb{G} \times \{A, (D_m)_{0\leq m\leq M}\}$ and $\mathbb{S}_{M,n} = \mathbb{G}_n \times \{A, (D_m)_{0\leq m\leq M}\}$.

Lemma 4.1. [Comparison finite-infinite] *Let $(z^{M,n}, z)$ be the bivariate process on $E_{M,n} \times E$ defined in (4.9), with initial laws in $\mathcal{T}_n, \mathcal{T}$ linked as in (2.9). Write $z = (z_u)_{u\in\mathbb{S}}$ and $z^{M,n} = (z_u)_{u\in\mathbb{S}_{M,n}}$, and denote by $\Lambda_{M,n}$ the diagonal of $\mathbb{S}_{M,n} \times \mathbb{S}_{M,n}$. Let $\mathcal{V}^{M,n}$ be the collection of all probability measures $\bar{\mu}$ on $E_{M,n} \times E$ such that*

$$(4.10) \qquad \bar{\mu}\left(z_u^{M,n} = z_u \ \forall u \in \Lambda_{M,n}\right) = 1.$$

Then

$$(4.11) \qquad \lim_{n\to\infty} \sup_{\bar{\mu}\in\mathcal{V}^{M,n}} \mathbb{E}_{\bar{\mu}}|z_u^{M,n}(t) - z_u(t)| = 0 \qquad \forall\, u \in \mathbb{S}_M \times \mathbb{S}_M,\, t > 0.$$

In particular, there exists a sequence of times $(\ell_n)_{n\in\mathbb{N}}$, satisfying $\lim_{n\to\infty} \ell_n = \infty$, such that

$$(4.12) \qquad \lim_{n\to\infty} \sup_{\bar{\mu}\in\mathcal{V}^{M,n}} \mathbb{E}_{\bar{\mu}}|z_u^{M,n}(\ell_n) - z_u(\ell_n)| = 0 \qquad \forall\, u \in \mathbb{S}_M \times \mathbb{S}_M.$$

Proof. We have to estimate the *Lyapunov function*

$$(4.13) \qquad t \mapsto \mathbb{E}\left(|x_i^{M,n}(t) - x_i(t)| + \sum_{m=0}^{M} K_m\, |y_{i,m}^{M,n}(t) - y_{i,m}(t)|\right), \qquad i \in \mathbb{G},$$

and show that the right-hand side tends to zero as $t \to \infty$. Using the computations on the coupling in [GHO22b, Section 5.3], carried out for the infinite system starting from different initial configurations, we can proceed exactly as in the proof of [CGS95, Proposition 2.4(a)], replacing the Green function $\hat{G}(0,0)$ of $\hat{a}(\cdot,\cdot)$ appearing there by the hazard integral $\hat{B}(0,0)$ defined in (A.2). The fact that the coupling is successful follows from Lemmas 3.3–3.4. $\qquad\square$

Step 3: Completion of the proof for $M < \infty$. We verify Assumption (A10) in Corollary B.3 in Appendix B, which will complete the proof of Theorem 2.10 for the case where $M < \infty$. A key observation made in [GHO22b, Section 5.1] is the fact that

$$(4.14) \qquad \mathcal{M}_{M,n}(t) = \lim_{n\to\infty} \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} \left[\frac{x_i^{M,n}(t) + \sum_{m=0}^M K_m y_{i,m}^{M,n}(t)}{1 + \sum_{m=0}^M K_m} \right], \qquad t \geq 0,$$

is a martingale that has (4.6) as increasing process. From this we want to conclude that: (i) the time-transformed processes converge in law and hence so do the martingales themselves; (ii) the limit is the diffusion with the claimed diffusion function.

No new ideas are needed. Lemmas 4.2–4.4 below are versions of [DGV95, Lemmas 4.8-4.10] reformulated for our purposes (see also Appendix B). Their proof is a straightforward adaption of the proofs given in [DGV95, Section 4(b)]. Both (i) and (ii) can be treated exactly as in the proofs of [DGV95, Lemmas 4.8–4.10], after we replace the martingale $(\Theta(s))_{s\geq 0}$ used there by the martingale in (4.14) adapted to the seed-bank, and use the facts derived in [GHO22b, Section 5.3]. Recall the definition of $\mathcal{T} = \mathcal{T}([0,1]^{\mathbb{G}\times\mathbb{N}_0})$ in (1.24) and $\mathcal{E}^{M,n}$ in (2.44).

Lemma 4.2. [Tightness of the macroscopic variable] *(a) The sequence*

$$(4.15) \qquad \mathcal{L}\left[\left(\left\{\int_0^t ds\, \frac{1}{|\mathbb{G}_n|} \sum_{i\in\mathbb{G}_n} g(x_i^{M,n}(s\beta_{M,n}))\right\}, \hat{\theta}_n(t\beta_{M,n})\right)_{t\geq 0}\right], \qquad n \in \mathbb{N},$$

is tight in $C([0,\infty), [0,\infty) \times [0,1])$.
(b) The sequence $\mathcal{L}[(\hat{\theta}^{M,n}(t\beta_{M,n}))_{t\geq 0}]$, $n \in \mathbb{N}$, is tight in $C([0,\infty), [0,1])$.
(c) The sequence $\mathcal{L}[(\int_0^t ds\, \mathcal{E}^{M,n}(s\beta_{M,n}))_{t\geq 0}]$, $n \in \mathbb{N}$, is tight in $C([0,\infty), \mathcal{T}([0,1]^{\mathbb{G}\times\{0,\dots,M\}}))$.

Lemma 4.3. [Convergence of the macroscopic variable] *Suppose that $(n_k)_{k\in\mathbb{N}}$ is such that*

$$(4.16) \qquad \lim_{k\to\infty} \mathcal{L}\left[\left(\hat{\theta}^{M,n_k}(t\beta_{M,n_k})\right)_{t\geq 0}\right] = \mathcal{L}[(\Theta(t))_{t\geq 0}].$$

Then

$$(4.17) \qquad (\Theta(t))_{t\geq 0} \quad \text{is a continuous martingale,}$$

$$(4.18) \qquad \left(\Theta(t)^2 - \int_0^t ds\, (\mathcal{F}g)(\Theta(s))\right)_{t\geq 0} \quad \text{is a martingale.}$$

Lemma 4.4. [Identification of the scaling limit] *The martingale problem in (4.17)–(4.18) has a unique solution, namely, the diffusion with diffusion function $\mathcal{F}g$.*

Lemma 4.3 guarantees that the time scale $\beta_{M,n}$ is proper, i.e., $(\Theta(t))_{t\geq 0}$ is a non-trivial random process that starts at θ and eventually converges to either 0 or 1.

Lemmas 4.3–4.4 imply the assumption in Corollary B.3, as proved in [DGV95, Section 4(d)] with the help of stochastic analysis of semi-martingales. To prove (4.5), we can follow the strategy in [DGV95, Section 4(d)] with only minor changes. To identify the time scale $\beta_{M,n}$ and verify (4.18), we start with the following observation about the dual process. The fraction of time the two random walks are jointly active equals $\kappa = 1/(1+\rho)^2$. Therefore the mixing property in (2.3) tells us that $\kappa|\mathbb{G}_n|$ is the average time it takes the two random walks to meet on $\mathbb{G}_n$ and be jointly active. When they do, they coalesce a rate d, and so coalescence eventually occurs on time scale $\beta_{M,n} = \kappa|\mathbb{G}_n|$.

The proof of Lemmas 4.2–4.4 follows the line of argument in [DGV95, Section 4(d)], and rests on Lemmas 3.3–3.4 and 3.8–3.9.

Step 4: Extension to $M = \infty$. We show how to drop the assumption $M < \infty$. Since $\rho < \infty$, $\hat{\theta}^n$ can be approximated uniformly by $\hat{\theta}^{M,n}$ as $M \to \infty$ uniformly in n on the time scale β_n, and we do not need to change the scale other than by the constant $\kappa = (1+\rho)^2$. Therefore we have uniform convergence to the model with $M = \infty$ when $\lim_{n\to\infty} M_n = \infty$. Hence the same argument as in Steps 1-3 goes through, provided we show that the equilibrium for the model with M_n converges to the equilibrium for the model with $M = \infty$. For $g = dg_{\mathrm{FW}}$, $d \in (0,\infty)$, this follows from duality, because all moments of the equilibria converge.

Indeed, denote by $(\Theta^M(s))_{s \geq 0}$ the process in (2.32) for Model 2 with M seed-banks, and let ν_θ^M be its equilibrium on $\mathbb{G}$. We need that

(4.19)
$$\lim_{M \to \infty} \nu_\theta^M = \nu_\theta$$

and

(4.20)
$$\lim_{M \to \infty} \mathcal{L}\left[(\hat{\theta}^M(t))_{t>0} \right] = \mathcal{L}\left[(\Theta(t))_{t>0} \right],$$

where the latter is the $(\mathcal{F}g)$-diffusion starting in θ. Clearly, (4.19) implies (4.20). Indeed, (4.19) yields $\lim_{M \to \infty} \mathcal{F}_M g = \mathcal{F} g$ (with $\mathcal{F}_M$ the renormalisation map for the model with M seed-banks). This in turn yields convergence of the solution of the corresponding martingale problem (by [EK86, Lemma 5.1, Chapter 4] and [JM86, Proposition 3.2.3]). It remains to show (4.19). For $g = dg_{\text{FW}}$, $d \in (0, \infty)$, this follows from convergence of the dual process, which implies convergence of moments.

Step 5: Extension to general g. Finally, we point out how to modify the argument for general g. Duality entered into the proof of (A1), and also in the proof that the limiting point arising in (A6) leads to the dynamics given by (4.18). The former arises in the standard construction of the model (see [GHO22b]). The latter requires us to verify that, conditional on $\hat{\theta}^{M,n}$, the increasing process defined in (4.6) satisfies a law of large numbers in L_2. This verification is based on the property that ν_θ is ergodic for every θ, irrespectively of $g \in \mathcal{G}$ (see [GHO22b]), which follows from Lemmas 3.5–3.6. As to the other assumptions, without duality we can work with the generator and verify these assumptions directly with the help of Lemma 3.9. See [DGV95, Section 1]. We also need that $\theta \mapsto \nu_\theta$ is continuous, a fact that was established in [GHO22b, Section 6.3.1, Lemma 6.6] with the help of coupling.

5   Proofs: $\rho = \infty$ and slow growing seed-bank

In this section we prove Theorem 2.11. We follow the same line of argument as in Section 4, but with a number of adaptations. In order to apply the abstract scheme in Appendix B, it no longer suffices to incorporate the seed-bank via an extension of the single-component state space, as was done for $\rho < \infty$ in Section 4. Instead, we need to *extend the geographic space with the seed-bank space*, namely, consider the state space $I^{\mathbb{S}}$ with $I = [0, 1]$, respectively, $I^{\mathbb{S}_n}$ with $I = [0, 1]$ (see Remark B.1). We will see below that, because the active population is a negligible fraction of the total population, new arguments are needed to control the deep seed-banks, based on the computations in Section 3. As shown in [GHO22b], the validity of the ergodic theorem for the infinite system requires the additional assumption of *colour regularity* for the initial law.

The main change is that the macroscopic variable and the active macroscopic variable (recall (2.43) and (2.50)) evolve on *different* times scales, namely, $\kappa_{M_n} |\mathbb{G}_n|$, respectively, $|\mathbb{G}_n|$ (provided the initial law is non-degenerate). Since the former is asymptotically larger than the latter, under certain conditions this opens up the possibility that correlations over typical distances in $\mathbb{G}_n$ change via migration before the macroscopic variable is able to move via the Brownian motions, leading to a break down of ergodicity in the geographic space. Consequently, if we try to follow the abstract scheme as for $\rho < \infty$ in Section 4, then we possibly run into problems.

Fix a sequence $(M_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} M_n = \infty$. We will see that, as long as M_n is in the slow growth regime (Regime (I) in (2.61)), the proper time scale is $\beta_n = \kappa_{M_n} |\mathbb{G}_n|$. The factor $\kappa_{M_n} = (1 + \sum_{m=0}^{M_n} K_m)^2$, which is the square of the size of the seed-bank, compensates for the fact that the seed-bank slows down the evolution and reduces the volatility. The proper time scale for the macroscopic variable must be calculated by looking at the *increasing process* $\langle \hat{\theta}^{M_n,n} \rangle$ and using Itô's formula. In particular, we must check tightness of $(\hat{\theta}^{M_n,n}(\cdot \beta_n))_{n \in \mathbb{N}}$ via (4.2). Because g is bounded, the expression we obtain for the increasing process at time $s\beta_n$ is bounded by $\|g\|_\infty s$. Our goal is to prove the analogue of (4.7) and show that the limit proces $(\Theta(s))_{s \geq 0}$ is non-trivial. For that we need to show that, on time scale β_n, $(\Theta(s))_{s \geq 0}$ has non-zero fluctuations, which gradually vanish as the boundary $\{0, 1\}$ is approached or hit. The latter is possible because of the slow growth of M_n.

To ensure that the line of argument used for $\rho < \infty$ and $M < \infty$ in Section 4 carries over to $\rho = \infty$ and $M = \infty$, we need the *new setting* of the abstract scheme mentioned above (extension of the geographic space with the seed-bank space). This also requires an adaptation of the sets from which we draw the initial

law. For $\rho < \infty$ these were the sets $\mathcal{T}, \mathcal{T}^{\mathrm{erg}}$ defined in (1.24) and $\mathcal{T}_\theta^{\mathrm{erg}}$ defined in (2.30), (2.56). For $\rho = \infty$ these sets must be replaced by $\mathcal{T}^\bullet, \mathcal{T}^{\mathrm{erg},\bullet}$ defined in (1.26) and $\mathcal{T}_\theta^{\mathrm{erg},\bullet}$ defined (2.59). Similarly, $\mathcal{R}_\theta^{(2)} \subset \mathcal{T}_\theta^{\mathrm{erg}}$ must be replaced by $\mathcal{R}_\theta^{(2),\bullet} \subset \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ (recall Definition 3.2).

In this setting, Assumptions (A1)–(A3) are straightforward in view of the results for the infinite system summarised in Section 1.4 (see the paragraph on equilibria) in combination with the preservation property stated in Lemma 3.4, which uses and extends the computations carried out in [GHO22b]. Assumptions (A4)–(A10), however, are not straightforward. There are *four tasks*:

(A) To settle (A4)–(A8), the key obstacle is that the seed-banks are not translation invariant and ergodic in the seed-bank direction. We use the results in Section 3 to gain control over the 'deep' seed-banks. To verify Assumptions (A4)–(A7), for $\rho = \infty$ we take into account that, in order to control the long-time behaviour of the infinite system, we must require that the initial law $\mu(0) \in \mathcal{T}^{\mathrm{erg}}$ (recall (2.59)) is *colour regular*, i.e., $\mu(0) \in \mathcal{T}^{\mathrm{erg},\bullet}$ (recall (1.25)). Furthermore, to verify Assumption (A8), we show that the *colour regularity is preserved over time* and that the *conserved quantity* is *continuous in the initial state*. Consequently, we have continuity of the ergodic theorem and we can apply the ergodic theorem that is known for the infinite system.

(B) To verify Assumption (A9), we first use a coupling argument to prove *uniformity of the ergodic theorem* in the *initial law*, over a time stretch $o(\beta_n)$. In other words, we show that for the slow growing seed-bank the problem of the breaking of ergodicity mentioned above does not occur, and the macroscopic variable depends continuously on the initial state.

(C) To complete the verification of Assumption (A9), we next check that the approximation of the infinite system by the finite system in Lemma 4.1 still holds in modified form for the slow growing seed-bank when $\rho = \infty$. For this we construct a coupling of the two systems that is successful because of the properties established in Task (B), and use that as time proceeds the deep seed-banks become deterministic.

(D) To verify Assumption (A10), we deal with the fact that $\hat{\theta}$ no longer is a continuous functional of the empirical measure, since the active population is negligible in the limit.

We will address these items in Sections 5.1–5.4.

5.1 Task (A)

We need to verify Assumptions (A4)–(A8). In order to transfer the ergodic theorem that is known for the infinite system to the sequence of finite systems, we use a restarting argument that requires modified versions of Assumptions (A4) and (A8) where the initial laws are drawn from $\mathcal{R}_\theta^{(2),\bullet}$ for some $\theta \in [0, 1]$ (see above) rather than from $\mathcal{T}_\theta^{\mathrm{erg},\bullet}$, as was done in Section 4. In particular, we need the fact that all weak limit points along sequences of times that are $o(\beta_n)$ satisfy the *Liggett conditions* (recall Definitions 3.1–3.2) and are *colour regular* (recall(1.25)), as was shown in Lemma 3.4.

Once we have established (A4)–(A5), (A6) will follow from the fact that $\hat{\theta}^{M_n,n}(t\beta_n)$ is a continuous-path martingale with bounded increasing process given by (2.49). In what follows we first verify on (A4), (A7) and then (A5), (A8).

To verify (A4) we argue as follows. For the system on $\mathbb{S}_n$ we determine the random variable Θ (the macroscopic variable) corresponding to the current state as required for (A4) via the estimator $\hat{\theta}^{M_n,n}$, which determines an associated equilibrium measure for the evolution on $\mathbb{S}$ that is our candidate for the limit of our sequence of scaled finite systems. In particular, (A4) requires that the estimator $\hat{\theta}^{M_n,n}$ defined in (4.2) converges in probability to a number Θ as $n \to \infty$, which for ergodic and colour regular initial laws is the parameter Θ selecting the extremal equilibrium measure ν_Θ in whose domain of attraction the system is. Assumption (A4) will follow for $\rho = \infty$, $M = \infty$ if we can establish L_2-convergence of the average in (4.3) as $M \to \infty$, under all possible current laws of our system, i.e.,

$$(5.1) \qquad \lim_{n\to\infty} \mathcal{L}[\hat{\theta}^{M_n,n}(t)] = \mathcal{L}[\hat{\theta}^{\infty,\infty}(t)] \qquad \text{in } L_2 \qquad \forall\, t \ge 0,$$

for some random variable $\hat{\theta}^{\infty,\infty}(t)$, playing the role of the macroscopic variable associated with the limit law of the finite system as $n \to \infty$.

As shown in [GHO22b, Section 6], subject to (2.58) we know that the infinite system is L_2-ergodic and that

$$(5.2) \qquad \lim_{M \to \infty} \mathbb{E}_{\mu(t)} \left[\frac{x_0 + \sum_{m=0}^{M} K_m \, y_{0,m}}{1 + \sum_{m=0}^{M} K_m} \right] = \theta \qquad \forall \, t \geq 0,$$

where $\mu(t)$ is the law of the system at time t. Via Lemma 3.8 this yields Assumption (A4).

Assumption (A7) amounts to showing that colour regularity is preserved under the evolution up to times of order β_n. But this follows from Lemma 3.9 via ergodic decomposition.

To settle Assumptions (A5) and (A8), which we treat together, we actually need more, namely, the existence of the limit in (5.2) for *sequences* of initial laws $(\mu_n)_{n \in \mathbb{N}}$ induced by the state of the finite system at time $t\beta_{M_n,n}$. Therefore two complications arise:

(1) In (5.2) we need convergence in law rather than in expectation.

(2) We need that *weak limit points* as $n \to \infty$ have the property that the limit of the macroscopic variable as $M \to \infty$ exists, i.e., $\hat{\theta}^{M_n,n}(s\beta_{M_n,n})$ converges in L_2 (and hence in law) to a random variable $\Theta(s)$, and conditional on $\Theta(s)$ colour regularity holds.

We want to show that, for a subsequence $(n_k)_{k \in \mathbb{N}}$ and for the law of the pair $((x^{M_n,n}, y^{M_n,n}), \hat{\theta}^{M_n,n})$ evaluated at time $s\beta_{M_n,n}$, we have

$$(5.3) \qquad \lim_{k \to \infty} \mathcal{L}_\nu \left[\hat{\theta}^{M_{n_k},n_k}(t\beta_{M_{n_k},n_k}) \right] = \mathcal{L}_{\nu(t)}[\Theta(t)] \quad \text{for some limit law } \nu(t).$$

In fact, we want to show the sharper statement

$$(5.4) \qquad \lim_{k \to \infty} \mathcal{L}_{\nu_{n_k}} \left[\hat{\theta}^{M_{n_k},n_k}(t\beta_{M_{n_k},n_k}) \right] = \mathcal{L}_{\nu(t)}[\Theta(t)] \qquad \forall \, (\nu_{n_k}) : \lim_{k \to \infty} \nu_{n_k} = \nu(t).$$

This we will get by coupling the evolution on $\mathbb{S}_n$ and the evolution on $\mathbb{S}$, which we provide in Task (C) via a restart argument for ν_{n_k}. For the latter we use the tightness part of Assumption (A6), which gives us convergence in law of the process $\hat{\theta}^{M(n_k),n_k}$ as $k \to \infty$ along a subsequence $(n_k)_{k \in \mathbb{N}}$. From (5.3)–(5.4) we conclude that

$$(5.5) \qquad \lim_{n \to \infty} \nu_n(t) = \mathcal{L}\left[\nu_{\Theta(t)}\right], \qquad \nu_n(t) = \mathcal{L}\left[\left(x^{M_n,n}(t\beta_{M_n,n}), y^{M_n,n}(t\beta_{M_n,n})\right)\right],$$

which gives Assumption (A5). Note that $\nu_n(s)$ is concentrated on configurations in $\mathbb{S}$ that are *periodic* in space and *constant* (equal to the estimator) in the seed-bank beyond colour M_n, since they are lifted from $\mathbb{S}_n$ to $\mathbb{S}$.

To prove (5.3), recall from Section 4.2 that for $\rho < \infty$ we used L_2-theory of stationary random fields. For $\rho = \infty$ this needs to be amended. We need to show that $\hat{\theta}^{M_n,n}(s\beta_n)$ converges in law and in L_2 to a random variable $\Theta(s)$ as $n \to \infty$. We also need to show that $\lim_{m \to \infty} \mathbb{E}_{\nu(s)}[y_{i,m}] = \theta$, i.e., on the set of invariant laws $\mathcal{I}$ (recall (1.24)) the property in (3.33) is valid even in the limit as $n \to \infty$. But this was already settled in Lemma 3.9. Since the averages take values in $[0,1]$, we have compactness of their laws, so that along a further subsequence we have weak convergence of their laws. The limit point must have mean θ, which arises from the existence as $M \to \infty$ of the mean of the initial law (recall (2.58)), which is an immediate consequence of first moment computations via the dual.

To prove (5.4), which is a *continuity property* of the limiting L_2-average as a function of ν_{n_k}, we use that, by Lemma 3.7, the deep seed-banks are *deterministic uniformly in n*. In fact, for any sequence of initial measures $(\nu_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} \nu_n = \nu$ for some ν, we also have that $\lim_{n \to \infty} \hat{\theta}^{M_n,n}(s\beta_n) = \theta(s)$ for some $\theta(s)$. To prove (5.5), note that every ergodic component of the state at time $s\beta_n$ satisfies $\lim_{n \to \infty} \lim_{m \to \infty} \mathbb{E}_{\nu_n}[y_{0,m}] = \theta(s)$, which fixes $\nu_{\theta(s)}$, the corresponding equilibrium approached after a restart. This is precisely the requirement in Assumption (A8), and altogether settles Assumption (A5).

5.2 Task (B)

We prove that, once we have a successful coupling as in Lemma 4.1 (which will be established in Task (C)), we can verify Assumption (A9). Here, we need to decompose the sequence $(\mu_n)_{n \in \mathbb{N}}$ and its limit points μ into ergodic components by using Lemma 3.9. For each of the ergodic components we can apply the version of Lemma 4.1 adapted to $\rho = \infty$, to get that $\hat{\theta}(\mu_n(t_n))$ and $\mu_n(t_n)$ have limit points along the *same* subsequential limits. This implies that we can replace $\mu_n(t_n)$ by $\mu(t_n)$ in the limit as $n \to \infty$.

5.3   Task (C)

We can define the bivariate dynamics in the same way as we did for Lemma 4.1, namely, as in (4.9) with $M = M_n$ for the finite system and $M = \infty$ for the infinite system. But the quantity in (4.13) with M replaced by M_n is no longer well defined after we have passed to the limit $n \to \infty$ for the bivariate system, except for configurations drawn for a *restricted class* (which we will achieve by restricting the class of initial laws). This must be resolved through the fact that the deep seed-banks become deterministic after a long time. Two problems come up (which for the infinite system were resolved in [GHO22b]):

(1) The Lyapunov function for the infinite system (recall (4.13)) is well defined only for certain classes of initial laws in the limit as $M_n \to \infty$.

(2) Instead of the monotone decreasing Lyapunov function for the infinite system, we now have a function where the derivative also involves positive terms, arising from the migration in the infinite system, transporting mass in and out of $\mathbb{G}_n$. In addition, there is exchange with the deep seed-banks of colour beyond M_n.

Ad (1): This observation restricts the set of initial measures in $\mathcal{T}_\theta^{\mathrm{erg}}$ that we can compare via coupling, which was already an issue in the proof of the ergodic theorem for the infinite system in [GHO22b]. Ad (2): For fixed t the extra positive terms vanish when $M_n, n \to \infty$, so this puts a restriction on the growth rate of the sequence $(t_n)_{n\in\mathbb{N}}$.

Therefore the following analogue of Lemma 4.1 is needed in order to provide the successful coupling used in Task (B). Recall from (3.57) that $\mathcal{T}_\theta^{\mathrm{erg},\blacksquare} \subset \mathcal{T}_\theta^{\mathrm{erg}}$ is the set of initial laws that are invariant and ergodic under translations such that

$$(*) \qquad \forall\, i \in \mathbb{G}: \quad \lim_{m\to\infty} y_{i,m} = \theta \quad \text{in probability.}$$

Lemma 5.1.   [Comparison finite-infinite: $\rho = \infty$, Regime (I)]

(a) *For every $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$, and every $(t_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} t_n/\beta_n = 0$, all weak limit points of $((\Phi_n\mu)S_n(t_n))_{n\in\mathbb{N}}$ are in $\mathcal{T}_\theta^{\mathrm{erg},\blacksquare}$.*

(b) *For every $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ the same properties as in Lemma 4.1 hold.*

Proof. Note that (a) and (b) do not imply that the conclusions in Lemma 4.1 hold for every $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$, which we used to obtain the claim in Task (B) for finite seed-banks. We need to restrict to $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ and construct a successful coupling by showing that the limit of $\mu(t)$ as $t \to \infty$ lies in $\mathcal{T}_\theta^{\mathrm{erg},\blacksquare}$. This is achieved by considering two time scales, which we call $(t_n^*)_{n\in\mathbb{N}}$ and $(\bar{t}_n)_{n\in\mathbb{N}}$.

(a) To prove part (a), we proceed as follows.

- Choose a sequence $(t_n^*)_{n\in\mathbb{N}}$ such that

  (i) all weak limit points of $\{\mu S(t_n^*)\}_{n\in\mathbb{N}}$ lie in $\mathcal{T}_\theta^{\mathrm{erg},\blacksquare}$,

  (ii) all weak limit points of $\{\tilde{\Phi}_n[(\Phi_n\mu)S_n(t_n^*)]\}_{n\in\mathbb{N}}$ lie in $\mathcal{T}_\theta^{\mathrm{erg},\blacksquare}$,

  (iii) $\lim_{n\to\infty} t_n^* = \infty$ and $\lim_{n\to\infty} t_n^*/\beta_n = 0$.

  where $(S(t))_{t\geq 0}$ and $(S_n(t))_{t\geq 0}$ are the semi-groups of the infinite, respectively, finite system, and $(\tilde{\Phi}_n)_{n\in\mathbb{N}}$ are the extension operators defined in (2.12) (see also the general set-up in Appendix B, in particular, (B.7)). In our setting, $\tilde{\Phi}_n$ is the periodic continuation from $\mathbb{G}_n$ to $\mathbb{G}$ in the geographic coordinate and the constant continuation by $\hat{\theta}^{M_n,n}$ from $\{0, \dots, M_n\}$ to $\mathbb{N}_0$ in the seed-bank coordinate.

- Choose a sequence $(\bar{t}_n)_{n\in\mathbb{N}}$ satisfying $\lim_{n\to\infty} \bar{t}_n = \infty$ and $\lim_{n\to\infty} \bar{t}_n/\beta_n = 0$ such that the Markov jump process Z on $\mathbb{G} \times \mathbb{N}_0$ with transition kernel $b^{(2)}(\cdot, \cdot)$ satisfies

  $$(**) \qquad \lim_{n\to\infty} \mathbb{P}_0\Big(Z \text{ leaves } \mathbb{G}_n \times \{0, \dots, M_n\} \text{ before time } \bar{t}_n\Big) = 0.$$

Property (i) holds for $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ by the ergodic theorem for $(S(t))_{t\geq 0}$ and properties of the equilibrium, stated in Lemma 3.4. Property (ii) holds by Lemma 3.7. Property (iii) is imposed to limit the time scale, so that we are in a position to carry out Task (B).

The choice of t_n^* and $\bar{t}_n$ must achieve that, along any subsequence $(n_k)_{k \in \mathbb{N}}$,

$$(***) \qquad \lim_{k \to \infty} \mathcal{L}[(z^{M_n,n}(t_{n_k}^* + \bar{t}_{n_k}), z(t_{n_k}^* + \bar{t}_{n_k})] = \Gamma \implies \Gamma(\{z_1, z_2 \in \text{Diag}(E \times E)\}) = 1.$$

On the time interval $[0, t_n^*]$ the system converges along subsequences to a state in which we can couple the two sequences of seed-banks at every $i \in \mathbb{G}$ because the two components $(z_i^{M_n,n}, z_i)$ agree on the tail sigma-field even for the limit state, as follows from Lemma 3.7. On the time interval $[t_n^*, \bar{t}_n]$ we want to use the bivariate dynamics for the finite, respectively, the infinite system with its Lyapunov function to obtain the successful coupling. We want to combine the two to arrive at $(***)$. Two problems arise:

- $(*)$ only holds for the weak limit points in (i) and (ii), and not necessarily for finite n.

- During the time interval $[0, t_n^*]$ discrepancies between the finite and the infinite system may build up in $\mathbb{G}_n \times \{A, (D_m)_{0 \le m \le M_n}\}$.

The second item will be dealt with in the proof of part (b).

Concerning the first item, it suffices to show that at the end of the interval $[t_n^*, t_n^* + \bar{t}_n]$ the finite system looks like the equilibrium measure of the infinite system in whose domain of attraction a weak limit point of $((\Phi_n \mu) S_n(t_n^*))_{n \in \mathbb{N}}$ lies. This property is enough because by running the system for a total time $t_n^* + \bar{t}_n$ we can conclude that $(***)$ holds. To obtain said property, we have to argue that we can make the replacement at time t_n^* for any weak limit point. In particular, we have to prove that any weak limit point of $((\Phi_n \mu) S_n(t_n^*))_{n \in \mathbb{N}}$ can be successfully coupled via the bivariate finite dynamics associated with $\mathcal{L}[\bar{\nu}_{\theta^\bullet}]$, where $\theta^\bullet$ is the limit in probability of the conserved quantity of the measure μ. For this we again use coupling.

For every $k \in \mathbb{N}$, for the bivariate finite dynamics

starting in $(\Phi_k \mu, \Phi_k \bar{\mu})$, with $\mu \in \mathcal{R}_\theta^{(2)}$ and $\bar{\mu}$ a weak limit point of $((\Phi_n \mu) S_n(t_n^*))_{n \in \mathbb{N}}$,

we have a Lyapunov function (recall (4.13)) that is well-defined and that is non-decreasing. We want to show that, for the bivariate dynamics of two finite systems,

$$(5.6) \qquad \lim_{k \to \infty} \mathcal{L}[(\Phi_k \mu) S_k(t_k^*), (\Phi_k \bar{\mu}) S_k(t_k^*)] = 0$$

when $\lim_{k \to \infty} t_k^* = \infty$. We know that for both marginals, with μ and $\bar{\mu}$ as marginal initial laws,

$$(5.7) \qquad \lim_{k \to \infty} \hat{\theta}^{M_k,k}(t_k^*) = \theta \quad \text{in law}$$

when $\lim_{k \to \infty} t_k^*/\beta_k = 0$. Hence, under this condition, correlations between typical sites tend to zero as $k \to \infty$. This implies that a weak limit point of the bivariate dynamics, at time t_k^* in the limit as $k \to \infty$, lies on the diagonal, and the Lyapunov function tends to zero. Indeed, the two configurations can be ordered in the limit, because they have the same value of $\hat{\theta}$ and are ergodic. In particular, this also holds when we consider times $(t_k^* + \bar{t}_k)_{k \in \mathbb{N}}$. We next put this together to get $(***)$.

Since the finite and the infinite bivariate dynamics constructed above allow for a coupling of $\Phi_k \bar{\mu}$ and $\mu \in \mathcal{R}_\theta^{(2),\bullet}$ that is successful when run for a time $\bar{t}_k$ with $k \to \infty$, we use this is an ingredient to conclude that $\Phi_k \mu$ and ν_θ can be successfully coupled when the bivariate dynamics is run for a time $t_k^* + \bar{t}_k$ with $k \to \infty$. The latter we may view as quadrivariate dynamics: namely, take the infinite bivariate dynamics whose components start from μ, ν_θ and the finite bivariate dynamics whose components start from $\Phi_k \mu, \Phi_k \nu_\theta$, and place them on the same probability space. The claim now follows by picking $t_k = t_k^* + \bar{t}_k$ and letting $k \to \infty$.

(b) To prove part (b), we need to show that also the positive terms in the derivative of the Lyapunov function tend to zero as $n \to \infty$. Here there are two effects:

- The truncated migration on $\mathbb{G}_n$ is different from the migration on $\mathbb{G}$.

- The truncated seed-bank $\{0, \ldots, M_n\}$ is different from the seed-bank $\mathbb{N}_0$.

We need to show that both differences are negligible in the limit as $n \to \infty$.

The first was already dealt with for systems without seed-bank, for which we refer to [DGV95, Section 4, pp. 2319–2322]. The same reasoning carries over to systems with seed-bank because in Regime (I) automatically $\lim_{n \to \infty} t_n/\beta_n^{**} = 0$ (recall that $\beta_n^{**} \gg \beta_n$ is the time scale on which two active lineages in the dual coalesce on the active layer $\mathbb{G}_n$).

For the second we need the restrictions posed in Regime (I) (recall (2.60)–(2.61)) in order to be able to show that perturbations arising from the truncation of the seed-bank are negligible in the limit as $n \to \infty$. If g is a multiple of the Fisher-Wright diffusion function, then we can use duality and restrict to the event $\mathcal{A}_n$ defined in (3.22), which is the event that a single Markov process does not visit a seed-bank with colour $> M_n$ until time t_n. The claim now follows from the fact that $\lim_{n\to\infty} \mathbb{P}_n(\mathcal{A}_n) = 1$ when we choose t_n such that $\lim_{n\to\infty} t_n/\beta_n^* = 0$ (recall that $\beta_n^* \ll \beta_n$ is the time scale on which a single dormant lineage in the dual starting from the deepest seed-bank with colour M_n becomes active on $\mathbb{G}_n$).

To get the claim for general $g \in \mathcal{G}$, we need to use the coupling argument in [GHO22b, Sections 5–6], which uses contraction via a Lyapunov function to show that the coupling is successful when t_n^* has the above properties. As is clear from Section 3, the function g plays no role in the estimates as long as $\|g\| < \infty$. $\qquad\square$

5.4 Task (D)

To show that there is only *one* limit point, we can again follow the *abstract scheme* developed in [DGV95, Section 4]. We work with the tightness of the process $(\hat{\theta}^{M(n_k),n_k}(t\beta_{M(n_k),n_k}))_{t\geq 0}$ and the fact that the weak limit points must solve a *well-posed* martingale problem (see Lemmas 4.2–4.4). The latter we can obtain from the very same argument as in Section 4 for $\rho < \infty$, $M < \infty$. There we showed that $(\hat{\theta}^{M_n,n}(t\beta_{M_n,n}))_{t\geq 0}$ is a martingale for which we can determine the increasing process. Namely, in the formula for the increasing process in (4.6) we can apply a *law of large numbers* to the empirical measure, which allows us to identify the limiting martingale problem as before. Here, the difficulty is that the active component of the empirical measure is negligible in the limit as $n \to \infty$.

To circumvent this obstacle, we look at $\mathcal{E}_{L_n}^{M_n,n}$, the truncated empirical measure that includes the active component and the first L_n dormant components, with $L_n \in \mathbb{N}_0$ chosen appropriately (see Section 2.5.3). For this quantity it suffices to obtain the increasing process and to describe the empirical measure in the weak topology. Indeed, for every $L \in \mathbb{N}_0$ the truncated empirical measure converges to the truncated equilibrium measure. Since the increasing process only depends on the active component, as before we get a unique limit point and so we have *full convergence* as $n \to \infty$ (not just along subsequences).

To identify the process $(\Theta(t))_{t\geq 0}$ as the $\mathcal{F}g$-diffusion, we can follow the same route as in Step 3 in Section 4.2, without any changes. Therefore we have verified (A10).

6 Proofs: $\rho = \infty$ and fast growing seed-bank

In this section we prove Theorem 2.12. The line of argument is different from that in Sections 4–5, and requires new ideas because the conditional spatial ergodicity used in Section 5 for slow growing seed-bank fails at times of order β_n^{**}. Hence the argument in Appendix B breaks down. The three time scales in Theorem 2.12 are treated in Section 6.1–6.3, respectively. The initial law satisfies $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg},\bullet} = \mathcal{R}_\theta^{(2),\bullet}$ for some $\theta \in [0,1]$.

6.1 Convergence to equilibrium on short time scales

In this section we prove Theorem 2.12(1).

1. Consider two independent Markov processes on $\mathbb{S}_n$ with transition kernel $b^{(2),n}(\cdot,\cdot)$, both starting in the active state at the origin. Let $\mathcal{E}(t)$ and $\mathcal{E}'(t)$ be the events that the Markov processes are active at time t, and let $T(t) = \int_0^t \mathrm{d}s\, 1_{\mathcal{E}(s)}$ and $T'(t) = \int_0^t \mathrm{d}s\, 1_{\mathcal{E}'(s)}$ be their total activity times up to time t. Then their *total joint activity time* at time T (i.e., the total time the two Markov processes are jointly active at the same site up to time T) is equal in law to

$$(6.1) \qquad I^{(2),n}(T) = \int_0^T \mathrm{d}t\, 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)}\, 1_{\{RW_1^{n,\uparrow}(T(t)) = RW_2^{n,\uparrow}(T'(t))\}},$$

where $(RW_1^{n,\uparrow}(t))_{t\geq 0}$ and $(RW_2^{n,\uparrow}(t))_{t\geq 0}$ are two independent Markov processes on $\mathbb{G}_n$ with transition kernel $b^n(\cdot,\cdot)$. This is true because the transition rates between active and dormant do not depend on the location in geographic space.

We want to compare this quantity with the same quantity for the infinite system, given by

$$(6.2) \qquad I^{(2)}(T) = \int_0^T dt \, 1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \, 1_{\{RW_1^\uparrow(T(t)) = RW_2^\uparrow(T'(t))\}},$$

where $(RW_1^\uparrow(t))_{t \geq 0}$ and $(RW_2^\uparrow(t))_{t \geq 0}$ are two independent random walks on $\mathbb{G}$ with transition kernel $a(\cdot, \cdot)$. We can couple the transitions between active and dormant, provided no transition from $A \to D_m \to A$ occurs in the infinite system until time T for any $m > M_n$. However, as long as $T = o(\beta_n^*)$, the probability of this event tends to zero as $n \to \infty$ (recall (3.22)). Therefore, under this coupling we have

$$
(6.3) \qquad
\begin{aligned}
0 &\leq I^{(2),n}(T) - I^{(2)}(T) \\
&= o(1) + \int_0^T dt \, 1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \left[1_{\{RW_1^{n,\uparrow}(T(t)) = RW_2^{n,\uparrow}(T'(t))\}} - 1_{\{RW_1^\uparrow(T(t)) = RW_2^\uparrow(T'(t))\}} \right].
\end{aligned}
$$

Put

$$(6.4) \qquad \Delta^n(T) = \hat{\mathbb{E}}_{(0,A),(0,A)}[I^{(2),n}(T) - I^{(2)}(T)],$$

where the expectation is with respect to the law of the coupling. Then

$$(6.5) \qquad 0 \leq \Delta^n(T) = o(1) + \int_0^T dt \, \mathbb{E}\left[1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \left[\sum_{i \in \mathbb{G}_n} a_{T(t)}^n(0,i) \, a_{T'(t)}^n(0,i) - \sum_{i \in \mathbb{G}} a_{T(t)}(0,i) \, a_{T'(t)}(0,i) \right] \right],$$

where $\mathbb{E}$ denotes expectation with respect to the law of the process $(1_{\mathcal{E}(t)}, 1_{\mathcal{E}'(t)})_{t \geq 0}$ for the infinite system starting from $(1,1)$. Since for Model 2 with $\rho = \infty$ we have assumed that the transition kernel is symmetric (recall (1.37)), this gives

$$(6.6) \qquad 0 \leq \Delta^n(T) = o(1) + \int_0^T dt \, \mathbb{E}\left[1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \left[a_{T(t)+T'(t)}^n(0,0) - a_{T(t)+T'(t)}(0,0) \right] \right].$$

2. Our goal is to show that the right-hand side tends to zero as $n \to \infty$ as long as $T = o(\beta_n^{**})$. To do so we will use that, as shown in [GHO22b, Section 6.2],

$$
(6.7) \qquad
\begin{aligned}
&\lim_{t \to \infty} \frac{1}{t^\gamma} T(t) = V, \qquad \lim_{t \to \infty} \frac{1}{t^\gamma} T'(t) = V' \quad \text{in } \mathbb{P}\text{-probability}, \\
&\lim_{t \to \infty} t^{1-\gamma} \mathbb{P}(\mathcal{E}(t)) = \mathbb{E}[V], \qquad \lim_{t \to \infty} t^{1-\gamma} \mathbb{P}(\mathcal{E}'(t)) = \mathbb{E}[V'],
\end{aligned}
$$

where $V = W^{-\gamma}/\chi$, with W a stable law random variable on $(0, \infty)$ with exponent γ, satisfying $\mathbb{E}[V] < \infty$, and V' is an independent copy of V. We split the integral into two parts,

$$
(6.8) \qquad
\begin{aligned}
\Delta_*^n &= \int_0^{o(\psi_n^{1/\gamma})} dt \, \mathbb{E}\left[1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \left[a_{T(t)+T'(t)}^n(0,0) - a_{T(t)+T'(t)}(0,0) \right] \right], \\
\Delta_*^n(T) &= \int_{o(\psi_n^{1/\gamma})}^T dt \, \mathbb{E}\left[1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \left[a_{T(t)+T'(t)}^n(0,0) - a_{T(t)+T'(t)}(0,0) \right] \right],
\end{aligned}
$$

and show that both parts vanish.

3. For the first integral in (6.8) we use (2.6) to estimate

$$(6.9) \qquad \Delta_*^n = \int_0^{o(\psi_n^{1/\gamma})} dt \, \mathbb{E}\left[1_{\mathcal{E}(t)} \, 1_{\mathcal{E}'(t)} \, o(a_{T(t)+T'(t)}(0,0)) \right],$$

where we use that $T(t) + T'(t) \asymp t^\gamma = o(\psi_n)$ in $\mathbb{P}$-probability. Combining (6.7) and (6.9), we get

$$(6.10) \qquad \Delta_*^n = o\left(\int_0^{o(\psi_n^{1/\gamma})} dt \, (1 \wedge t^{-2(1-\gamma)}) \, a_{t^\gamma}(0,0) \right),$$

where we use that $t \mapsto a_t(0,0)$ is regularly varying at infinity (recall (1.38)) to get rid of the factor $V + V'$ and the fluctuations of $T(t) + T'(t)$ on scale t^γ. Inserting the change of variable $s = t^\gamma$, we obtain

$$(6.11) \qquad \Delta_*^n = o\left(\int_0^{o(\psi_n)} ds \, (1 \wedge s^{-(1-\gamma)/\gamma}) \, a_s(0,0) \right).$$

But

$$(6.12) \qquad \int_0^\infty \mathrm{d}s \, (1 \wedge s^{-(1-\gamma)/\gamma}) \, a_s(0,0) < \infty$$

in the coexistence regime (recall (1.42)), and so we have shown that $\lim_{n\to\infty} \Delta_*^n = 0$.

4. For the second integral in (6.8) we use the definition of the mixing time in (2.3) in combination with the comparison property in (2.7) to estimate $C^{-1}|\mathbb{G}_n|^{-1} \le a_s^n(0,0) \le C|\mathbb{G}_n|^{-1}$ for $o(\psi_n) \le s \le T$ and some $C < \infty$. This gives

$$(6.13) \qquad \Delta_*^n(T) \asymp C|\mathbb{G}_n|^{-1} \int_{o(\psi_n)}^{T^\gamma} \mathrm{d}s \, (1 \wedge s^{-(1-\gamma)/\gamma}) \asymp |\mathbb{G}_n|^{-1} \begin{cases} T^{2\gamma-1}, & \gamma \in (\tfrac{1}{2}, 1], \\ \log T, & \gamma = \tfrac{1}{2}. \end{cases}$$

Since the latter tends to zero as long as $T = o(|\mathbb{G}_n|^{1/(2\gamma-1)})$, respectively, $T = \mathrm{e}^{o(|\mathbb{G}_n|)}$ (recall (2.60) and the remark below (2.61)), we have shown that $\lim_{n\to\infty} \sup_{T=o(\beta_n^{**})} \Delta_*^n(T) = 0$. Note that we need the assumption $\psi_n = o((\beta_n^{**})^\gamma)$ because this guarantees that the divergence of the integral in (6.13) indeed occurs on time scale β_n^{**} and not later.

5. Combining the estimates Steps 2–4, we arrive at

$$(6.14) \qquad \lim_{n\to\infty} \sup_{T=o(\beta_n^{**})} [I^{(2),n}(T) - I^{(2)}(T)] = 0 \quad \text{in probability.}$$

Hence, up to time $o(\beta_n^{**})$ the two Markov processes starting from $(0,A), (0,A)$ see no difference in their total joint activity time between the finite system and the infinite system. It is easy to extend this fact to arbitrary starting points $(0, R_1), (0, R_2)$ with $R_1, R_2 \in \{A, (D_m)_{0 \le m \le L_n}\}$. Indeed, we wait until the two Markov processes are jointly active for the first time, which occurs at a finite random time (whose law depends on L_n). At that time they start from $(I,A), (J,A)$ with $I, J \in \mathbb{G}_n$ some random locations. This amounts to replacing $a_s^n(0,0)$ and $a_s(0,0)$ by $a_s^n(I,J)$ and $a_s(I,J)$ in (6.6), (6.8) and (6.9). However, this does not affect the estimates in Steps 2 and 3 because of the approximation in (2.6) and the bound in (2.5), respectively.

6. The estimates in Steps 2–4 can be trivially extended to $k \in \mathbb{N}$ Markov processes, because for each of the $\binom{k}{2}$ pairs the discrepancy between the total joint activity times tends to zero in probability. In case $g = dg_{\mathrm{FW}}$, $d \in (0, \infty)$, the dual is available and is obtained by letting active Markov processes at the same site coalesce at rate d. Consequently, all the mixed moments at time $\bar\beta_n$ have the same limit in the finite system as in the infinite system as long as $\bar\beta_n \to \infty$ and $\bar\beta_n/\beta_n^{**} \to 0$. Hence there is local convergence to ν_θ, the equilibrium for the infinite system.

7. For $g \in \mathcal{G}$ we do not have a dual, and comparison duality (exploited in Section 6.2.1) does not work either because ν_θ depends on g. However, as in Section 5, we can follow the abstract scheme from Appendix B, as long as we consider times scales $\bar\beta_n = o(\beta_n^{**})$. For this we have to work with the macroscopic variable, which on times scales $\bar\beta_n = o(\beta_n^{**})$ is the constant process equal to θ. We again have to check Assumptions (A1)–(A10). The key tools are once again Lemmas 3.3–3.7, which hold for times scales $\bar\beta_n = o(\beta_n^{**})$, as shown in Section 3.5. Steps 1-5 are needed to couple the finite and the infinite system and show decorrelation.

6.2   Partial clustering and switching on intermediate time scales

6.2.1   Partial clustering

In this section we prove the partial clustering part of Theorem 2.12(2). The proof works for $g \in \mathcal{G}$ and does not need the dual because it is based on a second moment computation only. Recall (1.53) and write $u = (u^1, u^2)$ with $u^1 \in \mathbb{G}_n$ and $u^2 \in \{A, (D_m)_{0 \le m \le M_n}\}$.

Proof. The proof is an adaptation of the proof of [GHO22b, Lemma 5.5]. Pick $\beta_n^{**} \ll \bar\beta_n \ll \beta_n^* (= \beta_n)$ and $L \in \mathbb{N}_0$.

1. To prove asymptotic *partial clustering* to depth L_n we must show that, for any $\mu \in \mathcal{T}$,

$$(6.15) \qquad \lim_{n\to\infty} \sup_{u_1, u_2 \in \mathbb{S}_n^{L_n}} \mathbb{E}_{\Phi_n \mu} \left[z_{u_1}(\bar\beta_n)(1 - z_{u_2}(\bar\beta_n)) \right] = 0.$$

Indeed, this implies that, uniformly in $u_1, u_2 \in \mathbb{S}_n^{L_n}$,

$$
(6.16)
\begin{aligned}
u_1 = u_2 &: \lim_{n \to \infty} \mathbb{P}_{\Phi_n \mu} \big(z_{u_1}(\bar{\beta}_n) \in [0, \epsilon) \cup (1 - \epsilon, 1] \big) = 1 \quad \forall \epsilon > 0, \\
u_1 \neq u_2 &: \lim_{n \to \infty} \mathbb{P}_{\Phi_n \mu} \big(z_{u_2}(\bar{\beta}_n) \in (1 - \epsilon, 1] \,\big|\, z_{u_1}(\bar{\beta}_n) \in (1 - \epsilon, 1] \big) = 1 \quad \forall \epsilon > 0.
\end{aligned}
$$

Like in [CG94a], we will give the proof by *comparison duality*. Fix $\epsilon > 0$. Since $g \in \mathcal{G}$ we can choose a $c = c(\epsilon) > 0$ such that $g(x) \geq \tilde{g}(x) = c(x - \epsilon)(1 - \epsilon - x)$, $x \in [0, 1]$. Note that $\tilde{g}(x) < 0$ for $x \in [0, \epsilon) \cup (1 - \epsilon, 1]$, so we cannot replace g by $\tilde{g}$ in the evolution equations. Instead we use $\tilde{g}$ as an auxiliary function.

2. Let $(B(t))_{t \geq 0}$ be the Markov chain with state space $\{1, 2\} \times \mathbb{S}_n \times \mathbb{S}_n$ and $B(t) = (B_0(t), B_1(t), B_2(t))$, evolving according to

$$
(6.17)
\begin{aligned}
(1, u_1, u_1) &\to (1, u_3, u_3), \quad \text{at rate } b^{(2),n}(u_1, u_3), \\
(2, u_1, u_2) &\to
\begin{cases}
(2, u_3, u_2), & \text{at rate } b^{(2),}(u_1, u_3), \\
(2, u_1, u_4), & \text{at rate } b^{(2),n}(u_2, u_4), \\
(1, u_1, u_1), & \text{at rate } c\, 1_{\{u_1^1 = u_2^1\}}\, 1_{\{u_1^2 = u_2^2 = A\}}.
\end{cases}
\end{aligned}
$$

This describes two random walks, evolving independently according to the transition kernel $b^{(2),n}(\cdot, \cdot)$, that coalesce at rate $c > 0$ when they are at the same site and both active. We put $B_0(t) = 1$ when the two random walks have already coalesced by time t, and $B_0(t) = 2$ otherwise. Let $\mathbb{P}_{(2, u_1, u_2)}$ denote the law of the Markov chain $(B(t))_{t \geq 0}$ that starts in $(2, u_1, u_2)$. Note that

$$
(6.18)
\begin{aligned}
\mathbb{P}_{(2, u_1, u_2)} (B_1(t) = u_3) &= b_t^{(2),n}(u_1, u_3), \\
\mathbb{P}_{(2, u_1, u_2)} (B_2(t) = u_4) &= b_t^{(2),n}(u_2, u_4).
\end{aligned}
$$

Below we will show that, for $L_n \ll \bar{\beta}_n^{1/\beta}$,

$$
(6.19) \qquad \lim_{n \to \infty} \sup_{u_1, u_2 \in \mathbb{S}_n^{L_n}} \mathbb{P}_{(2, u_1, u_2)}(B_0(\bar{\beta}_n) = 2) = 0.
$$

3. The evolution equations read

$$
(6.20) \qquad \mathrm{d}z_{u_1}(t) = \sum_{u_3 \in \mathbb{S}_n} b^{(2),n}(u_1, u_3)\, [z_{u_3}(t) - z_{u_1}(t)]\, \mathrm{d}t + \sqrt{g(z_{u_1}(t))}\, 1_{\{R_i = A\}}\, \mathrm{d}w_i(t), \qquad u_1 = (i, R_i) \in \mathbb{S}_n.
$$

Put

$$
(6.21) \qquad z_u^- = z_u - \epsilon, \qquad z_u^+ = z_u + \epsilon.
$$

For $t \geq 0$, define $F_t \colon \{1, 2\} \times \mathbb{S}_n \times \mathbb{S}_n \to \mathbb{R}$ by

$$
(6.22) \qquad F_t(1, u_1, u_1) = \mathbb{E}_{\Phi_n \mu}[z_{u_1}^-(t)], \qquad F_t(2, u_1, u_2) = \mathbb{E}_{\Phi_n \mu}[z_{u_1}^-(t) z_{u_2}^+(t)].
$$

Using Itô-calculus, we obtain from (6.20) that

$$
(6.23) \qquad \frac{\mathrm{d}}{\mathrm{d}t} F_1(1, u_1, u_1) = \sum_{u_3 \in \mathbb{S}_n} b^{(2),n}(u_1, u_3)\, [F_t(1, u_3, u_3) - F_1(1, u_1, u_1)]
$$

and

$$
(6.24)
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} F_t(2, u_1, u_2) = {}& \mathbb{E}_{\Phi_n \mu}\left[g(z_{u_1}(t)) \right] 1_{\{u_1^1 = u_2^1\}}\, 1_{\{u_1^2 = u_2^2 = A\}} \\
&+ \sum_{u_3 \in \mathbb{S}_n} b^{(2),n}(u_1, u_3)\, [F_t(2, u_3, u_2) - F_t(2, u_1, u_2)] + \sum_{u_4 \in \mathbb{S}_n} b^{(2),n}(u_2, u_4)\, [F_t(2, u_1, u_4) - F_t(2, u_1, u_2)].
\end{aligned}
$$

Note that

$$
(6.25) \qquad \mathbb{E}_{\Phi_n \mu}\left[\tilde{g}(z_{u_1}(t)) \right] = \mathbb{E}_{\Phi_n \mu}\left[c z_{u_1}^-(t)(1 - z_{u_1}^+(t)) \right] = c F_t(2, u_1, u_1).
$$

4. For $t \geq 0$, define $G_t \colon \{1,2\} \times \mathbb{S}_n \times \mathbb{S}_n \to \mathbb{R}$ by

$$(6.26) \qquad G_t(1, u_1, u_1) = 0, \qquad G_t(2, u_1, u_2) = \mathbb{E}_{\Phi_n\mu} \left[(g(z_{u_1}(t)) - \tilde{g}(z_{u_1}(t))) \right] 1_{\{u_1^1 = u_2^1\}} 1_{\{u_1^2 = u_2^2 A\}}.$$

Let $\mathfrak{B}$ denote the generator of $(B(t))_{t \geq 0}$, and let $(\mathfrak{S}_t)_{t \geq 0}$ be the associated semigroup. Then, with the help of (6.25)–(6.26), (6.23)–(6.24) can be written as

$$(6.27) \qquad \frac{dF_t}{dt} = \mathfrak{B} F_t + G_t$$

and hence [Lig85, Theorem I.2.15]

$$(6.28) \qquad F_t = \mathfrak{S}_t F_0 + \int_0^t ds\, \mathfrak{S}_{t-s} G_s.$$

Since $G_t \geq 0$ for all $t \geq 0$, we obtain

$$(6.29) \qquad \begin{aligned} F_t(2, u_1, u_2) &\geq (\mathfrak{S}_t F_0)(2, u_1, u_2) = \mathbb{E}_{(2,u_1,u_2)} \left[F_0(B(t)) \right] \\ &= \mathbb{E}_{(2,u_1,u_2)} \left[F_0(B(t)) 1_{\{B_0(t)=1\}} \right] + \mathbb{E}_{(2,u_1,u_2)} \left[F_0(B(t)) 1_{\{B_0(t)=2\}} \right]. \end{aligned}$$

But

$$(6.30) \qquad \begin{aligned} \mathbb{E}_{(2,u_1,u_2)} \left[F_0(B(t)) 1_{\{B_0(t)=1\}} \right] &= \mathbb{E}_{(2,u_1,u_2)} \left[\mathbb{E}_{\Phi_n\mu} [z_{B_1(t)}^-] 1_{\{B_0(t)=1\}} \right] \\ &= \mathbb{E}_{(2,u_1,u_2)} \left[\mathbb{E}_{\Phi_n\mu} [z_{B_1(t)}^-] \right] - \mathbb{E}_{(2,u_1,u_2)} \left[\mathbb{E}_{\Phi_n\mu} [z_{B_1(t)}^-] 1_{\{B_0(t)=2\}} \right], \end{aligned}$$

while

$$(6.31) \qquad \mathbb{E}_{(2,u_1,u_2)} \left[F_0(B(t)) 1_{\{B_0(t)=2\}} \right] = \mathbb{E}_{(2,u_1,u_2)} \left[\mathbb{E}_{\Phi_n\mu} [z_{B_1(t)}^- z_{B_1(t)}^+] 1_{\{B_0(t)=2\}} \right].$$

Since $F_t(1, u_1, u_1) = \mathbb{E}_{(2,u_1,u_2)}[\mathbb{E}_{\Phi_n\mu} [z_{B_1(t)}^-]]$, we get

$$(6.32) \qquad \begin{aligned} \mathbb{E}_{\Phi_n\mu} \left[z_{u_1}^-(t)(1 - z_{u_2}^+(t)) \right] &= F_t(1, u_1, u_1) - F_t(2, u_1, u_2) \\ &\leq \mathbb{E}_{(2,u_1,u_2)} \left[\mathbb{E}_{\Phi_n\mu} [z_{B_1(t)}^-(1 - z_{B_1(t)}^+)] 1_{\{B_0(t)=2\}} \right] \leq \mathbb{P}_{(2,u_1,u_2)}(B_0(t) = 2). \end{aligned}$$

Combining (6.19) and (6.32), we obtain

$$(6.33) \qquad \liminf_{n \to \infty} \sup_{u_1, u_2 \in \mathbb{S}_n^{L_n}} \mathbb{E}_{\Phi_n\mu} \left[(z_{u_1}(\bar\beta_n) - \epsilon)(1 - \epsilon - z_{u_2}(\bar\beta_n)) \right] = 0.$$

Noting that $z_{u_1}(1 - z_{u_2}) \leq (z_{u_1} - \epsilon)(1 - \epsilon - z_{u_2}) + \epsilon(2 - \epsilon)$ and letting $\epsilon \downarrow 0$, we get (6.15).

5. It remains to prove (6.19). Let τ_1, τ_2 denote the first wake-up times of two independent Markov processes with transition kernel $b^{(2),n}(\cdot, \cdot)$ starting in u_1, u_2. If $u_1^2 = u_2^2 = A$, then $\tau_1 = \tau_2 = 0$. Otherwise, τ_1, τ_2 are positive. Since the wake-up time from the deepest seed-bank that is being monitored is of order $1/e_{L_n} \asymp L_n^\beta$, we have that τ_1, τ_2 are at most of order L_n^β uniformly in $u_1, u_2 \in \mathbb{S}_n^{L_n}$. Return to (6.1). We found in Step 1 in Section 6.1 that the average total joint activity time of the two Markov processes up to time T in the finite system satisfies

$$(6.34) \qquad \mathbb{E}_{u_1,u_2}[I^{(2),n}(T)] \asymp |\mathbb{G}_n|^{-1} \int_{\tau_1 \vee \tau_2}^T ds\, [1 \wedge (s - \tau_1)^{-(1-\gamma)}] [1 \wedge (s - \tau_2)^{-(1-\gamma)}].$$

when $T \gg \psi_n^{1/\gamma}$. Assume without loss of generality that $\tau_1 < \tau_2$. Then

$$(6.35) \qquad \mathbb{E}_{u_1,u_2}[I^{(2),n}(T)] = |\mathbb{G}_n|^{-1} \int_0^{T-\tau_2} ds\, [1 \wedge (s + (\tau_2 - \tau_1))^{-(1-\gamma)}] [1 \wedge s^{-(1-\gamma)}],$$

Pick $T = \bar\beta_n$, and note that $\bar\beta_n \gg \psi_n^{1/\gamma}$ by our assumption that $\psi_n = o((\beta_n^{**})^\gamma)$ because $\bar\beta_n \gg \beta_n^{**}$. Since $\bar\beta_n \gg L_n^\beta$, we have (recall (2.60))

$$(6.36) \qquad \mathbb{E}_{u_1,u_2}[I^{(2),n}(\bar\beta_n)] \asymp |\mathbb{G}_n|^{-1} \int_0^{\bar\beta_n} ds\, [1 \wedge s^{-2(1-\gamma)}] \asymp |\mathbb{G}_n|^{-1} (\bar\beta_n)^{2\gamma-1} = (\bar\beta_n/\beta_n^{**})^{2\gamma-1}$$

for $\gamma \in (\frac{1}{2}, 1]$, which diverges as $n \to \infty$. For $\gamma = \frac{1}{2}$ we have $\mathbb{E}[I^{(2),n}(\bar{\beta}_n)] \asymp |\mathbb{G}_n|^{-1} \log \bar{\beta}_n = \log(\bar{\beta}_n / \beta_n^{**})$, which again diverges as $n \to \infty$. Thus we have proved that the average total joint activity time of the two Markov processes on $\mathbb{S}_n$ diverges as $n \to \infty$. We will complete the proof by showing that the same is true in $\mathbb{P}$-probability. This settles (6.19) because the rate of coalescence c in (6.17) is strictly positive.

6. The process $(\mathcal{E}(s))_{s \geq 0}$ marks the times at which the Markov process on $\{A, (D_m)_{m \in \mathbb{N}_0}\}$ is in A. Since A is a renewal state for this Markov process, we can think of $(\mathcal{E}(s))_{s \geq 0}$ as a *Markov renewal process*. Once more return to (6.1). Let $u_1 = u_2 = (0, A)$. Pick any ϕ_n such that $\psi_n \ll \phi_n \ll \bar{\beta}_n^\gamma$, which is possible because $\psi_n = o((\beta_n^{**})^\gamma)$ and $\bar{\beta}_n \gg \beta_n^{**}$. Put $t_{k,n} = (k\phi_n)^{1/\gamma}$, $k \in \mathbb{N}_0$, and let K_n be such that $t_{K_n,n} \asymp \bar{\beta}_n$, i.e., $K_n \asymp \bar{\beta}_n^\gamma / \phi_n \gg 1$. By the *sample-path* version of (6.7), we have

$$(6.37) \qquad \lim_{n \to \infty} \left\{ \mathcal{L}\left[\left(T(t_{k,n}) - T(t_{k-1,n}), \mathcal{E}(t_{k,n}) \right)_{1 \leq k \leq K_n} \right] - \otimes_{1 \leq k \leq K_n} \left(\delta_{\tau_{k,n}^\gamma - \tau_{k,n}^\gamma}, \left[(1 - \bar{\tau}_{k,n}^{-(1-\gamma)}) \delta_0 + \bar{\tau}_{k,n}^{-(1-\gamma)} \delta_1 \right] \right) \right\}$$
$$= \delta_{0^{\mathbb{N}_0}}$$

for some random variables $\tau_{k,n}, \bar{\tau}_{k,n}$ satisfying $\tau_{k,n} \asymp t_{k,n} \asymp \bar{\tau}_{k,n}$, $1 \leq k \leq K_n$, in $\mathbb{P}_n$-probability. Here we use that

$$(6.38) \qquad t_{k,n} - t_{k-1,n} \gg (t_{k-1,n})^{1-\gamma}$$

to ensure that during each time interval $[t_{k-1,n}, t_{k,n}]$ many renewals to A take place. (The condition in (6.38) is satisfied uniformly in k because $t_{k,n}/t_{k-1,n} \asymp 1$ uniformly in k and $\phi_n \gg 1$.) Moreover, by (2.3),

$$(6.39) \qquad \lim_{n \to \infty} \left\{ \mathcal{L}\left[(RW_1^{n,\uparrow}(T(t_{k,n})))_{1 \leq k \leq K_n} \right] - U(\mathbb{G}_n)^{\otimes K_n} \right\} = \delta_{0^{\mathbb{G} \times \mathbb{N}_0}}$$

with $U(\mathbb{G}_n)$ the uniform distribution on $\mathbb{G}_n$. Here we use that $(t_{k,n})^\gamma - (t_{k-1,n})^\gamma = \phi_n \gg \psi_n$, $1 \leq k \leq K_n$ to ensure that during each time interval $[t_{k-1,n}, t_{k,n}]$ the Markov processes mix well over $\mathbb{G}_n$. It follows via the law of large numbers that

$$(6.40) \qquad I^{(2),n}(\bar{\beta}_n) \asymp |\mathbb{G}_n|^{-1} \sum_{1 \leq k \leq K_n} (t_{k,n} - t_{k-1,n}) (t_{k,n})^{-2(1-\gamma)} \quad \text{in } \mathbb{P}_n\text{-probability.}$$

The sum scales like the integral $\int_0^{\bar{\beta}_n} dt \, [1 \wedge t^{-2(1-\gamma)}] \asymp \bar{\beta}_n^{2\gamma-1}$, and so $I^{(2),n}(\bar{\beta}_n) \asymp \mathbb{E}_{u_1,u_2}[I^{(2),n}(\bar{\beta}_n)]$ in $\mathbb{P}_n$-probability. The same argument works for $u_1, u_2 \in \mathbb{S}_n^{L_n}$ because Markov processes starting at depth $\leq L_n$ become active in time $o(\bar{\beta}_n)$. $\qquad\square$

6.2.2 Switching

In this section we prove the switching part of Theorem 2.12(2).

Pick any time scale $\bar{\beta}_n$ and any seed-bank depths $1 \leq L_n^- < L_n^+ \leq M_n$ satisfying

$$(6.41) \qquad \beta_n^{**} \ll \bar{\beta}_n \ll \beta_n^*, \qquad (L_n^-)^\beta \ll \bar{\beta}_n \ll (L_n^+)^\beta.$$

We know from Section 6.2.1 that partial clustering has occurred to depth L_n^- at time $\bar{\beta}_n$. We show that none of the seed-banks with colour $L_n^+ < m \leq M_n$ has changed up to time $\bar{\beta}_n$.

Consider a Markov process with transition kernel $b^{(2),n}(\cdot, \cdot)$. Because the transitions into and out of the seed-bank do not depend on the location of the Markov process, it suffices to look at the Markov process projected onto $E = \{A, (D_m)_{0 \leq m \leq M_n}\}$. Transitions $A \to D_m$ occur at rate $K_m e_m$, transitions $D_m \to A$ occur at rate e_m. We know that the total activity time up to time T is $\asymp T^\gamma$. (This is the asymptotics for the infinite seed-bank, but as long as $T \ll \beta_n^*$ the probability that in the infinite seed-bank a jump to a colour $> M_n$ occurs up to time T is $o(1)$.) The rate to jump from A to D_m for some $L_n^+ < m \leq M_n$ is (recall (1.39))

$$(6.42) \qquad \sum_{L_n^+ < m \leq M_n} K_m e_m = \sum_{m > L_n^+} K_m e_m - \sum_{m > M_n} K_m e_m \sim \frac{AB}{\gamma\beta} [(L_n^+)^{-\gamma\beta} - (M_n)^{-\gamma\beta}] \asymp (L_n^+)^{-\gamma\beta}.$$

The total rate accumulated up to time T is $\asymp T^\gamma (L_n^+)^{-\gamma\beta}$. For $T = \bar{\beta}_n$ this equals $(\bar{\beta}_n / (L_n^+)^\beta)^\gamma$, which is $o(1)$, and so the claim follows.

Pick $\mu(0) = \mu \in \mathcal{R}_\theta^{(2),\bullet}$. We know from the above observations that

$$(6.43) \qquad \lim_{n \to \infty} \mathcal{L}\left[\Phi_n \mu(\bar{\beta}_n) - (1-\theta) \Phi_n \mu^{0,L_n^-,L_n^+}(\bar{\beta}_n) + \theta \Phi_n \mu_1 1, L_n^-, L_n^+(\bar{\beta}_n) \right] = \delta_{(0,0^{\mathbb{N}_0})^{\mathbb{G}}},$$

where $\Phi_n \mu^{0,L_n^-,L_n^+}(\bar{\beta}_n)$ and $\Phi_n \mu^{1,L_n^-,L_n^+}(\bar{\beta}_n)$ restricted to $\mathbb{G}_n \times \{(D_m)_{L_n^+ < m \leq M_n}\}$ coincide with $\Phi_n \mu(\bar{\beta}_n)$, but restricted to $\mathbb{G}_n \times \{A, (D_m)_{0 \leq m \leq L_n^-}\}$ are 'all 0', respectively, 'all 1', according to an independent draw $\tilde{\Upsilon} \in \{0,1\}$ with mean θ. Clearly, both are elements of $\mathcal{R}_\theta^{(2),\bullet}$, because the seed-banks beyond depth L_n^+ have not changed and $L_n^+ \ll M_n$. Therefore, at any time $\tilde{\beta}_n$ satisfying $\bar{\beta}_n \ll \tilde{\beta}_n \ll \beta_n^*$, partial clustering has occurred in both, to some depth $o(\tilde{\beta}_n^{1/\beta})$ that is larger than L_n^- (and possibly larger than L_n^+ as well), according to an independent draw $\tilde{\Upsilon} \in \{0,1\}$ with mean θ.

6.3 Complete clustering on large time scales

In this section we prove Theorem 2.12(3). The proof again works for $g \in \mathcal{G}$ because it uses the comparison duality exploited in Section 6.2.1. The core is the following lemma, which implies that the macroscopic variable $\hat{\theta}^{M_n,n}(\bar{\beta}_n)$ observed at time $\bar{\beta}_n$ converges to 0 or 1 in probability as $n \to \infty$ when $\bar{\beta}_n \gg \beta_n^*$. Since $t \mapsto \hat{\theta}^{M_n,n}(t)$ is a martingale, the convergence to 0 and 1 occurs at all times $t \geq \bar{\beta}_n$ simultaneously. From (2.43) we see that this implies complete fixation (= complete clustering).

Lemma 6.1. *For any $\mu \in \mathcal{T}$, $g \in \mathcal{G}$ and $\bar{\beta}_n \gg \beta_n^*$,*

$$(6.44) \qquad \lim_{n \to \infty} \mathbb{E}_{\Phi_n \mu} \left[\hat{\theta}^{M_n,n}(\bar{\beta}_n)(1 - \hat{\theta}^{M_n,n}(\bar{\beta}_n)) \right] = 0.$$

Proof. Write (recall (2.48))

$$(6.45) \qquad \mathbb{E}_{\Phi_n \mu} \left[\hat{\theta}^{M_n,n}(t)(1 - \hat{\theta}^{M_n,n}(t)) \right] = \frac{1}{|\mathbb{G}_n|^2} \frac{1}{\kappa_{M_n}} \sum_{u_1,u_2 \in \mathbb{S}_n} K(u_1) K(u_2) \, \mathbb{E}_{\Phi_n \mu}[z_{u_1}(t)(1 - z_{u_2}(t))].$$

The comparison duality in Section 6.2.1 gave us that, for every $t \geq 0$ and $\epsilon > 0$ (recall (6.21)),

$$(6.46) \qquad \mathbb{E}_{\Phi_n \mu}[z_{u_1}^-(t)(1 - z_{u_2}^+(t))] \leq \mathbb{P}_{(2,u_1,u_2)}(B_0(t) = 2),$$

where we recall the definition of the Markov process $(B(t))_{t \geq 0}$ on state space $\{1,2\} \times \mathbb{S}_n \times \mathbb{S}_n$ with transition rates defined in (6.17). Let $p_n(t) = \sup_{u_1,u_2 \in \mathbb{S}_n} \mathbb{P}_{(2,u_1,u_2)}(B_0(t) = 2)$. Then

$$(6.47) \qquad \mathbb{E}_{\Phi_n \mu} \left[\hat{\theta}^{M_n,n}(t)(1 - \hat{\theta}^{M_n,n}(t)) \right] \leq [p_n(t) + \epsilon(2 - \epsilon)] \frac{1}{|\mathbb{G}_n|^2} \frac{1}{\kappa_{M_n}} \sum_{u_1,u_2 \in \mathbb{S}_n} K(u_1) K(u_2) = p_n(t) + \epsilon(2 - \epsilon).$$

Letting $\epsilon \downarrow 0$, we see that it suffices to show that

$$(6.48) \qquad \lim_{n \to \infty} p_n(\bar{\beta}_n) = 0.$$

But this follows from the same type of argument as in the last steps of the proof in Section 6.2.1 with L_n replaced by M_n, where we use that $M_n^\beta = \beta_n^* \ll \bar{\beta}_n$. $\qquad \square$

A Preservation of diffusion function class and reduction of volatility

In this appendix we show that the renormalisation map $\mathcal{F}$ for the infinite system preserves the class $\mathcal{G}$ of proper diffusion functions (recall (1.34)). This is needed for the existence and uniqueness of the scaling limit in the finite-systems scheme. We further show that $\mathcal{F}$ maps the Fisher-Wright diffusion function to a smaller multiple of itself (recall (2.34)). This says that the seed-bank reduces the volatility.

We begin with Model 1, and proceed as in [CGS95]. Let

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

be the kernel for the motion of a lineage in the dual of Model 1. This describes a Markov process on $\mathbb{G}$ with migration kernel $\hat{a}(\cdot,\cdot)$ that becomes dormant (state D) at rate Ke (after which it stops moving) and active (state A) at rate e (after which it starts moving again). Let

$$(A.2) \qquad \hat{B}(i,j) = \int_0^\infty dt \sum_{k \in \mathbb{G}} b_t^{(1)}((k,A),(i,A)) \, b_t^{(1)}((k,A),(j,A)), \qquad i, j \in \mathbb{G},$$

where $b_t^{(1)}(\cdot, \cdot)$ is the time-t transition kernel of the Markov process with transition kernel $b^{(1)}(\cdot, \cdot)$ in (A.1). This is the mean of the *total joint activity time* of two copies of the Markov process starting from (i, A) and (j, A). The coexistence regime corresponds to $\hat{B}(0, 0) < \infty$ (recall Section 1.4).

The following lemma is the seed-bank analogue of [CGS95, Lemma 2.11].

Lemma A.1. [Moments and coupling] *Suppose that $\hat{B}(0, 0) < \infty$. Then*

1. For all $i, j \in \mathbb{G}$,

$$(A.3) \qquad \mathbb{E}_{\nu_\theta}[x_i] = \theta, \qquad \mathbb{E}_{\nu_\theta}[x_i x_j] = \theta^2 + \hat{B}(i, j)\,\mathbb{E}_{\nu_\theta}[g(x_0)].$$

2. For every $\theta, \theta' \in [0, 1]$ with $\theta < \theta'$ there exist a coupling $\mathbb{P}^{\theta,\theta'}$ of the random variables $z = (z_i)_{i \in \mathbb{G}}$ and $z' = (z'_i)_{i \in \mathbb{G}}$ such that $\mathcal{L}[z] = \nu_\theta$, $\mathcal{L}[z'] = \nu_{\theta'}$ and $\mathbb{P}^{\theta,\theta'}(x_i \leq x'_i \,\forall\, i \in \mathbb{G}) = 1$. Consequently, $\mathbb{E}^{\theta,\theta'}[|x_i - x'_i|] = |\theta - \theta'|$ for all $\theta, \theta' \in [0, 1]$.

Proof. Part 1 follows from [GHO22b, Lemma 6.1] by picking the equilibrium measure as the initial measure. Part 2 follows from [GHO22b, Lemma 6.6, Section 6.3.1], where the coupling is achieved by using the same Brownian motions. $\qquad\square$

For Model 2, the analogue of Lemma A.1 holds with $b^{(1)}(\cdot, \cdot)$ replaced by the kernel for the motion of a lineage in the dual of Model 2 (for symmetric migration):

$$(A.4) \qquad b^{(2)}((i, R_i), (j, R_j)) = \begin{cases} \hat{a}(i, j), & R_i = R_j = A, \\ K_m e_m, & i = j, R_i = A, R_j = D_m, m \in \mathbb{N}_0, \\ e_m, & i = j, R_i = D_m, R_j = A, m \in \mathbb{N}_0, \\ 0, & \text{otherwise.} \end{cases}$$

and in the dual of Model 3 (for symmetric migration):

$$(A.5) \qquad b^{(3)}((i, R_i), (j, R_j)) = \begin{cases} \hat{a}(i, j), & R_i = R_j = A, \\ K_m e_m \hat{a}_m(j, i), & R_i = A, R_j = D_m, m \in \mathbb{N}_0, \\ e_m \hat{a}_m(i, j), & R_i = D_m, R_j = A, m \in \mathbb{N}_0. \end{cases}$$

These kernels define appropriate analogues of $\hat{B}(i, j)$ in (A.2).

Our main result in this appendix is the following theorem.

Theorem A.2. [Preservation and reduction: Models 1–3]

(a) *If $g \in \mathcal{G}$, then $\mathcal{F}g \in \mathcal{G}$.*

(b) *If $g = d g_{\mathrm{FW}}$, $d \in (0, \infty)$, then $\mathcal{F}g = d^* g_{\mathrm{FW}}$ with $d^* \in (0, d)$.*

Proof. (a) We must check that $\mathcal{F}g$ satisfies the constraints defining the class $\mathcal{G}$ in (1.34).

- If $\theta \in (0, 1)$, then ν_θ puts positive mass in $(0, 1)$. Since $g(x) > 0$ for all $x \in (0, 1)$, it follows from (2.33) that $(\mathcal{F}g)(\theta) > 0$.

- By (A.3),

$$(A.6) \qquad (\mathcal{F}g)(\theta) = \mathbb{E}_{\nu_\theta}[g(x_0)] = \frac{1}{\hat{B}(0, 0)}\,\mathbb{E}_{\nu_\theta}[(x_0 - \theta)^2].$$

  Thus, if $(\mathcal{F}g)(\theta) = 0$, then by translation invariance we have $\nu_\theta = \delta_\theta$. Since $g(0) = g(1) = 0$, this implies that $(\mathcal{F}g)(0) = (\mathcal{F}g)(1) = 0$.

- For any $\theta, \theta' \in [0, 1]$,

$$(A.7) \qquad \begin{aligned} \left|(\mathcal{F}g)(\theta) - (\mathcal{F}g)(\theta')\right| &= \left|\mathbb{E}^{\theta,\theta'}[g(x_0) - g(x'_0)]\right| \leq \mathbb{E}^{\theta,\theta'}[|g(x_0) - g(x'_0)|] \\ &\leq \mathrm{Lip}[g]\,\mathbb{E}^{\theta,\theta'}[|x_0 - x'_0|] = \mathrm{Lip}[g]\,|\theta - \theta'|, \end{aligned}$$

  with $\mathrm{Lip}[g]$ the global Lipschitz constant of g. Hence $\mathrm{Lip}[\mathcal{F}g] \leq \mathrm{Lip}[g]$. Since g is globally Lipschitz, so is $\mathcal{F}g$.

(b) If $g = d g_{\mathrm{FW}}$, then

$$(A.8) \qquad (\mathcal{F}g)(\theta) = \mathbb{E}_{\nu_\theta}[g(x_0)] = \mathbb{E}_{\nu_\theta}[d x_0(1 - x_0)] = d\theta - d[\theta^2 + \hat{B}(0, 0)(\mathcal{F}g)(\theta)],$$

where we use (A.3). Hence $(\mathcal{F}g)(\theta) = d^* g$ with $d^* = d/(1 + d\hat{B}(0, 0))$. $\qquad\square$

B   General conditions for the finite-systems scheme

In this appendix we recall an *abstract scheme* from [CG94b] that lists general conditions for the validity of the finite-systems scheme. Together with ideas from [DGV95, Section 4], this allows us in Sections 4–5 to prove the finite-systems scheme described in Section 2. Model 2 with $\rho = \infty$ in regime (II) needs a different route, which is followed in Section 6.

Remark B.1. [Conversion of notation for abstract scheme with seed-bank] In what follows we stick to the notation used in [CG94b]. Since the latter was written for systems *without* seed-bank, the reader must everywhere insert the following objects:

- $\rho < \infty$:

$\mathbb{G}, \mathbb{G}_n$	$\rightarrow$	$\mathbb{G}, \mathbb{G}_n$
I	$\rightarrow$	$[0,1] \times [0,1]$ (Model 1)
	$\rightarrow$	$[0,1] \times [0,1]^{\mathbb{N}_0}, [0,1] \times [0,1]^{M_n+1}$ (Model 2–3)
$\mathcal{J}$	$\rightarrow$	$[0,1]$
E	$\rightarrow$	$I^{\mathbb{G}}$
$\mathcal{T}, \mathcal{T}^{\mathrm{erg}}, \mathcal{I}$	$\rightarrow$	$\mathcal{T}, \mathcal{T}^{\mathrm{erg}}, \mathcal{I}$ in (1.24), (1.27)
$\mathcal{T}_\theta^{\mathrm{erg}}$	$\rightarrow$	$\mathcal{T}_\theta^{\mathrm{erg}}$ in (2.30), (2.56)
μ_n	$\rightarrow$	μ_n in (2.9)
$\mathcal{T}_n$	$\rightarrow$	$\mathcal{T}_n$ in (2.11)

- $\rho = \infty$:

$\mathbb{G}, \mathbb{G}_n$	$\rightarrow$	$\mathbb{G} \times (A, D), \mathbb{G}_n \times (A, D)$ (Model 1)
	$\rightarrow$	$\mathbb{G} \times \{(A, (D_m)_{m \in \mathbb{N}_0}\}, \mathbb{G}_n \times \{(A, (D_m)_{0 \leq m \leq M_n}\}$ (Model 2–3)
I	$\rightarrow$	$[0,1]$
$\mathcal{J}$	$\rightarrow$	$[0,1]$
E	$\rightarrow$	$I^{\mathbb{G} \times (A,D)}$ (Model 1)
	$\rightarrow$	$I^{\mathbb{G} \times \{(A, (D_m)_{m \in \mathbb{N}_0}\}}$ (Model 2–3)
$\mathcal{T}, \mathcal{T}^{\mathrm{erg}}, \mathcal{I}$	$\rightarrow$	$\mathcal{T}^\bullet, \mathcal{T}^{\mathrm{erg},\bullet}, \mathcal{I}$ in (1.26), (1.27)
$\mathcal{T}_\theta^{\mathrm{erg}}$	$\rightarrow$	$\mathcal{T}_\theta^{\mathrm{erg},\bullet}, \mathcal{T}^{\mathrm{erg},\bullet,*}, \mathcal{T}_\theta^{\mathrm{erg},\blacksquare}$ in (1.28),(2.59), (3.49)
μ_n	$\rightarrow$	μ_n in (2.9)
$\mathcal{T}_n$	$\rightarrow$	$\mathcal{T}_n$ in (2.11)

With these replacements, the abstract scheme applies to systems *with* seed-bank. For $\rho < \infty$ the seed-bank can be incorporated via an extension of the single-component state space I, and no extension of the geographic space $\mathbb{G}$ is needed. For $\rho = \infty$, however, this is not enough and the geographic space $\mathbb{G}$ needs to be extended with the seed-bank space $\{(A, (D_m)_{m \in \mathbb{N}_0}\}$. This difference once more underpins the fact that for $\rho < \infty$ the effect of the seed-bank is minor, while for $\rho = \infty$ it is not. In the latter case, colour regularity is needed to control the deep seed-banks.   □

General set-up.   In order to be able to formulate the abstract theorem, we need to set up the framework and introduce the relevant notation. Let $\mathbb{G}$ be a countable set, let I be a Polish space, and let $(Z(t))_{t \geq 0}$ be a Markov processes with state space $I^{\mathbb{G}}$ (endowed with the product topology) i.e., $Z(t) = (Z_i(t))_{i \in \mathbb{G}}$, with associated semigroup of transition operators $(S(t))_{t \geq 0}$. Let $(\mathbb{G}_n)_{n \in \mathbb{N}}$ be a projective systems of discrete finite groups that increases to $\mathbb{G}$, and let $(Z_n(t))_{t \geq 0}$, $n \in \mathbb{N}$, be Markov processes with state space $E \subseteq I^{\mathbb{G}_n}$ and semigroup $(S_n(t))_{t \geq 0}$, $n \in \mathbb{N}$. For $Z \in I^{\mathbb{G}}$, define the element $Z|_{\mathbb{G}_n} \in I^{\mathbb{G}_n}$ by the restriction $(Z|_{\mathbb{G}_n})_i = Z_i, i \in \mathbb{G}_n$. Suppose that there are extension operators $\tilde{\phi}_n \colon I^{\mathbb{G}_n} \to I^{\mathbb{G}}$, $n \in \mathbb{N}$, satisfying $(\tilde{\phi}_n Z_n)|_{\mathbb{G}_n} = Z_n$, and write $\tilde{\phi}_n$ also for the corresponding push forward acting on measures.

If μ is a measure on $I^{\mathbb{G}}$ and μ_n is a measure on $I^{\mathbb{G}_n}$, $n \in \mathbb{N}$, then write $\lim_{n \to \infty} \mu_n = \mu$ as $\lim_{n \to \infty} \tilde{\phi}_n \mu_n = \mu$. A function $f \colon I^{\mathbb{G}} \to \mathbb{R}$ is called tame when it is bounded, continuous and depends on only *finitely many* coordinates. Write $\lim_{n \to \infty} \mathcal{L}[\mu_n - \nu_n] = 0$ as $n \to \infty$ when $\lim_{n \to \infty} [\langle \mu_n, f \rangle - \langle \nu_n, f \rangle] = 0$ for all tame f.

We need four groups of assumptions.

I: The processes $Z_n(\cdot)$, $n \in \mathbb{N}$, are finite versions of $Z(\cdot)$.   This first set of assumptions makes precise the statement that $Z_n(\cdot)$, $n \in \mathbb{N}$, can be considered *finite* versions of $Z(\cdot)$, and identifies certain classes of appropriate initial distributions $\mathcal{T}$ and $\mathcal{T}_n$ for $Z(t)$ and $Z_n(t)$.

(A1) For each $Z(0) \in I^{\mathbb{G}}$, with $Z_n(0) = Z(0)|_{\mathbb{G}_n}$,

(B.1) $\lim_{n\to\infty} \mathcal{L}[Z_n(t)] = \mathcal{L}[Z(t)]$, $\forall\, t \geq 0$.

(A2) There are sets of measures $\mathcal{T}$ and $\mathcal{T}_n$ on $I^{\mathbb{G}}$, respectively, $I^{\mathbb{G}_n}$, such that $\tilde{\phi}_n(\mathcal{T}_n) \subset \mathcal{T}$, and these sets are closed under the actions of the associate semigroups, i.e., for all $t \geq 0$, $\mu_n \in \mathcal{T}_n$ implies $\mu_n S_n(t) \in \mathcal{T}_n$ and $\mu \in \mathcal{T}$ implies $\mu S(t) \in \mathcal{T}$.

II: Basic ergodic properties of the infinite system.   The second set of assumptions are concerned with the equilibrium states and the ergodic properties of the infinite system. Let $\mathcal{I}$ denote the set of invariant measures for $Z(\cdot)$, i.e., all $\mu \in \mathcal{T}^{\mathrm{erg}}$ such that $\mu S(t) = \mu$ for all $t \geq 0$.

(A3) There is a Polish space $\mathcal{J}$ and a set of probability measures $\{\nu_\theta \colon \theta \in \mathcal{J}\}$ on $I^{\mathbb{G}}$ such that $\{\nu_\theta \colon , \theta \in \mathcal{J}\} \subset \mathcal{I}$. For $\theta \in \mathcal{J}$, let $\mathcal{T}_\theta^{\mathrm{erg}}$ be the set of all $\mu \in \mathcal{T}^{\mathrm{erg}}$ such that $\lim_{t\to\infty} \mu S(t) = \nu_\theta$. (Typically, $\{\nu_\theta \colon \theta \in \mathcal{J}\}$ is the set of *extreme points* of $\mathcal{I} \cap \mathcal{T}^{\mathrm{erg}}$.)

(A4) There is a random variable Θ on $I^{\mathbb{G}}$, called the *macroscopic variable*, and functions $\hat{\theta}^m \colon I^{\mathbb{G}_n} \to \mathcal{J}$ that are estimators of Θ, such that for all $\mu \in \mathcal{T}^{\mathrm{erg}}$ the following limit exists (saying that $\hat{\theta}$ is a consistent estimator of θ):

(B.2) $\lim_{n\to\infty} \hat{\theta}^n(Z|_{\mathbb{G}_n}) = \Theta(Z)$,   in law under μ.

Furthermore, if $\Theta(Z) = \theta$ μ-a.s., then $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$ and, for all $\mu \in \mathcal{T}$,

(B.3) $\mu = \displaystyle\int_{\mathcal{J}} \Gamma(\mathrm{d}\theta)\, \mu_\theta$,

where Γ is the law of $\Theta(Z)$ with respect to μ, and $\mu_\theta \in \mathcal{T}_\theta^{\mathrm{erg}}$, $\theta \in \mathcal{J}$.

III: Time scale.   In the third set of assumptions we impose certain regularity properties on $\hat{\theta}_n$, and identify a fundamental *time scale* β_n connected with $\hat{\theta}^n$. We write $\bar{\theta}^n(t)$ for $\hat{\theta}^n(Z_n(t))$.

(A5) Let $(T_n)_{n\in\mathbb{N}}$ be any sequence tending to infinity. Suppose that $\mathcal{L}[Z_n(0)] \in \mathcal{T}_n$, and $\mathcal{L}[Z(0)] \in \mathcal{T}_\theta^{\mathrm{erg}}$ for some $\theta \in \mathcal{J}$. Suppose further that, along some subsequence $(n_k)_{k\in\mathbb{N}}$,

(B.4) $\lim_{k\to\infty} \mathcal{L}[Z_{n_k}(T_{n_k})] = \mathcal{L}[Z]$.

Then

(B.5) $\lim_{k\to\infty} \mathcal{L}\left[\hat{\theta}^{n_k}(Z_{n_k}(T_{n_k}))\right] = \mathcal{L}[\Theta(Z)]$.

(A6) There is a sequence $(\beta_n)_{n\in\mathbb{N}}$ tending to infinity such that if $\mathcal{L}[Z_n(0)] \in \mathcal{T}_n$ and $\lim_{n\to\infty} \mathcal{L}[Z_n(0)] = \mathcal{L}[Z(0)] \in \mathcal{T}^{\mathrm{erg}}$, then

  – the family of measures $\mathcal{L}[\hat{\theta}^n(s\beta_n)_{s\geq 0}]$ on $D(\mathcal{J}, [0, \infty))$ is *tight*,
  – all weak limit points $(\Theta(s))_{s\geq 0}$ have the property that $s \mapsto E[f(\Theta(s))]$ is continuous for all bounded continuous functions f on $\mathcal{J}$.

We want that β_n is proper, i.e., leads to a non-trivial limit process. This has to be checked afterwards.

IV: Compactness and strengthened convergence properties.   In the fourth set of assumptions we impose compactness and require a strengthened form of (A1), as well as a *strengthened Feller property* for $(Z(t))_{t\geq 0}$.

(A7) Fix $T < \infty$, and let $(t_n)_{n\in\mathbb{N}}$ satisfy $0 \leq t_n \leq T\beta_n$ for $n \in \mathbb{N}$. If $\mathcal{L}[Z_n(0)] \in \mathcal{T}^n$, then the family $\mathcal{L}(Z_n(t_n))$, $n \in \mathbb{N}$, is tight and all its weak limit points belong to $\mathcal{T}$.

(A8) If $\mu, \mu_n \in \mathcal{T}^{\mathrm{erg}}$ are such that $\lim_{n\to\infty} \mu_n = \mu$ and $\lim_{n\to\infty} t_n = \infty$, then

(B.6) $\lim_{n\to\infty} [\mu_n S(t_n) - \mu S(t_n)] = 0$.

(A9) There is a sequence $(L_n)_{n \in \mathbb{N}}$ tending to infinity such that if $0 \le t_n \le L_n$ and $\mu_n \in \mathcal{T}_n$, then

(B.7) $\qquad \lim_{n \to \infty} [\mu_n S_n(t_n) - (\tilde{\phi}_n \mu_n) S(t_n)] = 0.$

Assumptions (A8) and (A9) should be viewed as statements about uniformity of the ergodic theorem in the initial measure and uniformity in the approximation of the infinite system by finite systems for large times. The most difficult assumptions to check are (A6), (A8) and (A9), and (A10) below.

Results. We are now ready to state our abstract theorem.

Theorem B.2. [Abstract theorem] *Suppose that $Z(t), Z_n(t)$, $n \in \mathbb{N}$, satisfy* (A1)–(A9), *and that* $\mathcal{L}(Z(0)) \in \mathcal{T}_\theta$ *for some* $\theta \in \mathcal{J}$. *Let* $(n_k)_{k \in \mathbb{N}}$ *be any sequence such that*

(B.8) $\qquad \lim_{k \to \infty} \mathcal{L}\left[(\hat{\theta}^{n_k}(s\beta_{n_k}))_{s \ge 0}\right] = \mathcal{L}[(\Theta(s))_{s \ge 0}] \text{ in } D(\mathcal{J}, [0, \infty)).$

For any $s \in (0, \infty)$, *if* $T(n) \to \infty$ *and* $T(n)/\beta_n \to s$ *as* $n \to \infty$, *then*

(B.9) $\qquad \lim_{k \to \infty} \mathcal{L}[Z_{n_k}(sT(n_k))] = \int_{\mathcal{J}} P(\Theta(s) \in d\theta) \, \nu_\theta,$

where P is the law of $(\Theta(s))_{s \ge 0}$.

Note that, because of (A6), sequences satisfying (B.8) exist. Suppose that we add the assumption:

(A10) The weak limit in (B.8) is unique.

Then the finite-systems scheme holds as claimed. The way we are able to check (A10) is as follows.

Corollary B.3. [Verification of (A10)] *If a weak limit point of* $\mathcal{L}[(\hat{\theta}^{n_k}(s\beta_{n_k}))_{s \ge 0}]$ *as* $k \to \infty$ *satisfies a well-posed martingale problem, independently of* $(n_k)_{k \in \mathbb{N}}$, *then* (B.9) *holds and* $\mathcal{L}[(\hat{\theta}^n(s\beta_n))_{s \ge 0}]$ *converges as* $n \to \infty$ *to the solution of this martingale problem.*

Once we have checked (A1)–(A10), we will want to verify that the macroscopic time scale β_n is proper, i.e., at time $s\beta_n$ the finite system locally converges to $\nu_{\Theta(s)}$, the equilibrium of the infinite system with density $\Theta(s)$, where $(\Theta(s))_{s \ge 0}$ is a non-trivial random process with cadlag paths hitting the traps of the evolution as $s \to \infty$, which in our models are 0 or 1.

C Fraction of time spent in the active state

In this appendix we show that, in the limit as $M \to \infty$, the fraction of time spent in the active state until time M^β is $\asymp f_M$ with

(C.1) $\qquad f_M = \dfrac{1}{1 + \sum_{0 \le m \le M} K_m}.$

Because the transitions into and out of the seed-bank do not depend on the location of the Markov process, it suffices to look at the Markov process projected onto $E = \{A, (D_m)_{\mathbb{N}_0}\}$. Transitions from dormant states with colour $> M$ occur at rate $O(M^{-\beta})$ as $M \to \infty$. Hence these transitions may be ignored, at the cost of a finite correction factor, so that we only need to look at the projection onto $C_M = \{A, (D_m)_{0 \le m \le M}\}$.

Let $(U_k)_{k \in \mathbb{N}}$ be the discrete-time Markov chain on C_M with transition probabilities 1 for $D_m \to A$ and p_m for $A \to D_m$ with

(C.2) $\qquad p_m = \dfrac{K_m e_m}{\sum_{0 \le \ell \le M} K_\ell e_\ell}, \qquad 0 \le m \le M.$

The transition times are equal in distribution to

(C.3) $\qquad A \to (D_m)_{0 \le m \le M} : \dfrac{1}{\sum_{0 \le m \le M} K_m e_m} Z, \qquad D_m \to A : \dfrac{1}{e_m} Z',$

with Z, Z' independent Exp(1) random variables. For $n \in \mathbb{N}$, let T_n^A and T_n^D denote the *total* active time and dormant time after n transitions out of and into A, when the Markov process starts in A. Then, in distribution,

$$(C.4) \qquad T_n^A = \sum_{k=0}^{n-1} \frac{1}{\sum_{0 \le m \le M} K_m e_m} Z_k, \qquad T_n^D = \sum_{k=0}^{n-1} \frac{1}{e_{M_k}} Z_k',$$

where $(Z_k)_{k \in \mathbb{N}_0}, (Z_k')_{k \in \mathbb{N}_0}$ are sequences of independent Exp(1) random variables. Since C_M is finite, by (C.2) we have

$$(C.5) \qquad \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{U_k} = \sum_{0 \le m \le M} p_m \delta_m \quad \text{in probability.}$$

Hence

$$(C.6) \qquad \lim_{n \to \infty} \frac{1}{n} T_n^A = \mu^A, \qquad \lim_{n \to \infty} \frac{1}{n} T_n^D = \mu^D \quad \text{in probability,}$$

with

$$(C.7) \qquad \mu^A = \frac{1}{\sum_{0 \le m \le M} K_m e_m}, \qquad \mu^D = \frac{1}{\sum_{0 \le m \le M} K_m e_m} \sum_{0 \le m \le M} K_m.$$

Note that

$$(C.8) \qquad \mu^D = \mu^A \sum_{0 \le m \le M} K_m.$$

Consequently,

$$(C.9) \qquad \lim_{n \to \infty} \frac{T_n^A}{T_n^A + T_n^D} = \frac{1}{1 + \sum_{0 \le m \le M} K_m} = f_M \quad \text{in probability.}$$

Hence a fraction f_M of time is spent in the active state, where we note that until time M^β *many* transitions $A \to D_m$ occur for *most* $0 \le m \le M$, so that the law of large numbers in (C.5) applies, at the cost of a finite correction factor.

D Speculations about clustering for fast growing seed-bank

In this appendix we speculate about how partial and complete clustering may come about in Model 2 with $\rho = \infty$ in regime (II). We will argue that both on time scale β_n^{**} and time scale β_n^* the macroscopic variable moves towards fixation according to a *jump process*, i.e., it follows a piecewise constant path that ends in 0 or 1. This is different from the diffusive clustering found in Model 1, Model 2 with $\rho < \infty$ and Model 2 with $\rho = \infty$ in regime (I) (see Fig. 5).

D.1 Partial clustering

Step 1: Hazard. Let $I(t) \in \{0, 1\}$ be the indicator of the event that two independent Markov processes with transition kernel $b^{(2),n}(\cdot, \cdot)$, starting at a distance of order n, at time t occupy the same colony in $\mathbb{G}_n$ and are jointly active. Let

$$(D.1) \qquad \zeta_0 = 0, \quad \eta_0 = 0, \quad \zeta_k = \inf\{t > \eta_{k-1} : I(t) = 1\}, \quad \eta_k = \inf\{t > \zeta_k : I(t) = 0\}, \qquad k \in \mathbb{N},$$

be the successive times at which joint active occupancy is switched on and off. The successive intervals of joint active occupancy are $(\Delta_k)_{k \in \mathbb{N}}$ with $\Delta_k = [\zeta_k, \eta_k)$, and so $I(t) = 1$ if and only if $t \in \cup_{k \in \mathbb{N}} \Delta_k$. The number of completed intervals up to time t is

$$(D.2) \qquad K(t) = \sup\{k \in \mathbb{N} : \eta_k \le t\}.$$

We are interested in the *total joint active occupancy time* up to time t,

$$(D.3) \qquad C(t) = \sum_{k=1}^{K(t)} |\Delta_k| + (t - \zeta_{K(t)+1}) \, 1_{\{\zeta_{K(t)+1} < t\}}.$$

This is the total time up to time t during which the two Markov processes coalesce at rate d. For $t \to \infty$ the boundary term is negligible (because $(t - \zeta_{K(t)+1}) \, 1_{\{\zeta_{K(t)+1} < t\}} \le |\Delta_{K(t)+1}|$). Therefore, the scaling of $C(t)$ is determined by the scaling of $K(t)$ and the law of $(|\Delta_k|)_{k \in \mathbb{N}}$. However, in order to find the scaling of $C(t)$ it is better to look at a *larger time scale*. Indeed, what happens is that the two Markov processes experience *bursts* of joint active occupancy times until they separate and meet again. These joint activity time intervals come in rapid succession, until the two Markov processes move far apart and loose track of each other. After that they both need to make order $|\mathbb{G}_n|$ steps until they meet again. In view of this scale separation, we divide time into two parts: the long stretches *between* the burst and the short stretches *during* the burst.

Step 2: Bursts.   We begin by focussing on the *long stretches*. Let us discretise time and consider a renewal process $\zeta = (0, \zeta_1, \zeta_1 + \zeta_2, \ldots)$ on $\mathbb{N}_0$ with i.i.d. increments such that $\mathbb{P}(\zeta_1 < \infty) = 1$ and $\mathbb{P}(\zeta_1 = n) \sim C n^{-(1+\gamma)}$, $n \to \infty$. Let ζ' be an independent copy of ζ, and put $\mathcal{G} = \zeta \cap \zeta'$. Then also $\mathcal{G}$ is a renewal process on $\mathbb{N}_0$, written $\mathcal{G} = (0, \mathcal{G}_1, \mathcal{G}_1 + \mathcal{G}_2, \ldots)$. It was shown in [AB16, Theorems 1.3 and 1.5] that, for $\gamma \in (\frac{1}{2}, 1)$,

$$(D.4) \qquad \mathbb{P}(\mathcal{G}_1 < \infty) = 1, \qquad \mathbb{P}(\mathcal{G}_1 = n) \sim C^* n^{-(1+\gamma^*)}, \quad n \to \infty,$$

with

$$(D.5) \qquad \gamma^* = 2\gamma - 1, \qquad C^* = \pi C^2 \left[\gamma^* \sin(\pi \gamma^*) \right] \left[\gamma \sin(\pi \gamma) \right]^{-2}.$$

Returning to continuous time, we see that (D.4)–(D.5) tell us that if we *ignore* the time lapses during which the two Markov processes *are active* (which are negligible because they have finite mean), then the time lapses $U_1, U_2, \ldots$ between the successive times at which they *become jointly active* are i.i.d. with

$$(D.6) \qquad \mathbb{P}(U_1 < \infty) = 1, \qquad \frac{\mathbb{P}(U_1 \in \mathrm{d}t)}{\mathrm{d}t} \sim C^* t^{-(1+\gamma^*)}, \quad t \to \infty.$$

However, by (2.3), on average at only one out of $|\mathbb{G}_n|$ such times the two Markov processes occupy the same site. Hence, the time lapses $V_1, V_2, \ldots$ between the successive times at which they *become jointly active at the same colony* are i.i.d. with

$$(D.7) \qquad \mathbb{P}(V_1 \in \mathrm{d}t) = \sum_{k \in \mathbb{N}} \frac{1}{|\mathbb{G}_n|} \left[1 - \frac{1}{|\mathbb{G}_n|} \right]^{k-1} \mathbb{P}\left(\sum_{\ell=1}^{k} U_\ell \in \mathrm{d}t \right).$$

Putting $t = s|\mathbb{G}_n|^{1/\gamma^*}$, $k = \bar{s}|\mathbb{G}_n|$ and letting $n \to \infty$, we obtain

$$(D.8) \qquad \lim_{n \to \infty} \frac{\mathbb{P}(|\mathbb{G}_n|^{-1/\gamma^*} V_1 \in \mathrm{d}s)}{\mathrm{d}s} = \int_0^\infty \mathrm{d}\bar{s} \, e^{-\bar{s}} \lim_{n \to \infty} \frac{\mathbb{P}(|\mathbb{G}_n|^{-1/\gamma^*} \sum_{\ell=1}^{\bar{s}|\mathbb{G}_n|} U_\ell \in \mathrm{d}s)}{\mathrm{d}s}.$$

It follows from (D.6) that

$$(D.9) \qquad \lim_{N \to \infty} \mathcal{L}\left[N^{-1/\gamma^*} \sum_{\ell=1}^{N} U_\ell \right] = \mathcal{L}[U^*]$$

with U^* a positive stable law random variable with exponent γ^*. Hence, putting $T_1^n = |\mathbb{G}_n|^{-1/\gamma^*} V_1 = V_1 / \beta_n^{**}$, we find that

$$(D.10) \qquad \lim_{n \to \infty} \frac{\mathbb{P}(T_1^n \in \mathrm{d}s)}{\mathrm{d}s} = \int_0^\infty \mathrm{d}\bar{s} \, e^{-\bar{s}} \frac{\mathbb{P}(\bar{s}^{1/\gamma^*} U^* \in ds)}{\mathrm{d}s}.$$

This says that T_1^n converges in distribution to a random variable T_1 whose density is given by the right-hand side of (D.10). In terms of the Laplace transform, we get

$$\mathbb{E}[e^{-\lambda T_1}] = \int_0^\infty \mathrm{d}s \, e^{-\lambda s} \, \mathbb{P}(T_1 \in \mathrm{d}s) = \int_0^\infty \mathrm{d}s \, e^{-\lambda s} \int_0^\infty \mathrm{d}\bar{s} \, e^{-\bar{s}} \, \mathbb{P}(\bar{s}^{1/\gamma^*} U^* \in ds)$$

$$(D.11) \qquad = \int_0^\infty \mathrm{d}\bar{s} \, e^{-\bar{s}} \, \mathbb{E}\left[e^{-\lambda \bar{s}^{1/\gamma^*} U^*} \right] = \int_0^\infty \mathrm{d}\bar{s} \, e^{-\bar{s}} \, e^{-D(\lambda \bar{s}^{1/\gamma^*})^{\gamma^*}}$$

$$= \int_0^\infty \mathrm{d}\bar{s} \, e^{-\bar{s}} \, e^{-\bar{s} D \lambda^{\gamma^*}} = (1 + D \lambda^{\gamma^*})^{-1}, \qquad \lambda > 0,$$

for some $D \in (0, \infty)$ depending on C^*.

Step 3: Short stretches.   Next we turn to the *short stretches*. In the limit as $n \to \infty$, the total joint active occupancy times during the successive bursts, which we denote by $W_1, W_2, \ldots$, are equal in distribution to the total joint active occupancy time of two Markov processes on $\mathbb{G}$ starting from the origin. Indeed, for a burst to end the two Markov processes must move apart and loose track of each other (as they do on $\mathbb{G}$). It is only after they meet again (as they do on $\mathbb{G}_n$) and are jointly active that the next burst starts, which takes a much longer time than the typical time of a burst. Thus, the law of W_1 is the same as that of the *total joint active occupancy times* we analyzed in [GHO22b, Section 6.2]. In particular, $\mathbb{E}[W_1] = I$ with I the integral in [GHO22b, Eq. (5.19)], which by [GHO22b, Eq. (6.25)] is finite because $I_{\hat{a}, \gamma} < \infty$ in the regime of coexistence (recall (1.42)).
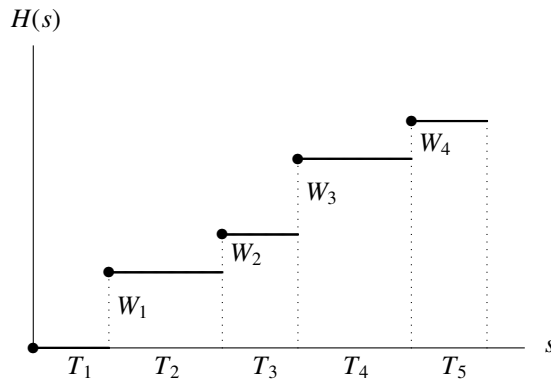

Figure 7: Accumulation of the joint activity time $s \mapsto H(s)$ of two Markov processes in the finite system on time scale β_n^{**} in the limit as $n \to \infty$. The time lapses $T_1, T_2, \ldots$ are i.i.d. with law given by (D.11), the increments $W_1, W_2, \ldots$ are i.i.d. with law given by the total joint activity time of two Markov processes in the infinite system.

Step 4: Towards partial clustering.   In view of Steps 1–3 we expect that, for $L_n = o((\beta_n^{**})^{1/\beta})$,

$$(D.12) \qquad \lim_{n \to \infty} \mathcal{L}[(\hat{\theta}^{L_n, n}(s\beta_n^{**}))_{s>0}] = (\Theta^*(s))_{s>0}$$

with

$$(D.13) \qquad \Theta^*(s) = \Theta(H(s)), \qquad s > 0,$$

where $(\Theta(s))_{s>0}$ is the diffusion on $[0, 1]$ defined in (2.32), with diffusion function $\mathcal{F} g$. Thus, partial clustering is achieved via i.i.d. random jumps $W_1, W_2, \ldots$ occurring after i.i.d. random time lapses $T_1, T_2, \ldots$.

D.2   Complete clustering

We expect the exact same behaviour as in (D.12) for $(\hat{\theta}^{M_n, n}(s\beta_n^*))_{s>0}$ on time scale β_n^*, modulo a constant that multiplies $T_1, T_2, \ldots$ and depends on the parameters controlling the deep seed-banks (recall in (1.39)). The hazard at time s is again growing in a bursty manner, now for Markov processes that start from the deepest seed-banks.

References

[AB16]   K. Alexander and Q. Berger. Local asymptotics for the first intersection of two independent renewals. *Electr. J. Probab.*, 21(66):1—20, 2016.

[BR76]   R.N. Bhattacharya and R. Ranga Rao. *Normal Approximation and Asymptotic Expansions.* Wiley & Sons, New York, 1976.

[Bre68]   L. Breiman. *Probability*. Addison-Wesley, Reading, Massachusetts, 1968.

[CG90]   J.T. Cox and A. Greven. On the long term behaviour of some finite particle systems. *Probab. Theory Relat. Fields*, 85:195–237, 1990.

[CG91]   J.T. Cox and A. Greven. On the long term behavior of finite particle systems:a critical dimension example. In *Random Walks, Brownian Motion, and Interacting Particle Systems*, volume 28 of *Progr. Probab.*, pages 203–213. Birkhäuser Boston, Boston, MA, 1991.

[CG94a]   J.T. Cox and A. Greven. Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.*, 22(2):833–853, 1994.

[CG94b]   J.T. Cox and A. Greven. The finite systems scheme: An abstract theorem and a new example. *CRM Proceedings and Lecture Notes*, 5:55–66, 1994.

[CGS95]   J.T. Cox, A. Greven, and T. Shiga. Finite and infinite systems of interacting diffusions. *Probab. Theory Relat. Fields*, 102:165–197, 1995.

[Cox89]   J.T. Cox. Coalescing random walks and voter model consensus times on the torus in zd. *Ann. Probab.*, 17(4):1333–1366, 1989.

[DG93]   D.A. Dawson and A. Greven. Multiple scale analysis of interacting diffusions. *Probab. Th. Relat. Fields*, 95:467–508, 1993.

[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(7):2277–2360, 1995.

[DGW04]   D.A. Dawson, L.G. Gorostiza, and A. Wakolbinger. Hierarchical random walks. In *Asymptotic methods in stochastics*, volume 44 of *Fields Inst. Commun.*, pages 173–193. Amer. Math. Soc., Providence, RI, 2004.

[DGW05]   D.A. Dawson, L. Gorostiza, and A. Wakolbinger. Degrees of transience and recurrence and hierarchical random walk. *Potential Anal.*, 22(4):305–350, 2005.

[EK86]   S.N. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. John Wiley, New York, 1986.

[GHO22a] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: renormalisation on the hierarchical group. *To appear in Memoirs Amer. Math. Soc.*, 2022.

[GHO22b] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: well-posedness, duality and equilibrium. *Electron. J. Probab.*, 27:(paper no. 18),1–88, 2022.

[Hug95]   B.D. Hughes. *Random Walk in Random Environment*, volume I. Clarendon Press, 1995.

[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.

[Kre85]   U. Krengel. *Ergodic Theorems*, volume 6 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1985.

[Lig85]   T.M. Liggett. *Interacting Particle Systems*. Springer, New York, 1985.

[OK73]   T. Ohta and M Kimura. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet. Res.*, 22:201–204, 1973.

[Spi64]   F. Spitzer. *Principles of Random Walk*, volume 34. Springer New York, New York, 1964.