Spatial populations with seed-bank: well-posedness, duality and equilibrium

Greven, A.; Hollander, W.T.F. den; Oomen, M.

Spatial populations with seed-bank:
well-posedness, duality and equilibrium[*]

Andreas Greven[†]   Frank den Hollander[‡]   Margriet Oomen[§]

Abstract

We consider a system of interacting Fisher-Wright diffusions with seed-bank. Individuals 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 labeling the colonies we consider a countable Abelian group $\mathbb{G}$ endowed with the discrete topology. The key example of interest is the Euclidean lattice $\mathbb{G} = \mathbb{Z}^d$, $d \in \mathbb{N}$. Our goal is to *classify* the long-time behaviour of the system in terms of the underlying model parameters. In particular, we want to understand in what way the seed-bank enhances genetic diversity. We introduce three models of increasing generality, namely, individuals become dormant: (1) in the seed-bank of their colony; (2) in the seed-bank of their colony while adopting a *random colour* that determines their wake-up time; (3) in the seed-bank of a *random colony* while adopting a *random colour*. The extension in (2) allows us to model wake-up times with fat tails while preserving the Markov property of the evolution. The extension in (3) allows us to place individuals in different colony when they become dormant. For each of the three models we show that the system of continuum stochastic differential equations, describing the population in the large-colony-size limit, has a unique strong solution. We also show that the system converges to a unique equilibrium depending on a single *density parameter* that is determined by the initial state, and exhibits a dichotomy of *coexistence* (= locally multi-type equilibrium) versus *clustering* (= locally mono-type equilibrium) depending on the parameters controlling the migration and the seed-bank. The seed-bank slows down the loss of genetic diversity. In model

(1), the dichotomy between clustering and coexistence is determined by migration only. In particular, clustering occurs for recurrent migration and coexistence occurs for transient migration, as for the system without seed-bank. In models (2) and (3), an interesting interplay between migration and seed-bank occurs. In particular, the dichotomy is affected by the seed-bank when the wake-up time has infinite mean. For instance, for *critically recurrent migration* the system exhibits clustering for finite mean wake-up time and coexistence for infinite mean wake-up time. Hence, at the *critical dimension* for the system without seed-bank, *new universality classes* appear when the seed-bank is added. If the wake-up time has a sufficiently fat tail, then the seed-bank determines the dichotomy and migration has no effect at all. The presence of the seed-bank makes the proof of convergence to a unique equilibrium a conceptually delicate issue. By combining duality arguments with coupling techniques, we show that our results also hold when we replace the Fisher-Wright diffusion function by a more general diffusion function, drawn from an appropriate class.

Contents

1 Background and outline

1.1 Background and goals

In populations with a seed-bank, individuals can become dormant and stop reproducing themselves, until they become active and start reproducing themselves again. In [BGCEK15] and [BGCKWB16], the evolution of a population evolving according to the Fisher-Wright model with a *seed-bank* was studied. In this model individuals are subject to *resampling* and can move in and out of a seed-bank. While in the seed-bank they suspend resampling, i.e., the seed-bank acts as a repository for the genetic information of the population. Individuals that do not reside in the seed-bank are called *active*, those that do are called *dormant*. In the present paper we extend the single-colony Fisher-Wright model with seed-bank introduced in [BGCKWB16] to a multi-colony setting in which individuals live in different colonies and move between colonies. In other words, we introduce *spatialness*.

Seed-banks are observed in many taxa, including plants, bacteria and other microorganisms. Typically, they arise as a response to unfavourable environmental conditions. The dormant state of an individual is characterised by low metabolic activity and interruption of phenotypic development (see e.g. Lennon and Jones [LJ11]). After a varying and possibly large number of generations, dormant individuals can be resuscitated under more favourable conditions and reprise reproduction after having become active. This strategy is known to have important implications for population persistence, maintenance of genetic variability and stability of ecosystems. It acts as a *buffer* against evolutionary forces such as genetic drift, selection and environmental variability.

The importance of this evolutionary trait has led to several attempts to model seed-banks from a mathematical perspective, see e.g. [KKL01], [BGCKS13], [GCAvWE+14], [BBGCWB19]. In [BGCKWB16] it was shown that the continuum model obtained by taking the large-colony-size limit of the individual-based model with seed-bank is the Fisher-Wright diffusion with seed-bank. Also the long-time behaviour and the genealogy of the continuum model with seed-bank were analysed in [BGCKWB16].

In the present paper we consider a *spatial* version of the continuum model with seed-bank, in which individuals live in colonies, each with their own seed-bank, and are allowed to *migrate* between colonies. Our goal is to understand the change in behaviour compared to the spatial model without seed-bank. The latter has been the object of intense study. A sample of relevant papers and overviews is [Shi80], [Daw93], [DG93], [DGV95], [DG96], [EF96], [HS98], [Hol06], [DGdH+08], [GdHKK14]. We expect the presence of the seed-bank to affect the long-time behaviour of the system not only quantitatively but also qualitatively. To understand how this comes about, we must find ways to deal with the *richer behaviour* of the population caused by the motion in and out of the seed-bank. Earlier work on a spatial model with seed-bank, migration and mutation was carried out in [dHP17], where the probability to be identical by decent for two individuals drawn randomly from two colonies was computed as a function of the distance between the colonies.

It has been recognised that qualitatively different behaviour may occur when the wake-up time in the seed-bank changes from having a thin tail to having a fat tail [LJ11]. One challenge in modelling seed-banks has been that fat tails destroy the Markov property for the evolution of the system. A key idea of the present paper is that we can *enrich the seed-bank with internal states* – which we call colours – to allow for fat tails and still preserve the Markov property for the evolution. We will see that fat tails induce *new universality classes*.

The main goals of the present paper are the following:

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

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

    (2a) Show that if the wake-up time has *finite mean*, then the dichotomy between coexistence and clustering is controlled by the migration only and the seed-bank has no effect. In particular, clustering prevails when the symmetrised migration kernel is recurrent while coexistence prevails when it is transient. This is the classical dichotomy for populations without seed-bank [CG94]. The critical dimension is $d = 2$.

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

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

We focus on the situation where the individuals can be of two types. The extension to infinitely many types, called the Fleming-Viot measure-valued diffusion, only requires standard adaptations and will not be considered here (see [DGV95]). Also, instead of Fisher-Wright resampling we will allow for state-dependent resampling, i.e., the rate of resampling in a colony depends on the fractions of the two types in that colony. In what follows we only work with *continuum* models, in which the components represent *type frequencies* in the colonies labelled by a discrete geographic space.

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

1.2 Outline

In Section 2 we introduce three models of increasing generality, establish their well-posedness via a martingale problem, and introduce their dual processes, which play a crucial role in the analysis. In Section 3 we state our main results. We focus on the long-time behaviour, prove convergence to equilibrium, and establish a *dichotomy* between clustering and coexistence. We show that this dichotomy is *affected by the presence of the seed-bank*, namely, the dichotomy depends not only on the migration rates, but can also depend on the relative sizes of the active and the dormant population and their rates of exchange. In particular, if the dormant population is much larger than the active population, then the residence time in the seed-bank has a fat tail that *enhances genetic diversity significantly*.

Sections 4–7 are devoted to the proofs of the theorems stated in Sections 2–3. In Appendix A we give the derivation of the single-colony continuum model from the single-colony individual-based Fisher-Wright model in the large-colony-size limit. In the individual-based model active individuals *exchange* with dormant individuals, i.e., for each active individual that becomes dormant a dormant individual becomes active. In Appendix B we look at the continuum limit of the single-colony individual-based Moran model in which active and dormant individuals no longer exchange state but rather change state independently. We show that change instead of exchange does *not* affect the long-time behaviour. Both Appendices A and (B) underpin the generality of the three models under investigation. Appendices C and D contain the proof of technical lemmas that are needed in the proof of the convergence to equilibrium.

In three upcoming companion papers [GHO20b], [GHO20c], [GHO20a] we deal with three further aspects:

(I) In [GHO20b] we establish the *finite-systems scheme*, i.e., we identify in the co-existence regime how a finite truncation of the system behaves as both the time and the truncation level tend to infinity, properly tuned together. This underlines the relevance of systems with an infinite geographic space and a seed-bank with infinitely many colours for the description of systems with a large finite geographic

space and a seed-bank with a large finite number of colours. We show that there is a *single universality class* for the scaling limit, represented by a Fisher-Wright diffusion whose volatility constant is *reduced* by the seed-bank. We show that if the wake-up time has finite mean, then the scaling time is *proportional* to the geographical volume of the system, while if the wake-up time has infinite mean, then the scaling time grows *faster* than the geographical volume of the system. We also investigate what happens for systems with a large finite geographic space and a seed-bank with infinitely many colours, where the behaviour turns out to be different.

(II) In [GHO20c] we consider the special case where the colonies are organised in a *hierarchical fashion*, i.e., the geographic space is the hierarchical group Ω_N of order N. We identify the parameter regime for clustering for all $N < \infty$, and analyse the *multi-scale behaviour* of the system in the hierarchical mean-field limit $N \to \infty$ by looking at block averages on successive hierarchical space-time scales. Playing with the migration kernel, we can choose the migration to be close to critically recurrent in the sense of potential theory. By letting $N \to \infty$ we can approach the *critical dimension*, so that the migration becomes similar to migration on the two-dimensional Euclidean geographic space. With the help of *renormalisation arguments* we show that, close to the critical dimension, the scaling behaviour on large space-time scales is *universal*.

(III) In [GHO20a] we identify the *pattern of cluster formation* in the clustering regime (= how fast mono-type clusters grow in time) and describe the *genealogy* of the population. The latter provides further insight into how the seed-bank enhances genetic diversity.

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

2 Introduction of the three models and their basic properties

In Section 2.1 we give a formal definition of the three models of increasing generality. In Section 2.2 we comment on their biological significance. In Section 2.3 we establish their well-posedness via a martingale problem (Theorem 2.4). In Section 2.4 we introduce the associated dual processes and state the relevant duality relations (Theorems 2.5, 2.8 and 2.10). In Section 2.5 we use these duality relations to formulate a criterion for clustering versus coexistence (Theorems 2.11 and 2.13).

2.1 Migration, resampling and seed-bank: three models

In this section we extend the model for a population with seed-bank from [BGCKWB16] to three models of increasing generality for spatial populations with seed-bank. In each of the three models, we consider populations of individuals of two types – either $\heartsuit$ or $\diamondsuit$ – located in a *geographic space* $\mathbb{G}$ that is a countable Abelian group endowed with the discrete topology. In each of the three models, the population in a colony consist of an *active* part and a *dormant* part. The *repository* of the dormant population at colony $i \in \mathbb{G}$ is called the seed-bank at $i \in \mathbb{G}$. Individuals in the active part of a colony $i \in \mathbb{G}$ can resample, migrate and exchange with a dormant population. Individuals in the dormant part of a colony $i \in \mathbb{G}$ only exchange with the active population. An active individual that resamples chooses uniformly at random another individual from its colony and adopts its type. (Alternatively, resampling may be viewed as the active individual being replaced by a copy of the active individual chosen. Because individuals carry a type and not a label, this gives the same model.) When an active individual at colony $i \in \mathbb{G}$ migrates, it

chooses a parent from another colony $j \in \mathbb{G}$ and adopts its type. In each of the three models the migration is described by a *migration kernel* $a(\cdot,\cdot)$, which is an irreducible $\mathbb{G} \times \mathbb{G}$ matrix of transition rates satisfying

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

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

Since we look at continuum models obtained from individual-based models, we are interested in the frequencies of type $\heartsuit$ in the different colonies. In Appendix A we discuss the individual-based models underlying the continuum models described below. In particular, we show that the continuum models arise as the *large-colony-size limit* of the individual-based models.

Remark 2.1 (Notation). Throughout the paper, in evolution equations we use lower case letters for *components* and upper case letters for *systems of components*. □

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

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

on the state space

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

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

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

$$\mathrm{d}y_i(t) = e\left[x_i(t) - y_i(t)\right] \mathrm{d}t, \qquad i \in \mathbb{G}, \tag{2.5}$$

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

The state space E equipped with the product topology is a Polish space. The factor $K \in (0, \infty)$ is defined by

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

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

and is the same for all colonies $i \in \mathbb{G}$. The factor K turns up in the scaling limit of the individual-based model when there is an *asymmetry* between the sizes of the active and the dormant population (see Appendix A). In Fig. 1 we give a schematic illustration of the process (2.4)–(2.5). A detailed description of the underlying individual-based model, as well as a derivation of the continuum limit (2.4)–(2.5) from the individual-based model following [BGCKWB16], can be found in Appendix A. The continuum limit is also referred to as the frequency limit or the diffusion limit.

Remark 2.2 (Interpretation of the state space). Note that the *state space* of the system can also be written as

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

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

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

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

Model 2: multi-layer seed-bank. In this model we give the seed-bank an internal structure by colouring the dormant individuals with countably many colours $m \in \mathbb{N}_0$. Active individuals that become dormant are assigned a colour m that is drawn randomly from an infinite sequence of colours labeled by $\mathbb{N}_0$ (see Fig. 2 for an illustration). As will be explained in Section 2.2, this captures the *different ways in which individuals can enter into the seed-bank*. In Section 2.4 we will show how this internal structure allows for fat tails in the wake-up times of individuals while preserving the Markov property.

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

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

$$\big(Z(t)\big)_{t \geq 0}, \qquad Z(t) = \big(z_i(t)\big)_{i \in \mathbb{G}}, \qquad z_i(t) = (x_i(t), (y_{i,m}(t))_{m \in \mathbb{N}_0}), \tag{2.9}$$
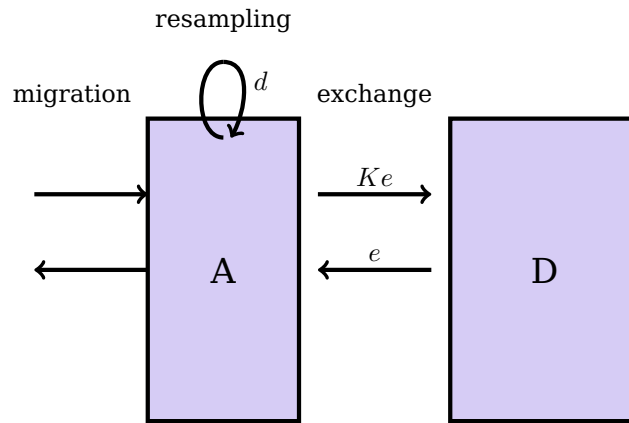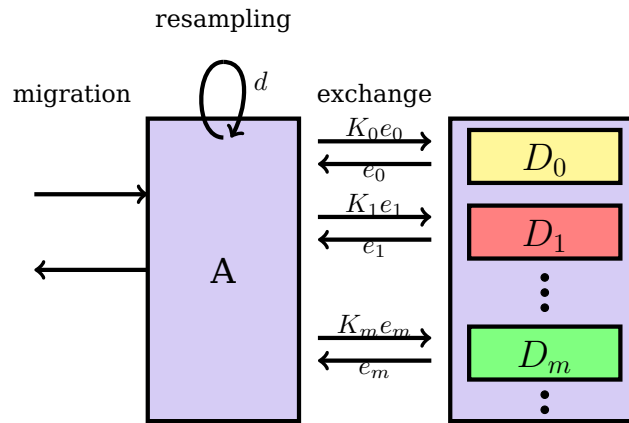
Figure 2: The evolution in a single colony in Model 2. Individuals are subject to migration, resampling and exchange with the seed-bank, as in Model 1. Additionally, when individuals become dormant they get a colour and when they become active they loose their colour.

on the state space

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

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

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

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

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

$$\mathrm{d}y_{i,m}(t) = e_m \left[x_i(t) - y_{i,m}(t) \right] \mathrm{d}t, \qquad m \in \mathbb{N}_0, \ i \in \mathbb{G}, \tag{2.13}$$

where we have to assume that

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

since otherwise active individuals become dormant instantly. Comparing (2.12)–(2.13) with the SSDE of Model 1 in (2.4)–(2.5), we see that active individuals migrate (the first term in (2.12)), resample (the second term in (2.12)), but now interact with a whole sequence of dormant populations (the third term in (2.12) and the term in (2.13)). See Fig. 3 for an illustration of the migration. As initial state $Z(0) = z$ we may again take any $z \in E$.

Remark 2.3 (Interpretation of the state space)**.** Note that, like in Remark 2.2, the *state space* of the system can also be written as

$$E = [0,1]^{\mathbb{S}}, \qquad \mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}. \tag{2.15}$$
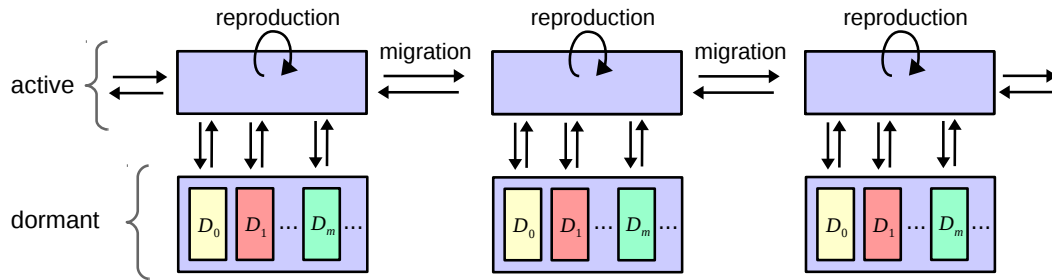
Figure 3: The spatial evolution in Model 2. Each colony has its own seed-bank. Reproduction occurs within the active population in each colony. Migration occurs between active populations in different colonies (top layer). Exchange between the active and the dormant population occurs in each colony (transitions between the top layer and the bottom layer.

With this interpretation, the process is denoted by

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

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

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

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

and replacing (2.12)–(2.13) by

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

$$\mathrm{d}y_{i,m}(t) = \sum_{j \in \mathbb{G}} e_m\, a_m(i, j)\, [x_j(t) - y_{i,m}(t)]\, \mathrm{d}t, \qquad m \in \mathbb{N}_0, \, i \in \mathbb{G}. \tag{2.19}$$

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

Initial laws. The evolution equations in (2.4)–(2.5), (2.12)–(2.13) and (2.18)–(2.19) must be supplemented with a choice of initial state. We draw the initial state according to a law $\mu(0)$ on $\mathcal{P}(E)$, the set of probability measures on E, and write $\mu(t)$ to denote the law evolved at time t.

We use the following classes of initial laws:

$$\mathcal{T} = \big\{\mu \in \mathcal{P}(E) \colon \ \mu \text{ is invariant under translations in } \mathbb{G}\big\},$$
$$\mathcal{T}^{\mathrm{erg}} = \big\{\mu \in \mathcal{T} \colon \ \mu \text{ is ergodic under translations in } \mathbb{G}\big\}, \tag{2.20}$$

where translation stands for *group action*, and

$$\mathcal{T}^{\bullet} = \big\{\mu \in \mathcal{T} \colon \ \mu \text{ is colour regular}\big\},$$
$$\mathcal{T}^{\mathrm{erg},\bullet} = \big\{\mu \in \mathcal{T}^{\mathrm{erg}} \colon \ \mu \text{ is colour regular}\big\}, \tag{2.21}$$

where *colour regular* means that $\lim_{m \to \infty} \mathbb{E}_\mu[y_{0,m}]$ exists, and

$$\mathcal{I} = \big\{\mu \in \mathcal{T} \colon \ \mu \text{ is invariant under the evolution}\big\}. \tag{2.22}$$

An example of an element that lies in all four classes in (2.20)–(2.21) is $\mu = \delta_{\underline{\theta}}$, with $\underline{\theta} \in \mathbb{S}$ the state whose components are all equal to some $\theta \in [0,1]$.

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

$$\chi = \sum_{m \in \mathbb{N}_0} K_m e_m < \infty \tag{2.23}$$

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

$$\rho = \sum_{m \in \mathbb{N}_0} K_m = \frac{\text{size dormant population}}{\text{size active population}}. \tag{2.24}$$

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

2.2 Comments

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

(2) As geographic space $\mathbb{G}$ we allow *any* countable Abelian group endowed with the discrete topology. Key examples are the Euclidean lattice $\mathbb{G} = \mathbb{Z}^d$, $d \in \mathbb{N}$, and the hierarchical lattice $\mathbb{G} = \Omega_N$, $N \in \mathbb{N}$. In this paper we will focus on $\mathbb{G} = \mathbb{Z}^d$. The case $\mathbb{G} = \Omega_N$ will be considered in more detail in [GHO20c].

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

In Model 2, the seed-bank is a repository for seeds with one of infinitely many colours. *The colours provide us with a tool to model different distributions for the time an individual stays dormant without loosing the Markov property for the evolution of the system.* Tuning the parameters K_m and e_m properly and subsequently forgetting about the colours, we can mimic different distributions for

the time an individual stays dormant. This is of biological significance, especially in colonies of bacteria, where individuals stay dormant for random times whose distribution is fat-tailed (see [LJ11]).

In Model 3, the seed may even be blown elsewhere. Individuals that displace before becoming dormant are observed in plant-species as well as in bacteria populations (see [LJ11]).

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

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

$$g_{\mathrm{FW}}(x) = x(1-x), \qquad x \in [0,1], \tag{2.25}$$

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

$$\mathcal{G} = \Big\{ g \colon [0,1] \to [0, \infty) \colon g(0) = g(1) = 0,\ g(x) > 0\ \forall\, x \in (0,1),\ g \text{ Lipschitz} \Big\}. \tag{2.26}$$

This class is appropriate because a diffusion with a diffusion function $g \in \mathcal{G}$ stays confined to $[0,1]$, yet can go everywhere in $[0,1]$ ([Bre68, Chapter 16, Section 7]). Picking $g \neq dg_{\mathrm{FW}}$ amounts to allowing the resampling rate to be *state-dependent*, which is a significant 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]. *In the sequel we allow for general diffusion functions $g \in \mathcal{G}$ in all three models, unless stated otherwise.*

2.3 Well-posedness

For every law on E, with E depending on the choice of model, we want the SSDE for Models 1, 2 and 3 to define a *Borel Markov process*, i.e., the law of the path is a Borel measurable function of the initial state [Daw93, p.62]. We use a martingale problem, in the sense of [EK86, p.173], to characterize the law of the SSDE. Let

$$\mathcal{F} = \Big\{ f \in C_b(E, \mathbb{R}) \colon f \text{ depends on finitely many components} \\ \text{and is twice continuously differentiable in each component} \Big\}. \tag{2.27}$$

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

$$G = \sum_{i \in \mathbb{G}} \Bigg(\Bigg[\sum_{j \in \mathbb{G}} a(i,j)(x_j - x_i) \Bigg] \frac{\partial}{\partial x_i} + g(x_i) \frac{\partial^2}{\partial x_i^2} \\ + Ke\,(y_i - x_i) \frac{\partial}{\partial x_i} + e\,(x_i - y_i) \frac{\partial}{\partial y_i} \Bigg), \tag{2.28}$$

for Model 2 ((2.12)–(2.13)) reads

$$G = \sum_{i \in \mathbb{G}} \left(\left[\sum_{j \in \mathbb{G}} a(i,j)(x_j - x_i) \right] \frac{\partial}{\partial x_i} + g(x_i) \frac{\partial^2}{\partial x_i^2} \right.$$
$$\left. + \sum_{m \in \mathbb{N}_0} \left[K_m e_m \left(y_{i,m} - x_i \right) \frac{\partial}{\partial x_i} + e_m \left(x_i - y_{i,m} \right) \frac{\partial}{\partial y_{i,m}} \right] \right), \tag{2.29}$$

while for Model 3 ((2.18)–(2.19)) the last term in the right-hand side of (2.29) is to be replaced by

$$\sum_{i,j \in \mathbb{G}} \sum_{m \in \mathbb{N}_0} \left[K_m e_m \, a_m(j,i) \left(y_{j,m} - x_i \right) \frac{\partial}{\partial x_i} + e_m \, a_m(i,j) \left(x_j - y_{i,m} \right) \frac{\partial}{\partial y_{i,m}} \right]. \tag{2.30}$$

Theorem 2.4 (Well-posedness: Models 1–3)**.** *For each of the three models the following holds:*
(a) The SSDE has a unique strong solution in $C([0,\infty), E)$. Its law is the unique solution of the $(G, \mathcal{F}, \delta_u)$-martingale problem for all $u \in E$.
(b) The process on E is Feller and strong Markov. Consequently, the SSDE defines a unique Borel Markov process starting from any initial law on E.

2.4 Duality

For $g = dg_{\mathrm{FW}}$ the three models have a tractable dual, which will be seen to play a crucial role in the analysis of their long-time behaviour. For $g \neq dg_{\mathrm{FW}}$ the three models do *not* have a tractable dual. However, as we will see later, we can *compare* them with models that do and use this comparison to determine their long-time behaviour.

In [BGCKWB16, Sections 2.2 and 3] it was shown that the non-spatial Fisher-Wright diffusion with seed-bank is dual to the so-called *block-counting process* of a seed-bank coalescent. The latter describes the evolution of the number of partition elements in a partition of $n \in \mathbb{N}$ individuals, sampled from the current population, into subgroups of individuals with the same ancestor (i.e., individuals that are identical by descent). The enriched dual generates the ancestral lineages of the individuals evolving according to a Fisher-Wright diffusion with seed-bank, i.e., generates their full genealogy. The corresponding block-counting process counts the number of ancestral lineages left when traveling backwards in time. In this section we will extend the duality results in [BGCKWB16] to the spatial setting.

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

$$\begin{aligned} &\pi = ((\pi_1, \eta_1), \ldots, (\pi_{\bar{n}}, \eta_{\bar{n}})), \quad \bar{n} = |\pi|, \\ &\pi_\ell \subset \{1, \ldots, n\}, \text{ and } \{\pi_1, \cdots \pi_{\bar{n}}\} \text{ is a partition of } \{1, \ldots, n\}, \\ &\eta_\ell \in \mathbb{S}, \quad \ell \in \{1, \ldots, \bar{n}\}, \quad 1 \leq \bar{n} \leq n. \end{aligned} \tag{2.31}$$

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

$$(\mathcal{C}^{(n)}(t))_{t \geq 0}, \tag{2.32}$$

and starts from some

$$\mathcal{C}^{(n)}(0) = \pi(0), \qquad \pi(0) = \{(\{1\}, \eta_1), \ldots, (\{n\}, \eta_n)\}, \qquad \eta_1, \ldots, \eta_n \in \mathbb{S}. \tag{2.33}$$
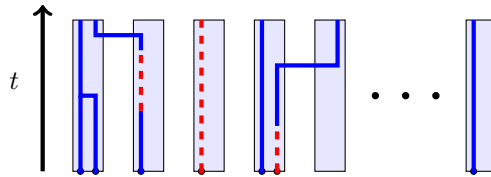
Figure 4: Picture of the evolution of lineages in the spatial coalescent. The purple blocks depict the colonies, the blue lines the active lineages, and the red dashed lines the dormant lineages. Blue lineages can migrate and become dormant, (i.e., become red dashed lineages). Two blue lineages can coalesce 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.

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

1. Each partition element moves independently of all other partition elements according the kernel

$$b^{(1)}((i, R_i), (j, R_j)) = \begin{cases} a(i,j), & \text{if } R_i = R_j = A, \\ Ke, & \text{if } i = j, \ R_i = A, \ R_j = D, \\ e, & \text{if } i = j, \ R_i = D, \ R_j = A, \\ 0, & \text{otherwise,} \end{cases} \qquad (2.34)$$

   where $a(\cdot, \cdot)$ is the migration kernel defined in (2.1), K is the relative size of the dormant population defined in (2.6), and e is the rate of exchange between the active and the dormant population defined in (2.4)–(2.5). What (2.34) says is that in the dual an active partition element migrates according to the transition kernel $a(\cdot, \cdot)$ and becomes dormant at rate Ke, while a dormant partition element can only become active and does so at rate e. In (2.34), the notation $b^{(1)}$ marks that the kernel refers to Model 1. Later we will use the notation $b^{(2)}$ for Model 2 and $b^{(3)}$ for Model 3.

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

The spatial seed-bank coalescent $(\mathcal{C}(t))_{t \geq 0}$ is defined as the projective limit of the n-spatial seed-bank coalescents $(\mathcal{C}^{(n)}(t))_{t \geq 0}$ as $n \to \infty$. This object is well-defined by Kolmogorov's extension theorem (see [BGCKWB16, Section 3]).

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

$$L(t) = \big(\ell_{(i,A)}(t), \ell_{(i,D)}(t) \big)_{i \in \mathbb{G}},$$

$$\ell_{(i,A)}(t) = \ell_{(i,A)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t) = (i,A)\}}, \qquad \ell_{(i,D)}(t) = \ell_{(i,D)}(\mathcal{C}^{(n)}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t) = (i,D)\}}.$$

$$(2.35)$$

Therefore $(L(t))_{t \geq 0}$ has state space $E' = (\mathbb{N}_0 \times \mathbb{N}_0)^{\mathbb{G}}$. We denote the elements of E' by sequences $(m_i, n_i)_{i \in \mathbb{G}}$, and define $\delta_{(j, R_j)} \in E'$ to be the element of E' that is 0 at

all sites $(i, R_i) \in \mathbb{G} \times \{A, D\} \backslash (j, R_j)$, but 1 at the site (j, R_j). From the evolution of $\mathcal{C}^{(n)}(t)$ described below (2.32) we see that the block-counting process has the following transition kernel:

$$(m_i, n_i)_{i \in \mathbb{G}} \rightarrow \begin{cases} (m_i, n_i)_{i \in \mathbb{G}} - \delta_{(j,A)} + \delta_{(k,A)}, & \text{at rate } m_j a(j, k) \text{ for } j, k \in \mathbb{G}, \\ (m_i, n_i)_{i \in \mathbb{G}} - \delta_{(j,A)}, & \text{at rate } d\binom{m_j}{2} \text{ for } j \in \mathbb{G}, \\ (m_i, n_i)_{i \in \mathbb{G}} - \delta_{(j,A)} + \delta_{(j,D)}, & \text{at rate } m_j K e \text{ for } j \in \mathbb{G}, \\ (m_i, n_i)_{i \in \mathbb{G}} + \delta_{(j,A)} - \delta_{(j,D)}, & \text{at rate } n_j e \text{ for } j \in \mathbb{G}. \end{cases} \tag{2.36}$$

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

$$H\Big((x_i, y_i)_{i \in \mathbb{G}}, (m_i, n_i)_{i \in \mathbb{G}}\Big) = \prod_{i \in \mathbb{G}} x_i^{m_i} y_i^{n_i} = \prod_{i \in \mathbb{G}} x_i^{\ell_{(i,A)}} y_i^{\ell_{(i,D)}}. \tag{2.37}$$

Note that $H(\cdot, (m_i, n_i)_{i \in \mathbb{G}})$ is in the domain of the generator G defined in (2.28) for every $(m_i, n_i)_{i \in \mathbb{G}}$.

The *duality relation* reads as follows.

Theorem 2.5 (Duality relation: Model 1). *Let* $g = d g_{\mathrm{FW}}$. *Let* H *be defined as in* (2.37). *Then for all* $(x_i, y_i)_{i \in \mathbb{G}} \in E$ *and* $(m_i, n_i)_{i \in \mathbb{G}} \in E'$,

$$\mathbb{E}_{(x_i, y_i)_{i \in \mathbb{G}}}\Big[H\Big((x_i(t), y_i(t))_{i \in \mathbb{G}}, (m_i, n_i)_{i \in \mathbb{G}}\Big)\Big] = \mathbb{E}_{(m_i, n_i)_{i \in \mathbb{G}}}\Big[H\Big((x_i, y_i)_{i \in \mathbb{G}}, (m_i(t), n_i(t))_{i \in \mathbb{G}}\Big)\Big] \tag{2.38}$$

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

Since the duality function H gives all the mixed moments of $(Z(t))_{t \geq 0}$, the duality relation in Theorem 2.5 is called a *moment dual*.

Remark 2.6 (Duality relation in terms of the effective geographic space). Interpreting $(Z(t))_{t \geq 0}$ with $Z(t) = (z_u(t))_{u \in \mathbb{S}}$ as a process on the effective geographic space $\mathbb{S}$ (recall Remark 2.2), we can rewrite the duality relation. Let the block-counting process $(L(t))_{t \geq 0} = (L(\mathcal{C}(t))_{t \geq 0}$ count at each site $u \in \mathbb{S}$ the number of partition elements of $\mathcal{C}(t)$, i.e.,

$$L(t) = (\ell_u(t))_{u \in \mathbb{S}},$$
$$\ell_u(t) = \ell_u(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t) = u\}}, \tag{2.39}$$

and rewrite the duality function H in (2.37) as

$$H(Z, L) = \prod_{u \in \mathbb{S}} z_u^{\ell_u}. \tag{2.40}$$

Then, on $\mathbb{E} \times \mathbb{E}'$, the *duality relation* reads

$$\mathbb{E}_Z\big[H\big(Z(t), L\big)\big] = \mathbb{E}_L\big[H\big(Z, L(t)\big)\big]. \tag{2.41}$$

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

Remark 2.7 (Wake-up times). Define (see Fig. 4)

$\sigma = $ typical time spent by an ancestral lineage in state A before switching to state D,

$\tau = $ typical time spent by an ancestral lineage in state D before switching to state A.

(2.42)

(Here, the word typical refers to the fact that σ and τ are random variables that do not depend on the geographic location. For a more precise definition we refer to Section 5.2 and Fig. 5.) It follows from (2.34) that

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

(2.43)

An ancestral lineage in the dual of the spatial seed-bank process behaves as an ancestral lineage in the dual of a spatial Fisher-Wright diffusion without seed-bank (see e.g. [FG96]), but becomes dormant every once in a while. In the long run we expect an ancestral lineage to be active only a fraction $\frac{1}{1+K}$ of the time. We will see in Section 5 that the effect of the seed-bank on the long-time behaviour of the ancestral lineages in the dual is a slow down by a factor $\frac{1}{1+K}$ compared to the long-time behaviour of the ancestral lineages in the dual of interacting Fisher-Wright diffusions without seed-bank. $\qquad\square$

Model 2. The dual for Model 2 arises naturally from the dual for Model 1 by adding internal states to the seed-bank and adapting the rates of becoming active and dormant accordingly. Recall that for Model 2 the effective geographic space is $\mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$. Migration and coalescence are as before, but at every colony switches between an active copy A and a dormant copy D_m now occur at rates e_m, respectively, $K_m e_m$. The spatial coalescent $(\mathcal{C}(t))_{t \geq 0}$ in (2.32) starts from an initial configuration like (2.33) and evolves according to the same two rules, but the transition kernel $b(\cdot, \cdot)$ must be replaced by

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

(2.44)

The corresponding block-counting process becomes

$$\begin{aligned} L(t) &= \left(\ell_{(i,A)}(t), \left(\ell_{(i,D_m)}(t) \right)_{m \in \mathbb{N}_0} \right)_{i \in \mathbb{G}}, \\ \ell_{(i,A)}(t) &= \ell_{(i,A)}(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=(i,A)\}}, \\ \ell_{(i,D_m)}(t) &= \ell_{(i,D_m)}(\mathcal{C}(t)) = \sum_{\ell=1}^{\bar{n}} 1_{\{\eta_\ell(t)=(i,D_m)\}}, \ m \in \mathbb{N}_0. \end{aligned}$$

(2.45)

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

$(m_i, (n_{i,m})_{m \in \mathbb{N}_0})_{i \in \mathbb{G}}$

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

(2.46)

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

$$H\Big((x_i, y_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0}, (m_i, n_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0}\Big) = \prod_{i\in\mathbb{G}} \prod_{m\in\mathbb{N}_0} x_i^{m_i} y_{i,m}^{n_{i,m}}. \tag{2.47}$$

Theorem 2.8 (Duality relation: Model 2). *For* $g = dg_{\mathrm{FW}}$, $(x_i, y_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0} \in E$, $(m_i, n_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0} \in E'$ *and* $t \geq 0$,

$$\begin{aligned}
&\mathbb{E}_{(x_i, y_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0}} \Big[H\Big((x_i(t), y_{i,m}(t))_{i\in\mathbb{G}, m\in\mathbb{N}_0}, (m_i, n_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0}\Big) \Big] \\
&= \mathbb{E}_{(m_i, n_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0}} \Big[H\Big((x_i, y_{i,m})_{i\in\mathbb{G}, m\in\mathbb{N}_0}, (m_i(t), n_{i,m}(t))_{i\in\mathbb{G}, m\in\mathbb{N}_0}\Big) \Big].
\end{aligned} \tag{2.48}$$

By rewriting the block-counting process as in Remark 2.6, the duality function can be rewritten as

$$H(Z, L) = \prod_{u\in\mathbb{S}} z_u^{l_u} \tag{2.49}$$

and the *duality relation* reads

$$\mathbb{E}_Z\big[H\big(Z(t), L\big)\big] = \mathbb{E}_L\big[H\big(Z, L(t)\big)\big]. \tag{2.50}$$

Remark 2.9 (Fat-tailed wake-up times). Recall the definition of χ in (2.23) and the definition of ρ in (2.24). Define

$$\sigma = \text{typical time spent by an ancestral lineage in the active state } A$$
$$\qquad \text{before switching to a dormant state } \cup_{m\in\mathbb{N}_0} D_m,$$
$$\tau = \text{typical time spent by an ancestral lineage in the dormant state } \cup_{m\in\mathbb{N}_0} D_m$$
$$\qquad \text{before switching to the active state } A. \tag{2.51}$$

(Note that the switch does not record the colour of the dormant state.) It follows from (2.44) that

$$\begin{aligned}
\mathbb{P}(\sigma > t) &= \mathrm{e}^{-\chi t}, \\
\mathbb{P}(\tau > t) &= \sum_{m\in\mathbb{N}_0} \frac{K_m e_m}{\chi} \mathrm{e}^{-e_m t},
\end{aligned} \tag{2.52}$$

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

$$\mathbb{E}\left[\tau\right] = \frac{\rho}{\chi}. \tag{2.53}$$

If $\rho < \infty$, then we invoke the seed-bank colours and use the balance equations for recurrent Markov chains to see that each ancestral lineage in the dual in the long run spends a fraction $\frac{\rho}{1+\rho}$ of the time in the dormant state. Like in Model 1, an ancestral lineage in the dual behaves like an ancestral lineage in the dual of interacting Fisher-Wright diffusions, but is slowed down by a factor $\frac{\rho}{1+\rho}$. However, if $\rho = \infty$, then (2.44) together with (2.53) imply that each ancestral lineage in the dual behaves like a null-recurrent Markov chain on $\{A, (D_m)_{m\in\mathbb{N}_0}\}$, and consequently the probability to be active tends to 0 as $t \to \infty$. Therefore we may expect that the long-time behaviour of the system is affected by the seed-bank. In particular, choosing

$$\begin{aligned}
&K_m \sim A\, m^{-\alpha}, \quad e_m \sim B\, m^{-\beta}, \quad m \to \infty, \\
&A, B \in (0, \infty), \quad \alpha, \beta \in \mathbb{R}\colon \alpha \leq 1 < \alpha + \beta,
\end{aligned} \tag{2.54}$$

we see that (2.52) implies

$$\mathbb{P}(\tau > t) \sim C t^{-\gamma}, \qquad t \to \infty, \tag{2.55}$$

with $\gamma = \frac{\alpha+\beta-1}{\beta}$ and $C = \frac{A}{\chi\beta} B^{1-\gamma} \Gamma(\gamma)$, where Γ is the Gamma-function. The conditions on α, β guarantee that $\rho = \infty$, $\chi < \infty$ (recall (2.23) and (2.24)). Examples are: $\alpha = 0$, $\beta > 1$ and $\alpha \in (0,1)$, $\beta > 1 - \alpha$. Thus, for $\rho = \infty$ we can model individuals with a fat-tailed wake-up time simply by not taking their colours into account. *The internal structure of the seed-bank captured by the colours allows us to model fat-tailed wake-up times without loosing the Markov property for the evolution.* The general mixture of exponentials in the second line of (2.52) covers a very large class of wake-up times. The formulas in (2.54)–(2.55) show that all possible tail exponents are included. As shown at the end of Section 3.2, even slowly varying modulations of the tail behaviour can be included, so that regular variation at infinity is fully covered. $\qquad\square$

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

$$b^{(3)}((i, R_i), (j, R_j)) = \begin{cases} a(i,j), & R_i = R_j = A, \\ K_m e_m a_m(j,i), & R_i = A,\ R_j = D_m,\ \text{for } m \in \mathbb{N}_0, \\ e_m a_m(i,j), & R_i = D_m,\ R_j = A,\ \text{for } m \in \mathbb{N}_0. \end{cases} \qquad (2.56)$$

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

$(m_i, (n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}}$

$$\to \begin{cases} (m_i, (n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} - \delta_{(j,A)} + \delta_{(k,A)}, & \text{at rate } m_j a(j,k) \text{ for } j, k \in \mathbb{G}, \\ (m_i, (n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} - \delta_{(j,A)}, & \text{at rate } d\binom{m_j}{2} \text{ for } j \in \mathbb{G}, \\ (m_i, (n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} - \delta_{(j,A)} + \delta_{(k,D_m)}, & \text{at rate } m_j K_m e_m a_m(k,j) \text{ for } j \in \mathbb{G}, \\ (m_i, (n_{i,m})_{m\in\mathbb{N}_0})_{i\in\mathbb{G}} + \delta_{(k,A)} - \delta_{(j,D_m)}, & \text{at rate } n_{j,m} e_m a_m(j,k) \text{ for } j \in \mathbb{G}. \end{cases} \qquad (2.57)$$

Theorem 2.10 (Duality relation: Model 3)**.** *Let $g = dg_{\mathrm{FW}}$. The same duality relation holds as in (2.48), where now the dual dynamics includes not only the exchange between active and dormant but also the accompanying displacement in space.*

2.5 Dichotomy criterion

For $g = dg_{\mathrm{FW}}$ the duality relations in Theorems 2.5, 2.8 and 2.10 provide us with the following criterion to characterise the long-term behaviour. If, in the limit as $t \to \infty$, locally only one type survives in the population, then we say that the system exhibits *clustering*. If, in the limit as $t \to \infty$, locally both types survive in the population, then we say that the system exhibits *coexistence*. Here, the word locally refers to the fact that weak convergence captures the behaviour in finite regions of geographic space only. We will see later that the evolution preserves the average fraction of types, so that if initially both types are present globally, then they are present globally at all later times.

For Model 1 the criterion reads as follows.

Theorem 2.11 (Dichotomy criterion: Model 1)**.** *Suppose that the initial law satisfies $\mu(0) \in \mathcal{T}^{\mathrm{erg}}$. Let $d \in (0, \infty)$. Then the system with $g = dg_{\mathrm{FW}}$ clusters if and only if in the dual two partition elements coalesce with probability 1.*

The idea behind Theorem 2.11 is as follows. If in the dual two partition elements coalesce with probability 1, then a random sample of n individuals drawn from the current population has a common ancestor some finite time backwards in time. Since individuals inherit their type from their parent individuals, this means that all n individuals have the same type. A formal proof will be given in Section 4.3.

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

Definition 2.12 (Colour regular initial measures). We say that a translation invariant initial law $\mu(0)$ is *colour regular* (recall (2.21)) when

$$\lim_{m \to \infty} \mathbb{E}_{\mu(0)}[y_{0,m}] \quad \text{exists,} \tag{2.58}$$

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

Thus, colour regularity is a condition on the deep seed-banks. This condition is needed because as time proceeds lineages starting from deeper and deeper seed-banks become active for the first time, and bring in their types into the active population. Without control on the initial states of the deep seed-banks, there may be no convergence to equilibrium, in which case we must work with multiple limit points. We will not pursue this generalisation, even though the limit points would still exhibit the dichotomy.

Theorem 2.13 (Dichotomy criterion: Models 2–3). *The same as in Theorem 2.11 is true for $\rho < \infty$, and also for $\rho = \infty$ when $\mu(0) \in \mathcal{T}^{\mathrm{erg},\bullet}$.*

Remark 2.14 (Clustering criterion general $g \in \mathcal{G}$). In Section 3 we will see that the dichotomy criterion in Theorems 2.11 and 2.13 for $g = dg_{\mathrm{FW}}$ does not depend on d, the rate of resampling. We will use *duality comparison arguments* to carry over the dichotomy criterion in Theorems 2.11 and 2.13 to $g \in \mathcal{G}$. We will see later that for all three models the system with g exhibits clustering if and only if the system with g_{FW} exhibits clustering. □

Remark 2.15 (Liggett conditions). We will see in Section 6.3 that, for Model 2 with $\rho = \infty$, if an initial measure satisfie $\mu \in \mathcal{T}^{\mathrm{erg},\bullet}$, then the Markov chain evolving according to $b^{(2)}(\cdot, \cdot)$ defined in (2.44) satisfies the following two conditions, for some $\theta \in [0,1]$ and for all (i, R_i), respectively, all $(i, R_i), (j, R_j)$:

(1)
$$\lim_{t \to \infty} \sum_{(k, R_k) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}\big((i, R_i), (k, R_k)\big) \, \mathbb{E}_\mu[z_{(k, R_k)}] = \theta, \tag{2.59}$$

(2)
$$\lim_{t \to \infty} \sum_{(k, R_k), (l, R_l) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}\big((i, R_i), (k, R_k)\big) \, b_t^{(2)}\big((j, R_j), (l, R_l)\big)$$
$$\times \mathbb{E}_\mu[z_{(k, R_k)} z_{(l, R_l)}] = \theta^2, \tag{2.60}$$

where $b_t^{(2)}(\cdot, \cdot)$ is the time-t transition kernel. These are precisely the conditions in [Lig85, Chapter V.1] necessary to determine the dichotomy in the long-time behaviour of the voter model. The parameter θ will be seen to play the role of the average density of type $\heartsuit$ in the population. The latter turns out to be preserved under the evolution. We will show that (1) and (2) imply convergence to a unique equilibrium that is invariant and ergodic under translations. It is difficult to identify exactly which initial measures μ satisfy (1) and (2). This is the reason why we work with sufficient conditions and need the notion of colour regularity.

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

3 Long-time behaviour

In this section we study the long-time behaviour of Models 1–3. In Sections 3.1–3.3 we prove convergence to a unique equilibrium measure, establish the dichotomy between clustering and coexistence, and identify which of the two occurs in terms of the migration kernel and the rates governing the exchange with the seed-bank (Theorems 3.1–3.8).

Throughout the sequel, g is a general diffusion function from the class $\mathcal{G}$ defined in (2.26). Special cases are the multiples of the standard Fisher-Wright diffusion function: $g = dg_{\mathrm{FW}}$, $d \in (0, \infty)$, with $g_{\mathrm{FW}}(x) = x(1-x)$, $x \in [0,1]$.

3.1 Long-time behaviour of Model 1

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

$$\hat{a}(i,j) = \tfrac{1}{2}[a(i,j) + a(j,i)], \qquad i, j \in \mathbb{G}, \tag{3.1}$$

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

$$t \mapsto \hat{a}_t(0,0) \text{ is regularly varying at infinity.} \tag{3.2}$$

(Examples can be found in [Hug95, Chapter 3].) Define

$$I_{\hat{a}} = \int_1^\infty \mathrm{d}t \, \hat{a}_t(0,0). \tag{3.3}$$

Note that $I_{\hat{a}} = \infty$ if and only if $\hat{a}(\cdot, \cdot)$ is recurrent (see e.g. [Spi64, Chapter 1]).

Define

$$\theta = \mathbb{E}_{\mu(0)} \left[\frac{x_0 + Ky_0}{1 + K} \right], \tag{3.4}$$

i.e., the average density of $\heartsuit$ at colony 0 under the initial law $\mu(0)$. If $\mu(0)$ is invariant and ergodic under translations, then θ is the average density of $\heartsuit$ in the entire population. From the SSDE in (2.4)–(2.5) we see that

$$\theta = \mathbb{E}_{\mu(t)} \left[\frac{x_0 + Ky_0}{1 + K} \right] = \mathbb{E}_{\mu(0)} \left[\frac{x_0(t) + Ky_0(t)}{1 + K} \right] \qquad \forall\, t \geq 0, \tag{3.5}$$

which says that the average density of $\heartsuit$ is a conserved quantity under the evolution. For $\theta \in [0,1]$, we define

$$\mathcal{T}_\theta^{\mathrm{erg}} = \left\{ \mu \in \mathcal{T}^{\mathrm{erg}} \colon \mathbb{E}_\mu \left[\frac{x_0 + Ky_0}{1 + K} \right] = \theta \right\}. \tag{3.6}$$

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

$$\mathbb{E}_\mu[f(z)g(z)] \geq \mathbb{E}_\mu[f(z)] \, \mathbb{E}_\mu[g(z)]. \tag{3.7}$$

Theorem 3.1 (Long-time behaviour: Model 1)**.** *Suppose that* $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$*.*

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

$$\lim_{t \to \infty} \mu(t) = \nu_\theta, \tag{3.8}$$

where

$$\nu_\theta \text{ is an equilibrium measure for the process on } E, \qquad (3.9)$$

$$\nu_\theta \text{ is invariant, ergodic and mixing under translations}, \qquad (3.10)$$

$$\nu_\theta \text{ is associated}, \qquad (3.11)$$

$$\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_0] = \theta, \qquad (3.12)$$

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

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

$$\lim_{t \to \infty} \mu(t) = \theta \, [\delta_{(1,1)}]^{\otimes \mathbb{G}} + (1 - \theta) \, [\delta_{(0,0)}]^{\otimes \mathbb{G}}. \qquad (3.13)$$

(Recall that convergence of laws is meant in the weak topology. In particular, (3.13) means that clustering occurs locally, with the average density of type $\heartsuit$ being equal to θ. Globally both types are present when $\theta \neq 0, 1$.)

The results in (3.8)–(3.13) say that the system converges to an equilibrium whose density of type $\heartsuit$ equals θ in (3.4), a parameter that is controlled by the initial state $\mu(0)$ and the asymmetry parameter K. The equilibrium can be either locally mono-type or locally multi-type, depending on whether the symmetrised migration kernel is recurrent or transient. If the equilibrium is mono-type, then the system grows large mono-type clusters (= clustering). If the equilibrium is multi-type, then the system allows $\heartsuit$ and $\diamondsuit$ to mix (= coexistence). In the case of coexistence, the equilibrium measure ν_θ also depends on the migration kernel $a(\cdot, \cdot)$, the values of the parameters e, K, and the diffusion function $g \in \mathcal{G}$ (recall (2.26)). The dichotomy itself, however, is *controlled by* $I_{\hat{a}}$ *only*. In particular, $g \in \mathcal{G}$ plays no role, a fact that will be shown with the help of a *duality comparison* argument. In view of Theorem 2.11, if $g = dg_{\mathrm{FW}}$, then $I_{\hat{a}} = \infty$ implies that with probability 1 two ancestral lineages in the dual coalesce. Therefore $I_{\hat{a}} = \infty$ is said to be *the total hazard of coalescence*. Remarkably, this dichotomy is the same as the dichotomy observed for systems without seed-bank (see [CG94]): clustering prevails for recurrent migration; coexistence prevails for transient migration; for $\mathbb{G} = \mathbb{Z}^d$ the critical dimension is $d = 2$. From the proof in Section 5.2 it will become clear that in the dual the ancestral lineages in the long run behave like the ancestral lineages without seed-bank, but are slowed down by a factor $\frac{1}{1+K}$. Consequently, the dormant periods of the ancestral lineages do not affect the dichotomy of the system. In particular, it does not affect the critical dimension separating clustering from coexistence.

Remark 3.2 (Continuity and ergodic decomposition)**.**
(1) We will show that $\theta \mapsto \nu_\theta$ is continuous in the weak topology.
(2) Because $\mathcal{T}$ is a Choquet simplex and $\theta \mapsto \nu_\theta$ is measurable, it follows that Theorem 3.1 carries over from $\mu(0) \in \mathcal{T}^{\mathrm{erg}} = \cup_{\theta \in [01]} \mathcal{T}_\theta^{\mathrm{erg}}$ to $\mu(0) \in \mathcal{T}$, after decomposition into ergodic components. $\qquad \square$

3.2 Long-time behaviour of Model 2

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

$$a(i, j) = a(j, i) \qquad \forall \, i, j \in \mathbb{G}. \qquad (3.14)$$

Note that $\hat{a}_t(0, 0) = a_t(0, 0)$ because of (3.14). (In Lemma 3.6 we comment on what happens when we drop this assumption.) Recall (2.23)–(2.24). It turns out that the long-time behaviour of Model 2 is different for $\rho < \infty$ and $\rho = \infty$.

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

$$\theta = \mathbb{E}_{\mu(0)}\left[\frac{x_0 + \sum_{m\in\mathbb{N}_0} K_m\, y_{0,m}}{1+\rho}\right],\qquad (3.15)$$

which is the counter part of (3.4) in Model 1. Like in Model 1, it follows from the SSDE in (2.12)–(2.13) that

$$\mathbb{E}_{\mu(t)}\left[\frac{x_0 + \sum_{m\in\mathbb{N}_0} K_m\, y_{0,m}}{1+\rho}\right] = \mathbb{E}_{\mu(0)}\left[\frac{x_0(t) + \sum_{m\in\mathbb{N}_0} K_m\, y_{0,m}(t)}{1+\rho}\right]\qquad \forall\, t\geq 0. \quad (3.16)$$

Hence the average density of type $\heartsuit$ is a preserved quantity under the evolution of the system. The dichotomy is controlled by the same integral $I_{\hat{a}}$ as defined in (3.3) for Model 1.

Case $\rho = \infty$. For an infinite seed-bank, we assume that (recall Remark 2.9)

$$\begin{aligned}
K_m \sim A\,m^{-\alpha}, \quad e_m \sim B\,m^{-\beta}, \quad m\to\infty,\\
A, B \in (0,\infty), \quad \alpha,\beta\in\mathbb{R}\colon \alpha\leq 1 < \alpha+\beta,
\end{aligned}\qquad (3.17)$$

for which

$$P(\tau > t) \sim C\,t^{-\gamma}, \quad t\to\infty,\qquad (3.18)$$

with $\gamma = \frac{\alpha+\beta-1}{\beta} \in (0,1)$ and $C = \frac{A}{\beta}B^{1-\gamma}\gamma\Gamma(\gamma) \in (0,\infty)$, where Γ is the Gamma-function. In addition, we assume that the initial measure $\mu(0)$ is colour regular (recall Definition 2.12), and define

$$\theta = \lim_{m\to\infty}\mathbb{E}_{\mu(0)}[y_{0,m}].\qquad (3.19)$$

This ensures the existence of the initial density

$$\theta = \lim_{M\to\infty}\mathbb{E}_{\mu(0)}\left[\frac{x_0 + \sum_{m=0}^{M} K_m\, y_{0,m}}{1+\sum_{m=0}^{M} K_m}\right].\qquad (3.20)$$

It turns out that the dichotomy is controlled by the integral

$$I_{\hat{a},\gamma} = \int_1^\infty \mathrm{d}t\ t^{-(1-\gamma)/\gamma}\,\hat{a}_t(0,0)\qquad (3.21)$$

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

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

$$\begin{aligned}
\rho < \infty\colon \quad \mathcal{T}_\theta^{\mathrm{erg}} &= \left\{\mu\in\mathcal{T}^{\mathrm{erg}}\colon \mathbb{E}_\mu\left[\frac{x_0 + \sum_{m\in\mathbb{N}_0} K_m y_{0,m}}{1+\sum_{m\in\mathbb{N}_0} K_m}\right] = \theta\right\},\\
\rho = \infty\colon \quad \mathcal{T}_\theta^{\mathrm{erg},\bullet} &= \left\{\mu\in\mathcal{T}^{\mathrm{erg},\bullet}\colon \lim_{M\to\infty}\mathbb{E}_\mu\left[\frac{x_0 + \sum_{m=0}^{M} K_m y_{0,m}}{1+\sum_{m=0}^{M} K_m}\right] = \theta\right\}.
\end{aligned}\qquad (3.22)$$

Theorem 3.3 (Long-time behaviour: Model 2). (I) *Let $\rho < \infty$. Assume (3.2) and (3.14). Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$.*

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

$$\lim_{t\to\infty}\mu(t) = \nu_\theta,\qquad (3.23)$$

where

$$\nu_\theta \text{ is an equilibrium measure for the process on } E, \tag{3.24}$$

$$\nu_\theta \text{ is invariant, ergodic and mixing under translations,} \tag{3.25}$$

$$\nu_\theta \text{ is associated,} \tag{3.26}$$

$$\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta \; \forall \, m \in \mathbb{N}_0, \tag{3.27}$$

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

$$\begin{aligned}
\liminf_{m \to \infty} e_m > 0 \colon \; \liminf_{m \to \infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) > 0, \\
\limsup_{m \to \infty} e_m = 0 \colon \; \limsup_{m \to \infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) = 0.
\end{aligned} \tag{3.28}$$

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

$$\lim_{t \to \infty} \mu(t) = \theta \, [\delta_{(1,1^{\mathbb{N}_0})}]^{\otimes \mathbb{G}} + (1 - \theta) \, [\delta_{(0,0^{\mathbb{N}_0})}]^{\otimes \mathbb{G}}. \tag{3.29}$$

(II) *Let $\rho = \infty$. Assume (3.2), (3.14) and (3.17). Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}, \bullet}$. Then the same results as in (I) hold after $I_{\hat{a}}$ in (3.3) is replaced by $I_{\hat{a}, \gamma}$ in (3.21). Moreover,*

$$\limsup_{m \to \infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) = 0, \tag{3.30}$$

The result in part (I) shows that for $\rho < \infty$ the long-time behaviour is similar to that of Model 1. Like in Model 1, the results in (3.23)–(3.29) say that the system converges to an equilibrium whose density of type $\heartsuit$ equals θ in (3.4), the density of $\heartsuit$ under the initial measure $\mu(0)$. Again, the equilibrium can be either mono-type or multi-type, depending on whether the symmetrised migration kernel is recurrent or transient. Like in Model 1, in both cases the equilibrium measure depends on θ. In the case of coexistence, the equilibrium measure ν_θ also depends on the migration kernel $a(\cdot, \cdot)$, the sequences of parameters $(e_m)_{m \in \mathbb{N}_0}$ and $(K_m)_{m \in \mathbb{N}_0}$, and the diffusion function $g \in \mathcal{G}$ (recall (2.26)). Again, the dichotomy itself is *controlled by $I_{\hat{a}}$ only*, and the resampling rate given by $g \in \mathcal{G}$ plays no role. Therefore if $g = dg_{FW}$, in view of Theorem 2.11, whether or not two ancestral lineages in the dual coalesce with probability 1 is still only determined by the migration kernel $a(\cdot, \cdot)$. The same dichotomy holds as for systems without seed-bank (see [CG94]). Therefore part (I) of Theorem 3.3 indicates that, as long as the dormant periods of the ancestral lineages in the dual have a finite mean (here $\frac{\rho}{1+\rho}$; recall Remark 2.9), the seed-bank does not affect the dichotomy of the system.

Even so, (3.28) indicates that there is interesting behaviour in the deep seed-banks. Indeed, when the exchange rate e_m between the m-dormant and the active population is bounded away from zero as $m \to \infty$ the deep seed-banks are asymptotically *random*, while when e_m tend to zero as $m \to \infty$ the deep seed-banks are asymptotically *deterministic*. The latter means that the deep seed-banks serve as a reservoir, containing a fixed mixture of types. For $\rho < \infty$ this reservoir is too small to influence the dichotomy of the system, but not for $\rho = \infty$.

The result in Part (II) shows that for $\rho = \infty$ the system again converges to an equilibrium whose density of type $\heartsuit$ equals θ in (3.20), the density of $\heartsuit$ under the initial measure $\mu(0)$. The equilibrium can be mono-type or multi-type, but *the dichotomy criterion has changed.* Instead of $I_{\hat{a}}$, the dichotomy is now controlled by the integral $I_{\hat{a}, \gamma}$ (recall (3.21)), where γ is the parameter determined by relative sizes K_m of the colour m-dormant populations with respect to the active population and the exchanges rates $(e_m)_{m \in \mathbb{N}_0}$ with the seed-bank, recall (3.17)–(3.18). If $g = dg_{FW}$, γ is the parameter of the tail of the wake-up time of an ancestral lineages in the dual (recall (2.9)). Therefore if

$g = dg_{Fw}$, in view of Theorem 2.13, we see that the dormant periods of the ancestral lineages in the dual do affect whether or not two ancestral lineages in the dual coalesce with probability 1. For general $g \in \mathcal{G}$, the integral $I_{\hat{a},\gamma}$ in (3.21) shows a *competition* between migration and exchange. The smaller γ is, the longer the individuals remain dormant in the seed-bank, the smaller $I_{\hat{a},\gamma}$ is, and the more coexistence becomes likely. As a consequence clustering requires more stringent conditions than recurrent migration; for $\mathbb{G} = \mathbb{Z}^d$ the critical dimension is $1 < d < 2$ for $\gamma \in [\frac{1}{2}, 1]$ and $d = 1$ for $\gamma \in (0, \frac{1}{2})$. The *seed-bank enhances genetic diversity*. Note that $\gamma \uparrow 1$ links up with the case $\rho < \infty$, where coexistence occurs if and only if the migration is transient. Also note that for $\gamma \in (0, \frac{1}{2})$ there is always coexistence *irrespective of the migration*.

In the case of clustering the equilibrium measure only depends on θ, while in the case of coexistence, like for $\rho < \infty$, ν_θ depends on the migration kernel $a(\cdot, \cdot)$, the sequences of parameters $(e_m)_{m \in \mathbb{N}_0}$, $(K_m)_{m \in \mathbb{N}_0}$, and the diffusion function $g \in \mathcal{G}$. Since we assumed (3.17), we have $\limsup_{m \to \infty} e_m = 0$, and so we are automatically in the second case of (3.28). Hence the deep seed-banks are asymptotically deterministic, i.e., the m-dormant population converges in law to a deterministic state θ as $m \to \infty$. Roughly speaking, in case $g = dg_{FW}$, in equilibrium the volatility of a colour is inversely proportional to its average wake-up time in the dual. Since $\rho = \infty$, for each $M \in \mathbb{N}_0$ we have $\sum_{m=M}^{\infty} K_m = \infty$, and in the coexistence regime the effect of the seed-bank can be interpreted as a migration towards an infinite reservoir with deterministic density θ.

Remark 3.4 (Continuity and ergodic decomposition).
(1) For Model 2 with $\rho < \infty$, we will show that $\theta \mapsto \nu_\theta$ is continuous in the weak topology. For Model 2 with $\rho = \infty$, we will show that $\theta \mapsto \nu_\theta$ is continuous in the *uniform weak topology* defined as follows: $(\mu_n)_{n \in \mathbb{N}}$ is said to converge to μ in $\mathcal{T}^\bullet$ in the uniform weak topology if it converges to μ in the weak topology and $\lim_{n \to \infty} \mathcal{L}_{\mu_n}[\hat{\theta}] = \mathcal{L}_\mu[\hat{\theta}]$, where

$$\hat{\theta} = \lim_{M \to \infty} \frac{x_0 + \sum_{m=0}^{M} K_m y_{0,m}}{1 + \sum_{m=0}^{M} K_m} \tag{3.31}$$

is the limiting density of $\heartsuit$. (Later we will see that $\hat{\theta}$ exists μ-a.s. for every $\mu \in \mathcal{T}^\bullet$.) We will in fact show that the convergence to equilibrium also holds in the uniform weak topology.
(2) Because $\mathcal{T}$ is a Choquet simplex and $\theta \mapsto \nu_\theta$ is measurable, Theorem 3.3(I) carries over from $\mathcal{T}^{\mathrm{erg}} = \cup_{\theta \in [0,1]} \mathcal{T}_\theta^{\mathrm{erg}}$ to $\mathcal{T}$, while Theorem 3.3(II) carries over from $\mathcal{T}^{\mathrm{erg},\bullet} = \cup_{\theta \in [0,1]} \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ to $\mathcal{T}^\bullet$. $\qquad \square$

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

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

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

Corollary 3.5 (Three regimes). *Under the conditions of Theorem 3.3, the system in* (2.12)–(2.13) *(Model 2) has three different parameter regimes:*

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

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

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

Role of symmetry in migration. Unlike in Model 1, it is *not* possible to remove the symmetry assumption in (3.14), as the following counterexample shows. We consider Model 2 with $\rho < \infty$ under assumption (3.2), but we do not assume (3.14).

Lemma 3.6 (Counterexample). *Let* $\mathbb{G} = \mathbb{Z}^2$, *and for* $\eta \in (0,1)$ *pick*

$$a(i,j) = \begin{cases} \frac{1}{4}(1+\eta), & j = i + (1,0) \text{ or } i + (0,1), \\ \frac{1}{4}(1-\eta), & j = i - (1,0) \text{ or } i - (0,1), \end{cases} \tag{3.32}$$

i.e., nearest-neighbour random walk with drift upward and rightward. Suppose that τ *in* (3.18) *has a one-sided stable distribution with parameter* $\gamma \in (1,2)$ *(obtained from* (3.17) *but with* $\alpha, \beta \in \mathbb{R}$: $1 < \alpha < 1 + \beta$*). Then coexistence occurs because* $I_a < \infty$*, while* $I_{\hat{a}} = \infty$ *because* $\hat{a}(\cdot, \cdot)$ *is the transition kernel of simple random walk.*

Recall that for the two-dimensional nearest-neighbour random walk without drift we get clustering according to Theorem 3.3, independently of the distribution of τ. The key feature of the counterexample is that it corresponds to $\mathbb{E}(\tau) < \infty$ and $\mathbb{E}(\tau^2) = \infty$. Hence the central limit theorem fails for τ. We will see in Section 6.5 that the failure of the central limit theorem for τ is responsible for turning clustering into coexistence.

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

Remark 3.7 (Dichotomy criterion and degrees of random walk). We can read the condition $I_{\hat{a},\gamma} < \infty$ for coexistence versus $I_{a,\gamma} = \infty$ for clustering in terms of the *degree* of the random walk. Namely, let $\hat{a}(\cdot, \cdot)$ be the transition kernel of an irreducible random walk on a countable Abelian group. Then the degree δ of $\hat{a}(\cdot, \cdot)$ is defined as

$$\delta = \sup \left\{ \zeta > -1 \colon \int_1^\infty \mathrm{d}t\, t^\zeta\, \hat{a}_t(0,0) < \infty \right\}. \tag{3.33}$$

A more refined notation is to define the degree as δ^+ when the integral is finite *at* the degree and δ^- when the integral is infinite *at* the degree. Hence we can rephrase the dichotomy criterion in Theorem 3.3 as

$$\text{clustering} \quad \Longleftrightarrow \quad \text{either } -\frac{1-\gamma}{\gamma} \geq \delta^- \text{ or } -\frac{1-\gamma}{\gamma} > \delta^+. \tag{3.34}$$

For further details we refer to [DGW04], [DGW05], which relate the degree of the random walk to the tail of its return time to the origin. $\qquad\square$

Modulation of wake-up time with slowly varying function. Under weak conditions it is possible to modulate (3.18) by a slowly varying function. Assume that

$$\frac{P(\tau \in \mathrm{d}t)}{\mathrm{d}t} \sim \varphi(t)\, t^{-(1+\gamma)}, \quad t \to \infty, \tag{3.35}$$

with φ *slowly varying at infinity*. Define

$$\hat{\varphi}(t) = \begin{cases} \varphi(t), & \gamma \in (0,1), \\ \int_1^t \mathrm{d}s\, \varphi(s) s^{-1}, & \gamma = 1. \end{cases} \tag{3.36}$$

As shown in [BGT87, Section 1.3], without loss of generality we may take $\hat{\varphi}$ to be infinitely differentiable and to be represented by the integral

$$\hat{\varphi}(t) = \exp \left[\int_1^t \frac{\mathrm{d}u}{u}\, \psi(u) \right] \tag{3.37}$$

for some $\psi\colon [0,\infty) \to \mathbb{R}$ such that $\lim_{u\to\infty} |\psi(u)| = 0$. If we assume that ψ eventually has a sign and satisfies $|\psi(u)| \leq C/\log u$ for some $C < \infty$, then (3.21) needs to be replaced by

$$I_{\hat{a},\gamma,\varphi} = \int_1^\infty \mathrm{d}t\, \hat{\varphi}(t)^{-1/\gamma}\, t^{-(1-\gamma)/\gamma}\, \hat{a}_t(0,0). \tag{3.38}$$

A proof is given in Section 6.6. The modulation of the wake-up time by a slowly varying function appears naturally for the model on the hierarchical group, analysed in [GHO20c]. There the integral criterion for the dichotomy in (3.38) is needed to apply Theorem 3.3.

3.3 Long-time behaviour of Model 3

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

$$a_m(\cdot,\cdot) = a^\dagger(\cdot,\cdot) \qquad \forall\, m \in \mathbb{N}_0, \tag{3.39}$$

with $a^\dagger(\cdot,\cdot)$ an irreducible *symmetric* random walk kernel on $\mathbb{G} \times \mathbb{G}$. Let $\hat{a}_t^\dagger(\cdot,\cdot)$ denote the time-t transition kernel of the random walk with symmetrised displacement kernel $\hat{a}^\dagger(\cdot,\cdot)$ $(= a^\dagger(\cdot,\cdot))$ and jump rate 1. Assume that (compare with (3.2))

$$\begin{aligned} &t \mapsto (\hat{a}_t * \hat{a}_t^\dagger)(0,0) \text{ is regularly varying at infinity,} \\ &(\hat{a}_{Ct} * \hat{a}_t^\dagger)(0,0) \asymp (\hat{a}_t * \hat{a}_t^\dagger)(0,0) \text{ as } t \to \infty \text{ for every } C \in (0,\infty), \end{aligned} \tag{3.40}$$

where $*$ stands for convolution. Let

$$I_{\hat{a}*\hat{a}^\dagger} = \int_1^\infty \mathrm{d}t\, (\hat{a}_t * \hat{a}_t^\dagger)(0,0) \tag{3.41}$$

and

$$I_{\hat{a}*\hat{a}^\dagger,\gamma} = \int_1^\infty \mathrm{d}t\, t^{-(1-\gamma)/\gamma}\, (\hat{a}_t * \hat{a}_t^\dagger)(0,0). \tag{3.42}$$

Theorem 3.8 (Long-time behaviour: Model 3)**.** *Suppose that, in addition to the assumptions of Theorem 3.3, both (3.39) and (3.40) hold. Then the same results as for Model* 2 *hold: (I) for $\rho < \infty$ after $I_{\hat{a}}$ in (3.3) is replaced by $I_{\hat{a}*\hat{a}^\dagger}$ in (3.41); (II) for $\rho = \infty$ after $I_{\hat{a},\gamma}$ in (3.21) is replaced by $I_{\hat{a}*\hat{a}^\dagger,\gamma}$ in (3.42).*

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

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

$$\hat{a}_t(i,j) \leq \hat{a}_t(0,0), \qquad \hat{a}_t^\dagger(i,j) \leq \hat{a}_t^\dagger(0,0) \qquad \forall\, i,j \in \mathbb{G}\ \forall\, t \geq 0. \tag{3.43}$$

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

If $a(\cdot,\cdot) = a^\dagger(\cdot,\cdot)$, then $(a_t * a_t^\dagger)(0,0) = a_{2t}(0,0)$ and therefore the dichotomy is the same as for Model 2. Hence the extra displacement has in this case no effect on the dichotomy. However, if the displacement is transient while the migration is recurrent, then there is a difference. For instance, if $\rho < \infty$, the migration is a simple random walk on $\mathbb{Z}$, and the displacement is a symmetric random walk on $\mathbb{Z}$ with infinite mean, e.g. $a^\dagger(0,x) = a^\dagger(0,-x) \sim D|x|^{-\delta}$, $D \in (0,\infty)$, $\delta \in (1,2)$, then $I_{\hat{a}} = \infty$, $I_{\hat{a}^\dagger} < \infty$ and $I_{\hat{a}*\hat{a}^\dagger} < \infty$ [Spi64, Section 8]. Therefore there is clustering in Model 2, but coexistence in Model 3.

4 Proofs: Well-posedness and duality

In Section 4.1 we prove Theorem 2.4, in Section 4.2 Theorems 2.5, 2.8 and 2.10, and in Section 4.3 Theorems 2.11 and 2.13. We first give the proof for Model 1, and afterwards indicate what changes need to be made for Models 2–3. We will follow this routine also in later chapters, because it is easier conceptually and notationwise.

4.1 Well-posedness

In this section we prove Theorem 2.4.

Proof. (a) We first prove Theorem 2.4(a): existence and uniqueness of solutions to the SSDE. We do this for each of the three models separately, before we prove Theorem 2.4(b).

Model 1. Existence of the process defined in (2.4)–(2.5) for Model 1 is a consequence of the assumptions in (2.1), (2.17) and (2.23), in combination with [SS80, Theorem 3.2], which reads as follows:

Theorem 4.1 (Unique strong solution)**.** *Let $\mathbb{S}$ be a countable set, and let $Z = \{z_u\}_{u \in \mathbb{S}} \in [0,1]^{\mathbb{S}}$. Consider the stochastic differential equation*

$$\mathrm{d}z_u(t) = \alpha_u(z_u(t))\,\mathrm{d}B_u(t) + f_u(Z(t))\,\mathrm{d}t, \qquad u \in \mathbb{S}, \tag{4.1}$$

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

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

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

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

$$\begin{aligned} |f_u(Z^1) - f_u(Z^2)| &\leq \sum_{v \in \mathbb{S}} Q_{u,v}|z_v^1 - z_v^2|, \\ \text{for} \quad Z^1 &= \{z_v^1\}_{v \in \mathbb{S}} \in [0,1]^{\mathbb{S}}, \ Z^2 = \{z_v^2\}_{v \in \mathbb{S}} \in [0,1]^{\mathbb{S}}. \end{aligned} \tag{4.2}$$

- *For $Z \in [0,1]^{\mathbb{S}}$ and $z_u = 0$,*
$$f_u(Z) \geq 0. \tag{4.3}$$

- *For $Z \in [0,1]^{\mathbb{S}}$ and $z_u = 1$,*
$$f_u(Z) \leq 0. \tag{4.4}$$

Then (4.1) has a unique $[0,1]^{\mathbb{S}}$-valued strong solution with a continuous path.

To apply Theorem 4.1 to Model 1, recall that

$$\mathbb{S} = \mathbb{G} \times \{A, D\}, \tag{4.5}$$

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

$z_u = x_i$ when $u = (i, A)$ and $z_u = y_i$ when $u = (i, D)$. We can rewrite (2.4)–(2.5) in the form of (4.1) by picking

$$\alpha_u(z_u) = \begin{cases} \sqrt{g(x_i)}, & u = (i, A), \\ 0, & u = (i, D), \end{cases} \tag{4.6}$$

and

$$f_u(Z) = \begin{cases} \sum_{j \in \mathbb{G}} a(i, j) (x_j - x_i) + Ke (y_i - x_i), & u = (i, A), \\ e (x_i - y_i), & u = (i, D). \end{cases} \tag{4.7}$$

Since $g \in \mathcal{G}$ (recall (2.26)), the conditions in (1) are satisfied. To check the conditions in (2), define the matrix $Q = \{Q_{u,v}\}_{u,v \in \mathbb{S}}$ by

$$Q_{u,v} = \begin{cases} \sum_{j \in \mathbb{G}} a(i, j) + Ke, & u = (i, A), \ v = (i, A), \\ a(i, j), & u = (i, A), \ v = (j, A), \\ Ke, & u = (i, A), \ v = (i, D), \\ e, & u = (i, D), \ v = (i, D) \text{ or } u = (i, D), \ v = (i, A), \\ 0, & \text{otherwise.} \end{cases} \tag{4.8}$$

Then

$$\sum_{v \in \mathbb{S}} Q_{u,v} = \begin{cases} 2 \sum_{j \in \mathbb{G}} a(i, j) + 2Ke, & u = (i, A), \\ 2e, & u = (i, D). \end{cases} \tag{4.9}$$

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

By Itô's formula, the law of the strong solution solves the martingale problem. Uniqueness of that solution follows from [RY99, Theorem IX 1.7(i)]. This in turn implies the Markov property, plus the fact that $z \mapsto \mathcal{L}[Z(t) \mid Z(0) = z]$ is Borel measurable for every $t \geq 0$.

Model 2. To apply Theorem 4.1 to Model 2, recall that

$$\mathbb{S} = \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}. \tag{4.10}$$

Pick

$$\alpha_u(z_u) = \begin{cases} \sqrt{g(x_i)}, & u = (i, A), \\ 0, & u = (i, D_m), \ m \in \mathbb{N}_0, \end{cases} \tag{4.11}$$

and

$$f_u(Z) = \begin{cases} \sum_{j \in \mathbb{G}} a(i, j) (x_j - x_i) + \sum_{m \in \mathbb{N}_0} K_m e_m (y_{i,m} - x_i), & u = (i, A), \\ e_m (x_i - y_{i,m}), & u = (i, D_m). \end{cases} \tag{4.12}$$

Set

$$Q_{u,v} = \begin{cases} \sum_{j \in \mathbb{G}} a(i, j) + \sum_{m \in \mathbb{N}_0} K_m e_m, & u = (i, A), \ v = (i, A), \\ a(i, j), & u = (i, A), \ v = (j, A), \ j \neq i, \\ K_m e_m, & u = (i, A), \ v = (i, D_m), \\ e_m, & u = (i, D_m), \ v = (i, D_m) \text{ or } u = (i, D_m), \ v = (i, A), \\ 0, & \text{otherwise.} \end{cases} \tag{4.13}$$

Then, by assumptions (2.1) and (2.23), Q, f and α satisfy the conditions of Theorem 4.1.

Model 3. The state space $\mathbb{S}$ and the function α are the same as in Model 2. When $u \in \mathbb{S}$ is of the form (i, A), we must adapt the function f_u such that it takes the displacement of seeds into account. The matrix Q must be adapted accordingly and, by assumption (2.17), the conditions of Theorem 4.1 are again satisfied.

(b) The proof of Theorem 2.4(b) is the same for Models 1–3. The Feller property can be proved by using duality if $g = dg_{\mathrm{FW}}$, $d \in (0, \infty)$. For general g we use [SS80, Remark 3.2] (see also [Lig85, Theorem 5.8]). The Feller property in turn implies the strong Markov property. $\qquad\square$

4.2 Duality

In this section we prove Theorems 2.5, 2.8 and 2.10.

Model 1: Proof of Theorem 2.5.

Proof. We use the generator criterion (see [EK86, p.190–193] or [JK14, Proposition 1.2]) to prove the duality relation given in (2.38). Let $H((X,Y),(M,N))$ be defined as in (2.37), with $(X,Y) = (x_i, y_i)_{i \in \mathbb{G}}$ and $(M,N) = (m_i, n_i)_{i \in \mathbb{G}}$. Let F be the generator of the spatial block-counting process defined in (2.36). Then

$$
\begin{aligned}
&F\big(H\big((X,Y),\cdot\big)\big)(M,N) \\
&= \sum_{i \in \mathbb{G}} \Bigg[\sum_{k \in \mathbb{G}} m_i a(i,k) \left[H\big((X,Y),(M,N) - \delta_{(i,A)} + \delta_{(k,A)}\big) - H\big((X,Y),(M,N)\big) \right] \\
&\quad + d \binom{m_i}{2} \left[H\big((X,Y),(M,N) - \delta_{(i,A)}\big) - H\big((X,Y),(M,N)\big) \right] \\
&\quad + m_i K e \left[H\big((X,Y),(M,N) - \delta_{(i,A)} + \delta_{(i,D)}\big) - H\big((X,Y),(M,N)\big) \right] \\
&\quad + n_i e \left[H\big((X,Y),(M,N) + \delta_{(i,A)} - \delta_{(i,D)}\big) - H\big((X,Y)(M,N)\big) \right] \Bigg].
\end{aligned}
\tag{4.14}
$$

Recall that G is the generator of the SSDE (recall (2.27)–(2.28)). Let $\mathcal{D}_G$ denote the domain of G and $\mathcal{D}_F$ the domain of F. Let $(S_t)_{t \geq 0}$ denote the semigroup of the process $(Z(t))_{t \geq 0}$ in (2.2) and $(R_t)_{t \geq 0}$ the semigroup of the process $(L(t))_{t \geq 0}$ in (2.35). Since

$$
\frac{d^2}{dt^2} R_t\big(H\big((X,Y),\cdot\big)\big)(M,N) = F^2\big(R_t\big(H\big((X,Y),\cdot\big)\big)\big)(M,N),
\tag{4.15}
$$

we see that $H((X,Y),\cdot) \in \mathcal{D}_G$ and $R_t(H((X,Y),\cdot)) \in \mathcal{D}_G$. Furthermore, it is immediate that $H((\cdot,(M,N)) \in \mathcal{D}_F$ and $S_t(H(\cdot,(M,N))) \in \mathcal{D}_F$. Applying the generator G in (2.28) with $g = \frac{1}{2} dg_{\mathrm{FW}}$ to (2.37), we find

$$
\begin{aligned}
&G\big(H\big(\cdot,(M,N)\big)\big)(X,Y) \\
&= \sum_{i \in \mathbb{G}} \Bigg\{ \left[\sum_{k \in \mathbb{G}} a(i,k)\,(x_k - x_i) \right] \frac{\partial}{\partial x_i} \Big(\prod_{j \in \mathbb{G}} x_j^{m_j} y_j^{n_j} \Big) \\
&\quad + \frac{1}{2} d\, x_i (1 - x_i) \frac{\partial^2}{\partial x_i^2} \Big(\prod_{j \in \mathbb{G}} x_j^{m_j} y_j^{n_j} \Big) + K e\,(y_i - x_i) \frac{\partial}{\partial x_i} \Big(\prod_{j \in \mathbb{G}} x_j^{m_j} y_j^{n_j} \Big) \\
&\quad + e\,(x_i - y_i) \frac{\partial}{\partial y_i} \Big(\prod_{j \in \mathbb{G}} x_j^{m_j} y_j^{n_j} \Big) \Bigg\} \\
&= \sum_{i \in \mathbb{G}} \Bigg\{ \left[\sum_{k \in \mathbb{G}} m_i a(i,k) \prod_{\substack{j \in \mathbb{G} \\ j \neq i \\ j \neq k}} x_j^{m_j} y_j^{n_j} \left(x_i^{m_i-1} y_i^{n_i} x_k^{m_k+1} y_k^{n_k} - x_i^{m_i} y_i^{n_i} x_k^{m_l} y_k^{n_k} \right) \right]
\end{aligned}
\tag{4.16}
$$

$$+ \prod_{\substack{j \in \mathbb{G} \\ j \neq i}} x_j^{m_j} y_j^{n_j} \tfrac{d}{2} \, m_i(m_i - 1) \left(x_i^{m_i - 1} y_i^{n_i} - x_i^{m_i} y_i^{n_i} \right) 1_{\{m_i \geq 2\}}$$

$$+ m_i K e \prod_{\substack{j \in \mathbb{G} \\ j \neq i}} x_j^{m_j} y_j^{n_j} \left(x_i^{m_i - 1} y_i^{n_i + 1} - x_i^{m_i} y_i^{n_i} \right)$$

$$+ n_i e \prod_{\substack{j \in \mathbb{G} \\ j \neq i}} x_j^{m_j} y_j^{n_j} \left(x_i^{m_i + 1} y_i^{n_i - 1} - x_i^{m_i} y_i^{n_i} \right) \Bigg\}$$

$$= F\big(H\big((X, Y), \cdot\big)\big)(M, N).$$

Consequently, it follows from the generator criterion that

$$\mathbb{E}_{(X,Y)}\big[H\big((X(t), Y(t)), (M, N)\big)\big] = \mathbb{E}_{(M,N)}\big[H\big((X, Y), (M(t), N(t))\big)\big]. \tag{4.17}$$

This settles Theorem 2.5. □

Model 2: Proof of Theorem 2.8.

Proof. Theorem 2.8 follows after replacing in the above proof the block-counting process in (2.36) by the one in (2.46), the duality function by the one in (2.47), and checking the generator criterion. □

Model 3: Proof of Theorem 2.10.

Proof. Theorem 2.10 follows after replacing the block-counting process in (2.36) by the one in (2.57), the duality function is by the one in (2.47), and checking the generator criterion. □

4.3 Dichotomy criterion

In this section we prove Theorems 2.11 and 2.13.

Model 1: Proof of Theorem 2.11.

Proof.

"$\Longleftarrow$" The proof uses the duality relation in Theorem 2.5. Abbreviate

$$\theta_x = \mathbb{E}_{\mu(0)}[x_0], \qquad \theta_y = \mathbb{E}_{\mu(0)}[y_0]. \tag{4.18}$$

Note that, since $\mu(0)$ is invariant under translations, we have $\mathbb{E}_{\mu(0)}[x_i] = \theta_x$ and $\mathbb{E}_{\mu(0)}[y_i] = \theta_y$ for all $i \in \mathbb{G}$. We proceed as in [BGCKWB16, Proposition 2.9]. Abbreviate

$$m(t) = \sum_{i \in \mathbb{G}} m_i(t), \qquad n(t) = \sum_{i \in \mathbb{G}} n_i(t). \tag{4.19}$$

Let $(m(0), n(0)) \in E'$ be such that $m(0) + n(0) < \infty$, and put

$$T = \inf \left\{ t \geq 0 \colon\ m(t) + n(t) = 1 \right\}. \tag{4.20}$$

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

$$
\lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\prod_{i \in \mathbb{G}} x_i(t)^{m_i} y_i(t)^{n_i} \right]
$$

$$
= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(x(0),y(0))} \left[\prod_{i \in \mathbb{G}} x_i(t)^{m_i} y_i(t)^{n_i} \right] \right]
$$

$$
= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(m(0),n(0))} \left[\prod_{i \in \mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \right] \right]
$$

$$
= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[E_{(m(0),n(0))} \left[\prod_{i \in \mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \,\Big|\, T < \infty \right] \right] \mathbb{P}(T < \infty)
$$

$$
+ \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(m(0),n(0)} \left[\prod_{i \in \mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \,\Big|\, T = \infty \right] \right] \mathbb{P}(T = \infty) \qquad (4.21)
$$

$$
= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(m(0),n(0))} \left[\prod_{i \in \mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \,\Big|\, T < \infty,\ m(t) = 1,\ n(t) = 0 \right] \right]
$$

$$
\times\, \mathbb{P}(m(t) = 1,\ n(t) = 0)
$$

$$
+ \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(m(0),n(0))} \left[\prod_{i \in \mathbb{G}} x_i^{m_i(t)} y_i^{n_i(t)} \,\Big|\, T < \infty,\ m(t) = 0,\ n(t) = 1 \right] \right]
$$

$$
\times\, \mathbb{P}(m(t) = 0,\ n(t) = 1)
$$

$$
= \theta_x \frac{1}{1 + K} + \theta_y \frac{K}{1 + K},
$$

where in the last step we use that a single lineage in the dual behaves like the Markov chain with transition kernel $b^{(1)}(\cdot, \cdot)$ defined in (2.34). It follows from (4.21) that, for all $i, j \in \mathbb{G}$,

$$
\lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\frac{x_i(t) + K y_i(t)}{1 + K} \left(1 - \frac{x_j(t) + K y_j(t)}{1 + K} \right) \right] = 0. \qquad (4.22)
$$

Hence, either $\lim_{t \to \infty}(x(t), y(t)) = (0,0)^{\mathbb{G}}$ or $\lim_{t \to \infty}(x(t), y(t)) = (1,1)^{\mathbb{G}}$. Computing $\lim_{t \to \infty} \mathbb{E}_{\mu(0)}[x_i(t)]$ with the help of (4.21), we find

$$
\lim_{t \to \infty} \mu(t) = (1 - \theta) \, [\delta_{(0,0)}]^{\otimes \mathbb{G}} + \theta \, [\delta_{(1,1)}]^{\otimes \mathbb{G}} \qquad (4.23)
$$

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

"$\Longrightarrow$" Suppose that the systems clusters. Then (4.23) holds, which means that

$$
\lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[z_u(t) \left(1 - z_v(t) \right) \right] = 0 \qquad \forall\, u, v \in \mathbb{S}. \qquad (4.24)
$$

Let

$$
|L(t)| = \sum_{u \in \mathbb{S}} \ell_u(t) \qquad (4.25)
$$

be the total number of lineages left at time t. Applying the duality relation in (2.41) to

(4.24), we find

$$
\begin{aligned}
0 &= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[z_u(t)(1 - z_v(t)) \right] \\
&= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u} \left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \right] \right] - \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u + \delta_v} \left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \right] \right] \\
&= \frac{\theta_x + K\theta_y}{1 + K} \left[1 - \lim_{t \to \infty} \mathbb{P}_{\delta_u + \delta_v} \left(|L(t)| = 1 \right) \right] \\
&\quad - \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u + \delta_v} \left[\prod_{w \in \mathbb{S}} z_u^{\ell_w(t)} \,\bigg|\, |L(t)| = 2 \right] \right] \mathbb{P}_{\delta_u + \delta_v}(|L(t)| = 2),
\end{aligned}
\tag{4.26}
$$

where $\mathbb{P}_{\delta_u}$ is the law of the block process starting from a single partition element in u, and $\mathbb{P}_{\delta_u + \delta_v}$ is the law of the block process starting from two partition elements in u and v. As to the last term in the right-hand side of (4.26), we note that

$$
\begin{aligned}
&\lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u + \delta_v} \left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\bigg|\, |L(t)| = 2 \right] \right] \\
&= \lim_{t \to \infty} \frac{1}{(1 + K)^2} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u + \delta_v} \left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\bigg|\, L(t) = \delta_{(i,A)} + \delta_{(j,A)}, \, i, j \in \mathbb{G} \right] \right] \\
&\quad + \lim_{t \to \infty} \frac{2K}{(1 + K)^2} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u + \delta_v} \left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\bigg|\, L(t) = \delta_{(i,A)} + \delta_{(j,D)}, \, i, j \in \mathbb{G} \right] \right] \\
&\quad + \lim_{t \to \infty} \frac{K^2}{(1 + K)^2} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{\delta_u + \delta_v} \left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\bigg|\, L(t) = \delta_{(i,D)} + \delta_{(j,D)}, i, j \in \mathbb{G} \right] \right] \\
&< \frac{\theta_x}{(1 + K)^2} + \frac{K\theta_x + K\theta_y}{(1 + K)^2} + \frac{K^2\theta_y}{(1 + K)^2} = \frac{\theta_x + K\theta_y}{1 + K} = \theta.
\end{aligned}
\tag{4.27}
$$

Here, the strict inequality follows from the non-trivial invariant initial distribution (ruling out $z \equiv 0$ and $z \equiv 1$), together with the fact that the swapping between active and dormant is driven by a positive recurrent Markov chain on $\{A, D\}$. Hence (4.24) holds if and only if $\lim_{t \to \infty} \mathbb{P}_{\delta_u + \delta_v}(|L(t)| = 2 \,|\, |L(0)| = 2) = 0$ for every $u, v \in \mathbb{S}$. Therefore every pair of lineages coalesces with probability 1.

Thus, we have proved Theorem 2.11.

Model 2: Proof of Theorem 2.13.

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

$$
\theta_x = \mathbb{E}_{\mu(0)}[x_0], \qquad \theta_{y,m} = \mathbb{E}_{\mu(0)}[y_{0,m}], \qquad \theta = \frac{\theta_x + \sum_{m \in \mathbb{N}_0} K_m \theta_{y,m}}{1 + \rho}.
\tag{4.28}
$$

For $\rho < \infty$, a lineage in the dual moves as a positive recurrent Markov chain on $\{A, (D_m)_{m \in \mathbb{N}_0}\}$. Therefore the argument for "$\Longleftarrow$" given for Model 1 goes through via the duality relation, which gives (compare with (4.21))

$$
\begin{aligned}
\lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(x(0),y(0))} \left[\prod_{u \in \mathbb{S}} z_u(t)^{\ell_u} \right] \right] &= \lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[\mathbb{E}_{(m(0),n(0))} \left[\prod_{u \in \mathbb{S}} z_u^{\ell_u(t)} \right] \right] \\
&= \frac{\theta_x + \sum_{m \in \mathbb{N}_0} K_m \theta_{y,m}}{1 + \sum_{m \in \mathbb{N}_0} K_m}.
\end{aligned}
\tag{4.29}
$$

With the duality relation in (2.50), the argument for "$\Longrightarrow$" given for Model 1 also goes through directly.

Case $\rho = \infty$. For $\rho = \infty$, a lineage in the dual moves as a null-recurrent Markov chain, which has no stationary distribution, and so (4.29) does not carry over. However, from [Lin92, Section 3] it follows that, for all $u_1, u_2 \in \mathbb{S}$,

$$\lim_{t \to \infty} \left\| \mathbb{P}_{\delta_{u_1}}(L(t) \in \cdot \mid |L(t)| = 1) - \mathbb{P}_{\delta_{u_2}}(L(t) \in \cdot \mid |L(t)| = 1) \right\|_{tv} = 0. \qquad \forall v \in \mathbb{S}. \quad (4.30)$$

Moreover, by null-recurrence, for all $v \in \mathbb{S}$,

$$\lim_{t \to \infty} \mathbb{P}_{\delta_v}\big(\exists\, i \in \mathbb{G}\colon L(t) = \delta_{(i,A)}\big) = 0,$$
$$\lim_{t \to \infty} \mathbb{P}_{\delta_v}\big(\exists\, i \in \mathbb{G}\colon L(t) = \delta_{(i,D_m)}\big) = 0 \qquad \forall m \in \mathbb{N}_0, \quad (4.31)$$
$$\lim_{t \to \infty} \sum_{m \geq M} \mathbb{P}_{\delta_v}\big(\exists\, i \in \mathbb{G}\colon L(t) = \delta_{(i,D_m)}\big) = 1 \qquad \forall M \in \mathbb{N}_0.$$

"$\Longleftarrow$" By duality, we have

$$\lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{(x(0),y(0))}\left[\prod_{u \in \mathbb{S}} z_u(t)^{l_u}\right]\right] = \lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{(m(0),n(0))}\left[\prod_{u \in \mathbb{S}} z_u^{\ell_u(t)}\right]\right]$$
$$= \lim_{t \to \infty} \theta_x \mathbb{P}_{(m(0),n(0))}\big(\exists\, i \in \mathbb{G}\colon L(t) = \delta_{(i,A)}\big) \quad (4.32)$$
$$+ \lim_{t \to \infty} \sum_{m \in \mathbb{N}_0} \theta_{y,m} \mathbb{P}_{(m(0),n(0))}\big(\exists\, i \in \mathbb{G}\colon L(t) = \delta_{(i,D_m)}\big),$$

where we follow an argument similar as in (4.21) and use that $\mathbb{P}(T < \infty) = 1$. Because the initial measure is colour regular, we know that $\lim_{m \to \infty} \theta_{y,m} = \theta$ (recall Definition 2.12). Hence (4.31)–(4.32) imply, after we let $M \to \infty$, that all moments tend to θ. In particular,

$$\lim_{t \to \infty} \mathbb{E}_{\mu(0)}[x_i(t)] = \theta = \lim_{t \to \infty} \mathbb{E}_{\mu(0)}[y_{i,m}(t)], \qquad i \in \mathbb{G},\ m \in \mathbb{N}_0. \quad (4.33)$$

"$\Longrightarrow$" By the duality relation in (2.50) and the assumption of clustering, we have

$$\lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[z_u(t)(1 - z_v(t))\right] = 0 \qquad \forall u, v \in \mathbb{S}. \quad (4.34)$$

Therefore

$$0 = \lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[z_u(t)(1 - z_v(t))\right]$$
$$= \lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u}\left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)}\right]\right] - \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u + \delta_v}\left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)}\right]\right]$$
$$= \theta\left[1 - \lim_{t \to \infty} \mathbb{P}_{\delta_u + \delta_v}\left(|L(t)| = 1\right)\right] \quad (4.35)$$
$$- \lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u + \delta_v}\left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\Big|\, |L(t)| = 2\right]\right] \mathbb{P}_{\delta_u + \delta_v}(|L(t)| = 2),$$

where the last equality uses (4.31) and colour regularity. Suppose that $\lim_{t \to \infty} \mathbb{P}_{\delta_u + \delta_v}(|L(t)| = 2) \neq 0$. Then $\lim_{t \to \infty} \mathbb{E}_{\delta_u + \delta_v}\left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \mid |L(t)| = 2\right] = \theta$, because otherwise (4.35) would be violated. However,

$$\lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u + \delta_v}\left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\Big|\, |L(t)| = 2\right]\right] < \lim_{t \to \infty} \mathbb{E}_{\mu(0)}\left[\mathbb{E}_{\delta_u + \delta_v}\left[\prod_{w \in \mathbb{S}} z_w^{\ell_w(t)} \,\Big|\, |L(t)| = 1\right]\right]$$
$$= \theta, \quad (4.36)$$

because we start from a nontrivial stationary distribution.

Thus, we have proved Theorem 2.13.

Model 3: Proof of Theorem 2.13. Since the duality relation for Model 3 is exactly the same as for Model 2, the same results hold by translation invariance and the extra displacement does not affect the dichotomy criterion. □

4.4 Outline remainder of paper

In Sections 5–7 we prove Theorems 3.1, 3.3 and 3.8, respectively. For each of the three models we split the proof into four parts:

1. Moment relations.
2. The clustering case.
3. The coexistence case.
4. Proof of the dichotomy.

5 Proofs: long-time behaviour for Model 1

In Section 5.1 we relate the first and second moments of the process $(Z(t))_{t \geq 0}$ in (2.4)–(2.5) to the random walk with internal states $\{A, D\}$ that evolves according to the transition kernel $b^{(1)}(\cdot, \cdot)$ given in (2.34) (Lemma 5.1 below). These moment relations hold for all $g \in \mathcal{G}$. In Section 5.2 we deal with the clustering case (Lemmas 5.4–5.5 below), in Section 5.3 with the coexistence case (Lemmas 5.7–5.13 below). In Section 5.4 we prove Theorem 3.1. In Sections 5.2 and 5.3 we will see that the moment relations are crucial when no duality is available.

Below we write $\mathbb{E}_z$ instead of $\mathbb{E}_{\delta_z}$ for the expectation given that $Z(0) = \delta_z$, $z \in E$.

5.1 Moment relations

Lemma 5.1 (First and second moment). *For $z \in E$, $t \geq 0$ and $(i, R_i), (j, R_j) \in \mathbb{G} \times \{A, D\}$,*

$$\mathbb{E}_z[z_{(i,R_i)}(t)] = \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}\big((i, R_i), (k, R_k)\big) z_{(k,R_k)} \tag{5.1}$$

and

$$\begin{aligned}
&\mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\
&= \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}\big((i, R_i), (k, R_k)\big) b_t^{(1)}\big((j, R_j), (l, R_l)\big) z_{(k,R_k)} z_{(l,R_l)} \\
&\quad + 2 \int_0^t \mathrm{d}s \sum_{k \in \mathbb{G}} b_{(t-s)}^{(1)}((i, R_i), (k, A)) b_{(t-s)}^{(1)}((j, R_j), (k, A)) \mathbb{E}_z[g(x_k(s))].
\end{aligned} \tag{5.2}$$

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

$$\begin{aligned}
(G_{RW} f)(i, R_i) &= \sum_{(j,R_j) \in \mathbb{G} \times \{A,D\}} b^{(1)}\big((i, R_i), (j, R_j)\big) [f(j, R_j) - f(i, R_i)], \\
f &: \mathbb{G} \times \{A, D\} \to \mathbb{R}.
\end{aligned} \tag{5.3}$$

Applying the generator (2.28) of the system in (2.4)–(2.5) to the function $f_{(i,R_i)} \colon E \to \mathbb{R}$, $f_{(i,R_i)}(z) = z_{(i,R_i)}$, we obtain by standard stochastic calculus

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}_z[z_{(i,R_i)}(t)]}{\mathrm{d}t} &= \left[\sum_{j \in \mathbb{G}} a(i,j) \big(\mathbb{E}_z[x_j(t)] - \mathbb{E}_z[x_i(t)] \big) + Ke \big(\mathbb{E}_z[y_i(t)] - \mathbb{E}_z[x_i(t)] \big) \right] 1_{(R_i = A)} \\
&\quad + e \big(\mathbb{E}_z[x_i(t)] - \mathbb{E}_z[y_i(t)] \big) 1_{(R_i = D)}.
\end{aligned} \tag{5.4}$$

Hence, denoting by $(S_t)_{t \geq 0}$ the semigroup of the system in (2.4)–(2.5), we see from (5.4) and the definition of $b^{(1)}(\cdot, \cdot)$ in (2.34) that $(S_t f_{(i,R_i)})$ solves the differential equation

$$F'(t) = (G_{RW} F)(t). \tag{5.5}$$

On the other hand, for each $f \in \mathcal{C}_b(\mathbb{G} \times \{A, D\})$, $RW_t f$ also solves (5.5). In particular, if for $z \in \mathbb{E}$ we define $f_z \colon \mathbb{G} \times \{A, D\} \to \mathbb{R}$ by $f_z(i, R_i) = z(i, R_i)$, then $RW_t f_z$ is a solution to (5.5). Since

$$(RW_0 f_z)(i, R_i) = z_{(i,R_i)} = (S_0 f_{(i,R_i)})(z), \tag{5.6}$$

we see that (5.1) holds. To prove (5.2), we derive a similar system of differential equations and again solve this in terms of the random walk moving according to the kernel $b(\cdot, \cdot)$. Let $f \colon E \to \mathbb{R}$ be given by $f(z) = z_{(i,R_i)} z_{(j,R_j)}$. Using the generator (2.28), we obtain via Itô-calculus that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] = &\sum_{k \in \mathbb{G}} a(i,k) \left(\mathbb{E}_z[x_k(t) z_{(j,R_j)}(t)] - \mathbb{E}_z[x_i(t) z_{(j,R_j)}(t)] \right) 1_{\{R_i = A\}} \\
&+ Ke \left(\mathbb{E}_z[y_i(t) z_{(j,R_j)}(t)] - \mathbb{E}_z[x_i(t) z_{(j,R_j)}(t)] \right) 1_{\{R_i = A\}} \\
&+ e \left(\mathbb{E}_z[x_i(t) z_{(j,R_j)}(t)] - \mathbb{E}_z[y_i(t) z_{(j,R_j)}(t)] \right) 1_{\{R_i = D\}} \\
&+ \sum_{l \in \mathbb{G}} a(j,l) \left(\mathbb{E}_z[x_l(t) z_{(i,R_i)}(t)] - \mathbb{E}_z[x_j(t) z_{(i,R_i)}(t)] \right) 1_{\{R_j = A\}} \\
&+ Ke \left(\mathbb{E}_z[y_j(t) z_{(i,R_i)}(t)] - \mathbb{E}_z[x_j(t) z_{(i,R_i)}(t)] \right) 1_{\{R_j = A\}} \\
&+ e \left(\mathbb{E}_z[x_j(t) z_{(i,R_i)}(t)] - \mathbb{E}_z[y_j(t) z_{(i,R_i)}(t)] \right) 1_{\{R_j = D\}} \\
&+ 2\, \mathbb{E}_z[g(x_i(t))] 1_{\{i=j\}} 1_{\{R_i = R_j = A\}}.
\end{aligned} \tag{5.7}$$

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

$$\begin{aligned}
(Uh)((i, R_i), (j, R_j)) = &\sum_{k \in \mathbb{G}} a(i,k) \left[h((k, A), (j, R_j)) - h((i, R_i), (j, R_j)) \right] 1_{\{i, R_i = A\}} \\
&+ Ke \left[h((i, D), (j, R_j)) - h((i, R_i), (j, R_j)) \right] 1_{\{i, R_i = A\}} \\
&+ e \left[h((i, A), (j, R_j)) - h((i, R_i), (j, R_j)) \right] 1_{\{i, R_i = D\}} \\
&+ \sum_{l \in \mathbb{G}} a(j,l) \left[h((i, R_i), (l, A)) - h((i, R_i), (j, R_j)) \right] 1_{\{R_j = A\}} \\
&+ Ke \left[h((i, R_i), (j, D)) - h((i, R_i), (j, R_j)) \right] 1_{\{R_j = A\}} \\
&+ e \left[h((i, R_i), (j, A)) - h((i, R_i), (j, D)) \right] 1_{\{R_j = D\}}.
\end{aligned} \tag{5.8}$$

Let $F(t) = \mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)]$ and $H(t) = 2\mathbb{E}_z[g(x_i(t))] 1_{\{i=j\}} 1_{\{R_i = R_j = A\}}$. Then we can rewrite (5.7) as

$$\frac{\mathrm{d}}{\mathrm{d}t} F(t) = (UF)(t) + H(t). \tag{5.9}$$

Denote by $(RW_t^{(2)})_{t \geq 0}$ the semigroup corresponding to U. Applying [Lig85, Theorem I.2.15], we obtain

$$F(t) = RW_t^{(2)} F(0) + \int_0^t \mathrm{d}s\, RW_{t-s}^{(2)} H(s). \tag{5.10}$$

Hence

$$\begin{aligned}
&\mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\
&= \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big) b_t^{(1)}\big((j,R_j),(l,R_l)\big) \mathbb{E}_z[z_{(k,R_k)} z_{(l,R_l)}] \\
&\quad + 2 \int_0^t \mathrm{d}s \sum_{k \in \mathbb{G}} b_{t-s}^{(1)}\big((i,R_i),(k,A)\big) b_{t-s}^{(1)}\big((j,R_j),(k,A)\big) \mathbb{E}_z[g(x_k(s))]. \quad \square
\end{aligned} \tag{5.11}$$

Remark 5.2 (Density). From Lemma 5.1 we obtain that if μ is a translation invariant measure such that $\mathbb{E}_\mu[x_0] = \theta_x$ and $\mathbb{E}_\mu[y_0] = \theta_y$, then

$$\mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta_x \sum_{(k,R_k) \in \mathbb{G} \times \{A\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big) + \theta_y \sum_{(k,R_k) \in \mathbb{G} \times \{D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big),$$
(5.12)

in particular, $\lim_{t \to \infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \frac{\theta_x + K\theta_y}{1+K} = \theta$, recall (3.4), and

$$\begin{aligned}
&\mathbb{E}_\mu[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\
&= \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}\big((i,R_i),(k,R_k)\big)\, b_t^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_\mu[z_{(k,R_k)} z_{(l,R_l)}] \\
&\quad + 2 \int_0^t \mathrm{d}s \sum_{k \in \mathbb{G}} b_{t-s}^{(1)}\big((i,R_i),(k,A)\big)\, b_{t-s}^{(1)}\big((j,R_j),(k,A)\big)\, \mathbb{E}_\mu[g(x_i(s))].
\end{aligned}$$
(5.13)

$\square$

Remark 5.3 (First moment duality). Note that (5.1) shows that even for general $g \in \mathcal{G}$ there is a *first moment duality* between the process $Z(t)$ and the random walk $RW(t)$, that moves according to the kernel $b^{(1)}(\cdot,\cdot)$. The duality function is given by

$$H : E \times \mathbb{G} \times \{A,D\} \to \mathbb{R}, \qquad H(z,(i,R_i)) = z_{(i,R_i)}.$$
(5.14)

Equation (5.1) in Lemma 5.1 tells us that

$$\mathbb{E}_{Z(0)}[H(Z(t),RW(0))] = \mathbb{E}_{RW(0)}[H(Z(0),RW(t))].$$
(5.15)

5.2 The clustering case

The proof that the system in (2.4)–(2.5) converges to a unique trivial equilibrium when $\hat{a}(\cdot,\cdot)$ is recurrent goes as follows. We first consider the case where $g = dg_{\mathrm{FW}}$, for which *duality* is available (Lemma 5.4). Afterwards we use a *duality comparison argument* to show that the dichotomy between coexistence and clustering does not depend on the choice of $g \in \mathcal{G}$ (Lemma 5.5).

• **Case** $g = dg_{\mathbf{FW}}$.

Lemma 5.4 (Clustering). *Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ and $g = dg_{\mathrm{FW}}$. Moreover, suppose that $\hat{a}(\cdot,\cdot)$ defined in (3.1) is recurrent, i.e., $I_{\hat{a}} = \infty$. Let $\mu(t)$ be the law at time t of the system defined in (2.4)–(2.5). Then*

$$\lim_{t \to \infty} \mu(t) = \theta\,[\delta_{(1,1)}]^{\otimes \mathbb{G}} + (1-\theta)\,[\delta_{(0,0)}]^{\otimes \mathbb{G}}.$$
(5.16)

Proof. Since $g = dg_{\mathrm{FW}}$, we can use duality. By the dichotomy criterion in Theorem 2.11, it is enough to show that in the dual two partition elements coalesce with probability 1. Recall from Section 2.4 that each of the partition elements in the dual moves according to the transition kernel $b^{(1)}(\cdot,\cdot)$ on $\mathbb{G} \times \{A,D\}$ defined by (2.34) (see Fig. 4). Recall from Section (2.4) that $b^{(1)}(\cdot,\cdot)$ describes a random walk on $\mathbb{G}$ with migration rate kernel $a(\cdot,\cdot)$ that becomes dormant (state D) at rate Ke (after which it stops moving), and becomes active (state A) at rate e (after which it can move again). When two partition elements in the dual are active and are at the same site, they coalesce at rate d, i.e., each time they are active and meet at the same site they coalesce with probability $d/[\sum_{j \in \mathbb{Z}^d} a(i,j) + Ke + d] > 0$. Hence, in order to show that two partition elements coalesce with probability 1, we have to show that *with probability 1 two partition elements meet infinitely often while being active*. The latter holds if and only

if the expected total time the random walks spend together at the same colony while being active is infinite. We will show that this occurs if and only the random walk with *symmetrised* transition rate kernel $\hat{a}(\cdot, \cdot)$ is recurrent. The proof comes in 4 Steps.

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

$$(\sigma_k)_{k \in \mathbb{N}}, \quad (\sigma'_k)_{k \in \mathbb{N}}, \tag{5.17}$$

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

$$(\tau_k)_{k \in \mathbb{N}}, \quad (\tau'_k)_{k \in \mathbb{N}}, \tag{5.18}$$

denote the successive time lapses during which they are dormant (see Fig. 5). These are mutually independent sequences of i.i.d. random variables with marginal laws

$$\begin{aligned} \mathbb{P}(\sigma_1 > t) &= \mathbb{P}(\sigma'_1 > t) &= \mathrm{e}^{-Ket}, && t \geq 0, \\ \mathbb{P}(\tau_1 > t) &= \mathbb{P}(\tau'_1 > t) &= \mathrm{e}^{-et} && t \geq 0, \end{aligned} \tag{5.19}$$

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

σ_1	τ_1	σ_2	τ_2	σ_3	τ_3
σ'_1	τ'_1	σ'_2	τ'_2		σ'_3	τ'_3
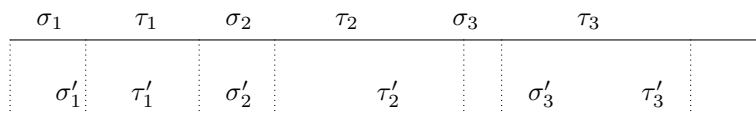
Figure 5: Successive periods during which the two random walks are active and dormant. The time lapses between the dotted lines represent periods of joint activity.

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

$$\begin{aligned} \mathcal{E}(k,t) &= \left\{ \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) \leq t < \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) + \sigma_{k+1} \right\}, \\ \mathcal{E}'(k',t) &= \left\{ \sum_{\ell=1}^{k'} (\sigma'_\ell + \tau'_\ell) \leq t < \sum_{\ell=1}^{k'} (\sigma'_\ell + \tau'_\ell) + \sigma'_{k+1} \right\}, \end{aligned} \tag{5.20}$$

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

$$\begin{aligned} T(k,t) &= \sum_{\ell=1}^{k} \sigma_\ell + \left(\left(t - \sum_{\ell=1}^{k} (\sigma_\ell + \tau_\ell) \right) \wedge \sigma_{k+1} \right), \\ T'(k',t) &= \sum_{\ell=1}^{k'} \sigma'_\ell + \left(\left(t - \sum_{\ell=1}^{k'} (\sigma'_\ell + \tau'_\ell) \right) \wedge \sigma_{k+1} \right), \end{aligned} \tag{5.21}$$

be the total accumulated activity times up to time t given that k, k' transitions between dormant and active have taken place. Note that the terms between brackets in (5.21) are at most σ_{k+1}, respectively, $\sigma'_{k'+1}$, and therefore are negligible as $k, k' \to \infty$.

Given the outcome of the sequences in (5.17)–(5.18), the probability that at time t both random walks are active and are at the same colony equals

$$\sum_{k,k'\in\mathbb{N}} \left(\sum_{i\in\mathbb{G}} a_{T(k,t)}(0,i)\, a_{T'(k',t)}(0,i) \right) 1_{\mathcal{E}(k,t)}\, 1_{\mathcal{E}(k',t)}. \qquad (5.22)$$

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

$$I = \int_0^\infty \mathrm{d}t \sum_{k,k'\in\mathbb{N}} \mathbb{E}_{(0,A),(0,A)} \left[\left(\sum_{i\in\mathbb{G}} a_{T(k,t)}(0,i)\, a_{T'(k',t)}(0,i) \right) 1_{\mathcal{E}(k,t)}\, 1_{\mathcal{E}'(k',t)} \right], \qquad (5.23)$$

where $\mathbb{E}$ is the expectation over the sequences in (5.17). Let

$$N(t) = \max\left\{ k \in \mathbb{N}\colon \sum_{\ell=1}^{k}(\sigma_\ell + \tau_\ell) \le t \right\}, \qquad N'(t) = \max\left\{ k' \in \mathbb{N}\colon \sum_{\ell=1}^{k'}(\sigma_\ell + \tau_\ell) \le t \right\}, \qquad (5.24)$$

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

$$T(t) = T(N(t),t), \quad T'(t) = T'(N'(t),t), \quad \mathcal{E}(t) = \mathcal{E}(N(t),t), \quad \mathcal{E}'(t) = \mathcal{E}'(N'(t),t), \quad (5.25)$$

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

$$I = \int_0^\infty \mathrm{d}t\, \mathbb{E}_{(0,A),(0,A)} \left[\left(\sum_{i\in\mathbb{G}} a_{T(t)}(0,i)\, a_{T'(t)}(0,i) \right) 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)} \right]. \qquad (5.26)$$

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

2. Fourier analysis. Define

$$M(t) = T(t) \wedge T'(t), \qquad \Delta(t) = [T(t) \vee T'(t)] - [T(t) \wedge T'(t)]. \qquad (5.27)$$

Then

$$\sum_{i\in\mathbb{G}} a_{T(t)}(0,i)\, a_{T'(t)}(0,i) = \sum_{j\in\mathbb{G}} \hat{a}_{2M(t)}(0,j)\, a_{\Delta(t)}(j,0). \qquad (5.28)$$

Indeed, the difference of the two random walks at time $M(t)$ has distribution $\hat{a}_{2M(t)}(0,\cdot)$, and in order for the random walk with the largest activity time to meet the random walk with the smallest activity time at time $2M(t) + \Delta(t)$, it must bridge this difference in time $\Delta(t)$. To work out (5.28), we assume without loss of generality that $\sum_{j\in\mathbb{G}} a(0,j) = 1$, and use *Fourier analysis*. For ease of exposition we focus on the special case where $\mathbb{G} = \mathbb{Z}^d$, but the argument below extends to *any* countable Abelian group endowed with the discrete topology, because these properties ensure that there is a version of Fourier analysis on $\mathbb{G}$ [Rud62, Section 1.2]. For $\phi \in [-\pi,\pi]^d$, define

$$a(\phi) = \sum_{j\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}(\phi,j)} a(0,j), \qquad \hat{a}(\phi) = \operatorname{Re} a(\phi), \qquad \tilde{a}(\phi) = \operatorname{Im} a(\phi). \qquad (5.29)$$

Then, by Fourier inversion,

$$\hat{a}_t(0,j) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi\, \mathrm{e}^{-\mathrm{i}(\phi,j)}\, \mathrm{e}^{-t[1-\hat{a}(\phi)]},$$

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

$$(5.30)$$

where we use that $a(\phi) = \hat{a}(\phi) + i\tilde{a}(\phi)$. Inserting these representations into (5.28), we get

$$\sum_{i \in \mathbb{Z}^d} a_{T(t)}(0, i) \, a_{T'(t)}(0, i) = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-[2M(t)+\Delta(t)] \, [1-\hat{a}(\phi)]} \cos(\Delta(t)\tilde{a}(\phi)), \quad (5.31)$$

where we use that $\sum_{j \in \mathbb{Z}^d} \mathrm{e}^{i(\phi'-\phi,j)} = (2\pi)^d \delta(\phi' - \phi)$, with $\delta(\cdot)$ the Dirac distribution (Folland [Fol92, Chapter 7]).

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

$$\lim_{k \to \infty} \frac{1}{k} \sum_{\ell=1}^{k} \sigma_\ell = \frac{1}{Ke} \quad \mathbb{P}\text{-a.s.}, \qquad \lim_{k \to \infty} \frac{1}{k} \sum_{\ell=1}^{k} \tau_\ell = \frac{1}{e} \quad \mathbb{P}\text{-a.s.} \qquad (5.32)$$

Therefore, by the standard renewal theorem (Asmussen [Asm03, Chapter I, Theorem 2.2]),

$$\lim_{t \to \infty} \frac{1}{t} N(t) = \lim_{t \to \infty} \frac{1}{t} N'(t) = A \quad \mathbb{P}\text{-a.s.},$$

$$\lim_{t \to \infty} \frac{1}{t} T(t) = \lim_{t \to \infty} \frac{1}{t} T'(t) = B \quad \mathbb{P}\text{-a.s.}, \qquad (5.33)$$

$$\lim_{t \to \infty} \mathbb{P}\big(\mathcal{E}(t)\big) = \lim_{t \to \infty} \mathbb{P}\big(\mathcal{E}'(t)\big) = B,$$

with

$$A = \frac{1}{\frac{1}{Ke} + \frac{1}{e}} = \frac{K}{1+K} \, e, \qquad B = \frac{\frac{1}{Ke}}{\frac{1}{Ke} + \frac{1}{e}} = \frac{1}{1+K}. \qquad (5.34)$$

Moreover, by the central limit theorem, we have

$$\left(\frac{T(t) - Bt}{c\sqrt{t}}, \frac{T'(t) - Bt}{c\sqrt{t}} \right) \quad \Longrightarrow \quad (Z, Z') \quad \text{in } \mathbb{P}\text{-distribution as } t \to \infty \qquad (5.35)$$

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

$$c^2 = A \left[(1-B)^2 \operatorname{Var}(\sigma_1) + B^2 \operatorname{Var}(\tau_1) \right] \qquad (5.36)$$

(see [Smi55] or [Asm03, Theorem VI.3.2]). Since $T(t), \mathcal{E}(t)$ and $T'(t), \mathcal{E}'(t)$ are independent, and each pair is asymptotically independent as well, we find that

$$\mathbb{E}_{(0,A),(0,A)} \left[\left(\sum_{i \in \mathbb{Z}^d} a_{T(t)}(0, i) \, a_{T'(t)}(0, i) \right) 1_{\mathcal{E}(t)} 1_{\mathcal{E}'(t)} \right] \sim B^2 f(t), \qquad t \to \infty, \qquad (5.37)$$

with

$$\begin{aligned}
f(t) &= \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-[1+o(1)] \, 2Bt \, [1-\hat{a}(\phi)]} \, \mathbb{E} \left[\cos \left([1+o(1)] \, c(Z - Z')\sqrt{t} \, \tilde{a}(\phi) \right) \right] \\
&= \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-[1+o(1)] \, 2Bt \, [1-\hat{a}(\phi)]} \, \mathrm{e}^{-[1+o(1)] \, c^2 t \, \tilde{a}(\phi)^2},
\end{aligned} \qquad (5.38)$$

where we use that $\cos$ is symmetric, $Z - Z' = \sqrt{2} \, Z''$ in $\mathbb{P}$-distribution with Z'' standard normal, and $\mathbb{E}(\mathrm{e}^{i\mu Z''}) = \mathrm{e}^{-\mu^2/2}$, $\mu \in \mathbb{R}$. From (5.26) and (5.37) we have that $I < \infty$ if and only if $t \mapsto f(t)$ is integrable. By Cramér's theorem, deviations of $T(t)/t$ and $T'(t)/t$ away from B are exponentially costly in t. Hence the error terms in (5.38), arising from (5.33) and (5.35), do *not* affect the integrability of $t \mapsto f(t)$. Note that, because $a(\cdot, \cdot)$ is assumed to be *irreducible* (recall (2.1)), $\hat{a}(\phi) = 1$ if and only if $\phi = 0$. Hence the integrability of $t \mapsto f(t)$ is determined by the behaviour of $\hat{a}(\phi)$ and $\tilde{a}(\phi)$ as $\phi \to 0$.

4. Irrelevance of asymmetric part of migration. We next observe that $\tilde{a}(\phi)^2 \le 1 - \hat{a}(\phi)^2 \le 2[1 - \hat{a}(\phi)]$. Hence, $t\,\tilde{a}(\phi)^2 \le 2t\,[1 - \hat{a}(\phi)]$. Therefore we see from (5.38) that for sufficiently large $T \in \mathbb{R}$ we can bound $t \mapsto f(t)$ on $[T, \infty)$ from above and below by functions of the form $t \mapsto g_C(t)$ with

$$g_C(t) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{d}\phi \, \mathrm{e}^{-Ct\,[1-\hat{a}(\phi)]}, \qquad C \in (0, \infty). \tag{5.39}$$

From (5.30) we have

$$g_C(t) = \hat{a}_{Ct}(0,0) \asymp \hat{a}_t(0,0), \tag{5.40}$$

where the last asymptotics uses that $t \mapsto \hat{a}_t(0,0)$ is regularly varying at infinity (recall (3.2)). Combining (5.26), (5.37) and (5.39)–(5.40), we get

$$I = \infty \quad \Longleftrightarrow \quad I_{\hat{a}} = \infty \tag{5.41}$$

with $I_{\hat{a}} = \int_1^\infty \mathrm{d}t \, \hat{a}_t(0,0)$. Thus, if $\hat{a}(\cdot, \cdot)$ is recurrent, then $I = \infty$ and the system clusters. Moreover, we see from the bounds on $f(t)$ (recall (5.38)) that *the asymmetric part of the migration kernel has no effect on the integrability*.

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

• **Case** $g \ne dg_{\mathbf{FW}}$**.** For $g \ne dg_{\mathrm{FW}}$ the proof of Lemma 5.4 does not go through. However, the *moment relations* in Lemma 5.1 hold for general $g \in \mathcal{G}$. Using these moment relations and a technique called duality comparison (see [CG94]), we prove Lemma 5.4 for general $g \in \mathcal{G}$.

Lemma 5.5 (Duality comparison). *Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ and $g \in \mathcal{G}$. Moreover, suppose that $\hat{a}(\cdot, \cdot)$ defined in (3.1) is recurrent, i.e., $I_{\hat{a}} = \infty$. Let $\mu(t)$ be the law at time t of the system defined in (2.4)–(2.5). Then*

$$\lim_{t \to \infty} \mu(t) = \theta \, [\delta_{(1,1)}]^{\otimes \mathbb{G}} + (1 - \theta) \, [\delta_{(0,0)}]^{\otimes \mathbb{G}}. \tag{5.42}$$

Proof. We proceed as in the proof of [CG94, Theorems 1–2]. First assume that $\mu(0) = \delta_z$ for some $z \in E$, which satisfies

$$\lim_{t \to \infty} \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}((i, R_i), (k, R_k)) \, z_{(k,R_k)} = \theta. \tag{5.43}$$

By Lemma 5.1, we have

$$\mathbb{E}_z \left[z_{(i,R_i)}(t) \right] = \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}((i, R_i), (k, R_k)) \, z_{(k,R_k)}. \tag{5.44}$$

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

$$\lim_{t \to \infty} \mathbb{E}_z \left[z_{(i,R_i)}(t) \right] = \theta. \tag{5.45}$$

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

$$\lim_{t \to \infty} \mathbb{E}_z \left[z_{(i,R_i)}(t)(1 - z_{(j,R_j)}(t)) \right] = 0, \tag{5.46}$$

because this implies that

$$i = j \colon \lim_{t \to \infty} \mathbb{P}_z \left(z_{(i,R_i)}(t) \in [0, \epsilon) \cup (1 - \epsilon, 1] \right) = 1 \quad \forall \, \epsilon > 0,$$
$$i \ne j \colon \lim_{t \to \infty} \mathbb{P}_z \left(z_{(j,R_j)}(t) \in (1 - \epsilon, 1] \mid \lim_{t \to \infty} z_{(i,R_i)}(t) \in (1 - \epsilon, 1] \right) = 1 \quad \forall \, \epsilon > 0. \tag{5.47}$$

Thus we are left to prove that

$$\lim_{t \to \infty} \mathbb{E}_z \left[z_{(i,R_i)}(t) z_{(j,R_j)}(t) \right] = \theta. \tag{5.48}$$

Since (5.45) implies that $\limsup_{t \to \infty} \mathbb{E}_z[z_{(i,R_i)} z_{(j,R_j)}] \leq \theta$, we are left to prove that

$$\liminf_{t \to \infty} \mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \geq \theta. \tag{5.49}$$

Like in [CG94], we will prove (5.49) by *comparison duality*.

Fix $\epsilon > 0$. Since $g \in \mathcal{G}$ we can choose a $c = c(\epsilon) > 0$ such that $g(x) \geq \tilde{g}(x) = c(x - \epsilon)(1 - (x + \epsilon))$, $x \in [0,1]$. Note that $\tilde{g}(x) < 0$ for $x \in [0, \epsilon) \cup (1 - \epsilon, 1]$, so we cannot replace g by $\tilde{g}$ in the SSDE. Instead we use $\tilde{g}$ as an auxiliary function.

Consider the Markov chain $(B(t))_{t \geq 0}$, with state space $\{1, 2\} \times (\mathbb{G} \times \{A, D\}) \times (\mathbb{G} \times \{A, D\})$ and $B(t) = (B_0(t), B_1(t), B_2(t))$, evolving according to

$$(1, (i, R_i), (i, R_i)) \to (1, (k, R_k), (k, R_k)), \quad \text{at rate } b^{(1)}((i, R_i), (k, R_k)),$$

$$(2, (i, R_i), (j, R_j)) \to \begin{cases} (2, (k, R_k), (j, R_j)), & \text{at rate } b^{(1)}((i, R_i), (k, R_k)), \\ (2, (i, R_i), (l, R_l)), & \text{at rate } b^{(1)}((j, R_j), (l, R_l)), \\ (1, (i, R_i), (i, R_i)), & \text{at rate } c 1_{\{i=j\}} 1_{\{R_i = R_j = A\}}. \end{cases} \tag{5.50}$$

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

$$\mathbb{P}_{(2,(i,R_i),(j,R_j))} \left(B_1(t) = (k, R_k) \right) = b_t^{(1)}((i, R_i), (k, R_k)), \tag{5.51}$$

and similarly

$$\mathbb{P}_{(2,(i,R_i),(j,R_j))} \left(B_2(t) = (l, R_l) \right) = b_t^{(1)}((j, R_j), (l, R_l)). \tag{5.52}$$

By the proof of Lemma 5.4, since we have assumed that $\hat{a}(\cdot, \cdot)$ is recurrent, i.e., $I_{\hat{a}} = \infty$, the two random walks meet infinitely often at the same site while being active and hence in the dual coalesce with probability 1. Therefore

$$\lim_{t \to \infty} \mathbb{P}_{(2,(i,R_i),(j,R_j))} \left(B_0(t) = 2 \right) = 0. \tag{5.53}$$

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

$$\mathrm{d}z_{(i,R_i)}(t) = \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b^{(1)}((i, R_i), (j, R_j))[z_{(j,R_j)}(t) - z_{(i,R_i)}(t)] \, \mathrm{d}t$$
$$+ \sqrt{g(z_{(i,R_i)}(t))} \, 1_{\{R_i = A\}} \, \mathrm{d}w_i(t). \tag{5.54}$$

Using (5.54) and Itô-calculus, we obtain

$$\frac{\mathrm{d}\mathbb{E}_z[z_{(i,R_i)}(t) - \epsilon]}{\mathrm{d}t} = \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b^{(1)}((i, R_i), (k, R_k)) \, \mathbb{E}_z \left[(z_{(k,R_k)}(t) - \epsilon) - (z_{(i,R_i)}(t) - \epsilon) \right]$$
$$\tag{5.55}$$

and

$$
\begin{aligned}
\frac{\mathrm{d}\mathbb{E}_z[(z_{(i,R_i)}(t) - \epsilon)(z_{(j,R_j)}(t) + \epsilon)]}{\mathrm{d}t} \\
= \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} & b^{(1)}((i,R_i),(k,R_k)) \\
& \times \mathbb{E}_z\left[(z_{(j,R_j)}(t) + \epsilon)(z_{(k,R_k)}(t) - \epsilon) - (z_{(j,R_j)}(t) + \epsilon)(z_{(i,R_i)}(t) - \epsilon)\right] \\
+ \sum_{(l,R_l)\in\mathbb{G}\times\{A,D\}} & b^{(1)}((j,R_j),(k,R_k)) \\
& \times \mathbb{E}_z\left[(z_{(i,R_i)}(t) - \epsilon)(z_{(l,R_l)}(t) + \epsilon) - (z_{(i,R_i)}(t) - \epsilon)(z_{(j,R_j)}(t) + \epsilon)\right] \\
+ \mathbb{E}_z & \left[c(z_{(i,R_i)}(t) - \epsilon)(1 - (z_{(j,R_j)}(t) + \epsilon))\,1_{\{i=j\}}1_{\{R_i=R_j=A\}}\right] \\
+ \mathbb{E}_z & \left[\left(g(z_{(i,R_i)}(t)) - \tilde{g}(z_{(i,R_i)}(t))\right)1_{\{i=j\}}1_{\{R_i=R_j=A\}}\right].
\end{aligned}
\tag{5.56}
$$

For $t \geq 0$, define $F_t\colon \{1,2\} \times (\mathbb{G} \times \{A,D\}) \times (\mathbb{G} \times \{A,D\}) \to \mathbb{R}$ by

$$
\begin{aligned}
F_t(1,(i,R_i),(i,R_i)) &= \mathbb{E}_z\left[z_{(i,R_i)}(t) - \epsilon\right] \\
F_t(2,(i,R_i),(j,R_j)), &= \mathbb{E}_z\left[(z_{(i,R_i)}(t) - \epsilon)(z_{(j,R_j)}(t) + \epsilon)\right],
\end{aligned}
\tag{5.57}
$$

and $H_t\colon \{1,2\} \times (\mathbb{G} \times \{A,D\}) \times (\mathbb{G} \times \{A,D\}) \to \mathbb{R}$ by

$$
\begin{aligned}
H_t(1,(i,R_i),(i,R_i)) &= 0, \\
H_t(2,(i,R_i),(j,R_j)) &= \mathbb{E}_z\left[\left(g(z_{(i,R_i)}(t)) - \tilde{g}(z_{(i,R_i)}(t))\right)1_{\{i=j\}}\,1_{\{R_i=R_j=A\}}\right].
\end{aligned}
\tag{5.58}
$$

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

$$
\frac{\mathrm{d}F_t}{\mathrm{d}t} = \mathfrak{B}F_t + H_t.
\tag{5.59}
$$

Hence, by [Lig85, Theorem I.2.15], it follows that

$$
F_t = V_t F_0 + \int_0^t \mathrm{d}s\, V_{(t-s)} H_s.
\tag{5.60}
$$

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

$$
\begin{aligned}
F_t&(2,(i,R_i),(j,R_j)) \geq V_t F_0(2,(i,R_i),(j,R_j)) \\
&= \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))\right] \\
&= \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))1_{\{B_0(t)=1\}} + F_0(B(t))\,1_{\{B_0(t)=2\}}\right] \\
&= \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_0(t) = 1, B_1(t) = (k,R_k)\right](z_{(k,R_k)} - \epsilon) \\
&\quad + \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))1_{\{B_0(t)=2\}}\right] \\
&= \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_1(t) = (k,R_k)\right](z_{(k,R_k)} - \epsilon) \\
&\quad - \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} \mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_0(t) = 2, B_1(t) = (k,R_k)\right](z_{(k,R_k)} - \epsilon) \\
&\quad + \mathbb{E}_{(2,(i,R_i),(j,R_j))}\left[F_0(B(t))1_{\{B_0(t)=2\}}\right] \\
&\geq \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_t^{(1)}((i,R_i),(k,R_k))\,(z_{(k,R_k)} - \epsilon) - (1 + \epsilon^2)\,\mathbb{P}_{(2,(i,R_i),(j,R_j))}\left[B_0(t) = 2\right].
\end{aligned}
\tag{5.61}
$$

Hence, by (5.43), (5.53) and (5.57), we obtain

$$
\liminf_{t\to\infty} \mathbb{E}_z\left[(z_{(i,R_i)}(t) - \epsilon)(z_{(j,R_j)}(t) + \epsilon)\right] \geq \theta - \epsilon.
\tag{5.62}
$$

Letting $\epsilon \downarrow 0$, we get (5.48).

To get rid of the assumption $\mu(0) = \delta_z$, note that for $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ we have, by Remark 5.2,

$$\lim_{t \to \infty} \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b_t((i,R_i),(k,R_k)) \mathbb{E}_{\mu(0)}[z_{(k,R_k)}] = \theta. \tag{5.63}$$

Hence, by the above argument,

$$\mathbb{E}_{\mu(0)} \left[(z_{(i,R_i)}(t) - \epsilon)(z_{j,R_j}(t) + \epsilon) \right] = \int \mathbb{E}_z \left[(z_{(i,R_i)}(t) - \epsilon)(z_{j,R_j}(t) + \epsilon) \right] \mu(0)(\mathrm{d}z)$$

$$\geq \int \sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}((i,R_i),(k,R_k))(z_{(k,R_k)} - \epsilon)$$

$$- (1 + \epsilon^2) \mathbb{P}_{(2,(i,R_i),(j,R_j))} \left[B_1(t) = 2 \right] \mu(0)(\mathrm{d}z). \tag{5.64}$$

Letting first $t \to \infty$ and then $\epsilon \downarrow 0$ we find that, for all $(i,R_i),(j,R_j) \in \mathbb{G} \times \{A,D\}$,

$$\lim_{t \to \infty} \mathbb{E}_{\mu(0)} \left[z_{(i,R_i)}(t)(1 - z_{(j,R_j)}(t)) \right] = 0. \tag{5.65}$$

$\square$

5.3 The coexistence case

For the coexistence case we proceed as in [CG94] with small adaptations. For the convenience of the reader we have written out the full proof. The proof relies on the moment relations in Lemma 5.1 and no distinction between $g = d g_{\mathrm{FW}}$ and general $g \in \mathcal{G}$ is needed. The proof consist of several lemmas (Lemmas 5.7–5.13 below), organised into 4 Steps. In Step 1 we use the moment relations in Lemma 5.1 to define a set of measures that are preserved under the evolution. In Step 2 we use coupling to prove that, for each given θ, the system converges to a unique equilibrium. In Step 3 we show that, for each given θ, each initial measure under the evolution converges to an invariant measure. In Step 4 we show that the limiting measure is invariant, ergodic and mixing under translations, and is associated.

1. Properties of measures preserved under the evolution. Let θ be defined as in (3.4) such that $\theta = \mathbb{E}_{\mu(0)} \left[\frac{x_0 + K y_0}{1+K} \right] = \frac{\theta_x + K \theta_y}{1+K}$.

Definition 5.6 (Preserved class of measure). Let $\mathcal{R}_\theta^{(1)}$ denote the set of measures $\mu \in \mathcal{T}$ satisfying:

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

$$\lim_{t \to \infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta. \tag{5.66}$$

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

$$\lim_{t \to \infty} \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}((i,R_i),(k,R_k)) \, b_t^{(1)}((j,R_j),(l,R_l)) \, \mathbb{E}_\mu[z_{(k,R_k)} z_{(l,R_l)}] = \theta^2. \tag{5.67}$$

$\square$

Clearly, if $\mu \in \mathcal{R}_\theta^{(1)}$, then (1) and (2), together with Lemma 5.1 and Remark 5.2, imply

$$\lim_{t \to \infty} \mathbb{E}_\mu \left[\left(\sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b_t^{(1)}((i,R_i),(k,R_k)) \, z_{(k,R_k)} - \theta \right)^2 \right] = 0, \tag{5.68}$$

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

On the other hand, suppose that (5.68) holds for *some* $(i, R_i) \in \mathbb{G} \times \{A, D\}$. Then, by Lemma 5.1 and Remark 5.2, we can rewrite (5.68) as

$$\lim_{t\to\infty} \mathbb{E}_\mu \left[\left(\mathbb{E}_z[z_{(i,R_i)}(t)] - \theta \right)^2 \right] = 0. \tag{5.69}$$

This implies

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta, \tag{5.70}$$

and hence, by translation invariance,

$$\lim_{t\to\infty} \mathbb{E}_\mu[z_{(k,R_i)}(t)] = \theta \qquad \forall\, k \in \mathbb{G}. \tag{5.71}$$

Using that switches between the active state at the dormant state occur at a positive rate, we can use the strong Markov property to obtain that (5.71) holds both for $R_i = A$ and for $R_i = D$. Hence (5.66) holds. Combining (5.66) and (5.68), we see that also (5.67) holds.

Lemma 5.7. $\mathcal{T}_\theta^{\mathrm{erg}} \subset \mathcal{R}_\theta^{(1)}$ *(recall (3.6)).*

Proof. The proof relies on *Fourier analysis* and the existence of spectral measures. As in Section 5.2, for ease of exposition we focus on the special case where $\mathbb{G} = \mathbb{Z}^d$, but the argument below extends to *any* countable Abelian group endowed with the discrete topology.

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

$$\begin{aligned} \mathbb{E}_\mu\left[(x_j - \theta_x)(x_k - \theta_x)\right] &= \int_{(-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}(j-k,\phi)} \mathrm{d}\lambda_A(\phi), \\ \mathbb{E}_\mu\left[(y_j - \theta_y)(y_k - \theta_y)\right] &= \int_{(-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}(j-k,\phi)} \mathrm{d}\lambda_D(\phi). \end{aligned} \tag{5.72}$$

Let $a(\phi) = \sum_{k \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i}(\phi,j)} a(0,k)$ be the characteristic function of the kernel $a(\cdot, \cdot)$ (recall (5.29)), and $T(t)$ the activity time of the random walk up to time t (recall (5.21)). Then

$$\begin{aligned} \sum_{k \in \mathbb{Z}^d} a_{T(t)}(0,k)\, \mathrm{e}^{\mathrm{i}(k,\phi)} &= \sum_{n \in \mathbb{N}_0} \frac{\mathrm{e}^{-T(t)}[T(t)]^n}{n!} \sum_{k \in \mathbb{Z}^d} a^n(0,k)\, \mathrm{e}^{\mathrm{i}(k,\phi)} \\ &= \sum_{n \in \mathbb{N}_0} \frac{\mathrm{e}^{-T(t)}[T(t)\, a(\phi)]^n}{n!} \\ &= \mathrm{e}^{-T(t)(1-a(\phi))}. \end{aligned} \tag{5.73}$$

Let $\mathcal{E}(t)$ be defined as in (5.25). Then, for fixed $t > 0$,

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

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

$$= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\sum_{k,l \in \mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))\,b_t^{(1)}((0,A),(l,A))\,\mathbb{E}_\mu\left[(x_k - \theta_x)(x_l - \theta_x)\right]$$

$$= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\sum_{k,l \in \mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))\,b_t^{(1)}((0,A),(l,A))\int_{(-\pi,\pi]^d} e^{i(k-l,\phi)}\mathrm{d}\lambda_A(\phi)$$

$$= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\sum_{k,l \in \mathbb{Z}^d}$$

$$\times\,\mathbb{E}_{(0,A),(0,A)}\left[a_{T(t)}(0,k)\,a_{T'(t)}(0,l)\,1_{\mathcal{E}(t)}1_{\mathcal{E}'(t)}\right]\int_{(-\pi,\pi]^d} e^{i(k-l,\phi)}\mathrm{d}\lambda_A(\phi)$$

$$= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}$$

$$\times\int_{(-\pi,\pi]^d}\mathbb{E}_{(0,A),(0,A)}\left[\sum_{k \in \mathbb{Z}^d} a_{T(t)}\,e^{i(k,\phi)}(0,k)1_{\mathcal{E}(t)}\sum_{l \in \mathbb{Z}^d} a_{T'(t)}\,e^{-i(l,\phi)}(0,l)1_{\mathcal{E}'(t)}\right]\mathrm{d}\lambda_A(\phi)$$

$$= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2}\int_{(-\pi,\pi]^d}\mathbb{E}_{(0,A),(0,A)}\left[e^{-T(t)(1-a(\phi))}1_{\mathcal{E}(t)}\,e^{-T'(t)(1-\bar{a}(\phi))}1_{\mathcal{E}'(t)}\right]\mathrm{d}\lambda_A(\phi).$$

$$(5.74)$$

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

$$\lim_{t \to \infty}\mathbb{E}_\mu\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\sum_{k \in \mathbb{Z}^d} b_t^{(1)}((0,A),(k,A))x_k - \theta_x\right)^2\right] = \lambda_A(\{0\}). \qquad (5.75)$$

Similarly,

$$\lim_{t \to \infty}\mathbb{E}_\mu\left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))}\sum_{k \in \mathbb{Z}^d} b_t^{(1)}((0,A),(k,D))y_k - \theta_y\right)^2\right] = \lambda_D(\{0\}). \qquad (5.76)$$

Hence

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

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

$$\left.\left.+\,\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))\sum_{k \in \mathbb{Z}^d}\frac{b_t^{(1)}((0,A),(k,D))}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))}y_k - \frac{K\theta_y}{1+K}\right)^2\right]$$

$$\leq \lim_{t \to \infty}\mathbb{P}_{(0,A)}(\mathcal{E}(t))\,\mathbb{E}_\mu\left[\left(\sum_{k \in \mathbb{Z}^d}\frac{b_t^{(1)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}x_k - \frac{\theta_x}{(1+K)}\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\right)^2\right] \quad (5.77)$$

Spatial populations with seed-bank: well-posedness, duality and equilibrium

$$+ \mathbb{P}_{(0,A)}(\mathcal{E}^c(t)) \, \mathbb{E}_\mu \left[\left(\sum_{k \in \mathbb{Z}^d} \frac{b_t^{(1)}((0,A),(k,D))}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))} \, y_k - \frac{K\theta_y}{1+K} \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}^c(t))} \right)^2 \right]$$

$$= \frac{1}{1+K} \lambda_A(\{0\}) + \frac{K}{1+K} \lambda_D(\{0\}).$$

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

Let $\Lambda_N = [0,N)^d \cap \mathbb{Z}^d$. By the L^2-ergodic theorem, we have, for $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$,

$$\lim_{N \to \infty} \mathbb{E}_\mu \left[\left(\frac{1}{\Lambda_N} \sum_{j \in \Lambda_N} x_j - \theta_x \right)^2 \right] = 0. \tag{5.78}$$

(For general $\mathbb{G}$ note that countable groups endowed with the discrete topology are amenable. For amenable groups $\mathbb{G}$, $(\Lambda_N)_{N \in \mathbb{N}}$ must be replaced by a so-called Følner sequence, i.e., a sequence of finite subsets of $\mathbb{G}$ that exhaust $\mathbb{G}$ and satisfy $\lim_{N \to \infty} |\mathfrak{g}\Lambda_N \triangle \Lambda_N|/|\Lambda_N| = 0$ for any $\mathfrak{g} \in \mathbb{G}$ [Lin99].) Using the spectral measure, we can write

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

$$= \lim_{N \to \infty} \frac{1}{\Lambda_N^2} \sum_{j,k \in \Lambda_N} \int_{(-\pi,\pi]^d} e^{i(j-k,\phi)} \mathrm{d}\lambda_A \tag{5.79}$$

$$= \lim_{N \to \infty} \int_{(-\pi,\pi]^d} \left(\frac{1}{\Lambda_N} \sum_{j \in \Lambda_N} e^{i(j,\phi)} \right) \left(\frac{1}{\Lambda_N} \sum_{k \in \Lambda_N} e^{-i(k,\phi)} \right) \mathrm{d}\lambda_A$$

$$= \lambda_A\{0\}.$$

In the last equality we use dominated convergence and

1. For all $\phi \in (-\pi,\pi]^d$,

$$\lim_{N \to \infty} \frac{1}{\Lambda_N} \sum_{j,k \in \Lambda_N} e^{-i(k,\phi)} = 1_{\{0\}}(\phi). \tag{5.80}$$

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

$$\left| \frac{1}{\Lambda_N} \sum_{j,k \in \Lambda_N} e^{-i(k,\phi)} - 1_{\{0\}}(\phi) \right| \le 1_{J_\delta}(\phi) + \epsilon(N,\delta), \tag{5.81}$$

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

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

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

Lemma 5.8 (Preservation). *If $b(\cdot,\cdot)$ is transient and $\mu \in \mathcal{R}_\theta^{(1)}$, then the following hold:*

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

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

Proof. Our dynamics preserve translation invariance. To check property (1) of $\mathcal{R}_\theta^{(1)}$ (see (5.66)), set $f(z) = z_{(i,R_i)}$. Since $\mu \in \mathcal{R}_\theta^{(1)}$, applying Lemma 5.1 multiple times, we obtain

$$
\begin{aligned}
\lim_{s\to\infty} \mathbb{E}_{\mu S_t}[z_{(i,R_i)}(s)] &= \lim_{s\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, \mathbb{E}_{\mu S_t}[z_{(k,R_k)}] \\
&= \lim_{s\to\infty} \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, \mathbb{E}_{\mu}[z_{(k,R_k)}(t)] \\
&= \lim_{s\to\infty} \sum_{(k',R'_{k'})\in\mathbb{G}\times\{A,D\}} b_{s+t}^{(1)}\big((i,R_i),(k',R_{k'})\big)\, \mathbb{E}_{\mu}[z_{(k',R_{k'})}] \\
&= \lim_{s\to\infty} \mathbb{E}_{\mu}[z_{(i,R_i)}(t+s)] = \theta.
\end{aligned}
\tag{5.82}
$$

To check property (2) of $\mathcal{R}_\theta^{(1)}$ (see (5.67)), we set $f(z) = z_{(i,R_i)} z_{(j,R_j)}$. Then, again by applying Lemma 5.1, we find

$$
\begin{aligned}
&\lim_{s\to\infty} \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, b_s^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_{\mu S_t}[z_{(k,R_k)} z_{(l,R_l)}] \\
&= \lim_{s\to\infty} \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, b_s^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_{\mu}[z_{(k,R_k)}(t) z_{(l,R_l)}(t)] \\
&= \lim_{s\to\infty} \Bigg[\sum_{(k',R_{k'}),(l',R_{l'})\in\mathbb{G}\times\{A,D\}} b_{t+s}^{(1)}\big((i,R_i),(k',R_{k'})\big)\, b_{t+s}^{(1)}\big((j,R_j),(l',R_{l'})\big)\, \mathbb{E}_{\mu}[z_{(k',R_{k'})} z_{(l',R_{l'})}] \\
&\quad + 2\int_0^t dr \sum_{k'\in\mathbb{G}} b_{t-r+s}^{(1)}\big((i,R_i),(k',A)\big)\, b_{t-r+s}^{(1)}\big((j,R_j),(k',A)\big)\, \mathbb{E}_{\mu}[g(x_{k'}(r))] \Bigg].
\end{aligned}
\tag{5.83}
$$

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

$$
\lim_{s\to\infty} \int_s^{t+s} du \sum_{k'\in\mathbb{G}} b_u^{(1)}\big((i,R_i),(k',A)\big)\, b_u^{(1)}\big((j,R_j),(k',A)\big)\, \mathbb{E}_{\mu}[g(x_{k'}(t+s-u))] = 0.
\tag{5.84}
$$

Using the notation of Section 5.2, we get ($\|g\|$ denotes the supremum norm of g)

$$
\begin{aligned}
&\lim_{s\to\infty} \int_s^{t+s} du \sum_{k'\in\mathbb{G}} b_u^{(1)}\big((i,R_i),(k',A)\big)\, b_u^{(1)}\big((j,R_j),(k',A)\big)\, \mathbb{E}_{\mu}[g(x_{k'}(t+s-u))] \\
&\leq \|g\| \lim_{s\to\infty} \int_s^{t+s} du \sum_{k'\in\mathbb{G}} b_u^{(1)}\big((i,R_i),(k',A)\big)\, b_u^{(1)}\big((j,R_j),(k',A)\big) \\
&= \|g\| \lim_{s\to\infty} \int_s^{t+s} du\, \mathbb{E}_{(i,R_i),(j,R_j)}\Bigg[\sum_{k'\in\mathbb{G}} a_{T(u)}(i,k')\, \mathbf{1}_{\mathcal{E}(u)}\, a_{T'(u)}(j,k')\, \mathbf{1}_{\mathcal{E}'(u)} \Bigg] \\
&\leq \|g\| \lim_{s\to\infty} \int_s^{t+s} du\, \mathbb{E}_{(0,A),(0,A)}\Bigg[\sum_{k'\in\mathbb{G}} a_{T(u)}(i,k')\, \mathbf{1}_{\mathcal{E}(u)}\, a_{T'(u)}(j,k')\, \mathbf{1}_{\mathcal{E}'(u)} \Bigg] = 0,
\end{aligned}
\tag{5.85}
$$

where the last equality follows from the assumption $I_{\hat{a}} < \infty$ in Theorem 3.1, (5.23) and (5.41). The last inequality follows from the Markov property and the observation that, in order to get a contribution to the integral, the two random walks first have to meet at the same site and both be active. We conclude that $\mu S_t \in \mathcal{R}_\theta^{(1)}$ for all $t \geq 0$.

To show that $\mu(\infty) \in \mathcal{R}_\theta^{(1)}$, we proceed like in (5.82), to obtain

$$
\lim_{s\to\infty} \mathbb{E}_{\mu(\infty)}[z_{(i,R_i)}(s)] = \lim_{s\to\infty} \lim_{n\to\infty} \mathbb{E}_{\mu S_{t_n}}[z_{(i,R_i)}(s)] = \lim_{s\to\infty} \lim_{n\to\infty} \mathbb{E}_{\mu}[z_{(i,R_i)}(t_n+s)] = \theta,
\tag{5.86}
$$

and so (5.66) is satisfied. To get (5.67), we note that, by Lemma 5.1,

$$
\begin{aligned}
&\sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_{t_n}^{(1)}\big((i,R_i),(k,R_k)\big)\, b_{t_n}^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}] \\
&\leq \mathbb{E}_\mu[z_{(i,R_i)}(t_n)z_{(j,R_j)}(t_n)] \\
&\leq \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_{t_n}^{(1)}\big((i,R_i),(k,R_k)\big)\, b_{t_n}^{(1)}\big((j,R_j),(l,R_l)\big)\, \mathbb{E}_\mu[z_{(k,R_k)}z_{(l,R_l)}] \\
&\quad + 2\|g\|\int_0^{t_n} \mathrm{d}s \sum_{k\in\mathbb{G}} b_{t_n-s}^{(1)}\big((i,R_i),(k,A)\big)\, b_{t_n-s}^{(1)}\big((j,R_j),(k,A)\big).
\end{aligned}
\tag{5.87}
$$

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

$$
\begin{aligned}
\theta^2 &\leq \mathbb{E}_{\mu(\infty)}[z_{(i,R_i)}z_{(j,R_j)}] \\
&\leq \theta^2 + 2\|g\|\int_0^\infty \mathrm{d}r \sum_{k\in\mathbb{G}} b_r^{(1)}\big((i,R_i),(k,A)\big)\, b_r^{(1)}\big((j,R_j),(k,A)\big).
\end{aligned}
\tag{5.88}
$$

Inserting (5.88) into (5.67), we see that it is enough to show that

$$
\begin{aligned}
&\lim_{s\to\infty} \sum_{(k,R_k),(l,R_l)\in\mathbb{G}\times\{A,D\}} b_s^{(1)}\big((i,R_i),(k,R_k)\big)\, b_s^{(1)}\big((j,R_j),(l,R_l)\big) \\
&\qquad\times\, 2\|g\|\int_0^\infty \mathrm{d}r \sum_{k'\in\mathbb{G}} b_r^{(1)}((k,R_k),(k',A))\, b_r^{(1)}((l,R_l),(k',A)) \\
&= \lim_{s\to\infty} 2\|g\|\int_0^\infty \mathrm{d}r \sum_{k'\in\mathbb{G}} b_{r+s}^{(1)}\big((i,R_i),(k',A)\big)\, b_{r+s}^{(1)}\big((j,R_j),(k',A)\big) = 0.
\end{aligned}
\tag{5.89}
$$

However, from the assumption $I_{\hat{a}}<\infty$ in Theorem 3.1, (5.23) and (5.41), we have

$$
\begin{aligned}
&\lim_{s\to\infty} \|g\|\int_0^\infty \mathrm{d}r \sum_{k'\in\mathbb{G}} b_{r+s}^{(1)}\big((i,R_i),(k',A)\big)\, b_{r+s}^{(1)}\big((j,R_j),(k',A)\big) \\
&= \lim_{s\to\infty} \|g\|\int_s^\infty \mathrm{d}r\, \mathbb{E}_{(i,R_i),(j,R_j)}\left[\sum_{k'\in\mathbb{G}} a_{T(r)}(i,k')\mathbf{1}_{\mathcal{E}(r)} a_{T'(r)}(j,k')\mathbf{1}_{\mathcal{E}'(r)}\right] = 0. \quad\square
\end{aligned}
\tag{5.90}
$$

2. Uniqueness of the equilibrium. In this section we show that, for given θ, the equilibrium when it exists is unique. To prove this we extend the coupling argument in [CG94]. Consider two copies of the system (2.4)–(2.5) *coupled via their Brownian motions*:

$$
\begin{aligned}
\mathrm{d}x_i^k(t) &= \sum_{j\in\mathbb{G}} a(i,j)\,[x_j^k(t)-x_i^k(t)]\,\mathrm{d}t + \sqrt{g(x_i^k(t))}\,\mathrm{d}w_i(t) \\
&\quad + Ke\,[y_i^k(t)-x_i^k(t)]\,\mathrm{d}t,
\end{aligned}
\tag{5.91}
$$

$$
\mathrm{d}y_i^k(t) = e\,[x_i^k(t)-y_i^k(t)]\,\mathrm{d}t, \qquad k\in\{1,2\}.
\tag{5.92}
$$

Here, k labels the copy, and the two copies are driven by the same set of independent Brownian motions $(w_i(t))_{t\geq0}$, $i\in\mathbb{G}$. As initial probability distributions we choose $\mu^1(0)$ and $\mu^2(0)$ that are both invariant and ergodic under translations.

Let

$$
\bar{z}_i(t) = (z_i^1(t), z_i^2(t)), \qquad z_i^k(t) = (x_i^k(t), y_i^k(t)), \quad k\in\{1,2\}.
\tag{5.93}
$$

The coupled system $(\bar{z}_i(t))_{i\in\mathbb{G}}$ has a unique strong solution (the marginals are the single-component systems, which by [SS80, Theorem 3.2] have a unique strong solution). Write $\hat{\mathbb{P}}$ to denote the law of the coupled system, and let $\Delta_i(t) = x_i^1(t)-x_i^2(t)$ and $\delta_i(t) = y_i^1(t)-y_i^2(t)$.

Lemma 5.9 (Coupling dynamics). *For every $t \geq 0$,*

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big] = -2 \sum_{j \in \mathbb{G}} a(i,j) \, \hat{\mathbb{E}}\left[|\Delta_j(t)| \, 1_{\{\mathrm{sgn}\,\Delta_i(t) \neq \mathrm{sgn}\,\Delta_j(t)\}}\right]$$
$$- 2Ke\,\hat{\mathbb{E}}\left[\big(|\Delta_i(t)| + |\delta_i(t)|\big) \, 1_{\{\mathrm{sgn}\,\Delta_i(t) \neq \mathrm{sgn}\,\delta_i(t)\}}\right]. \tag{5.94}$$

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

$$\mathrm{d}|\Delta_i(t)| = \mathrm{sgn}\,\Delta_i(t)\,\mathrm{d}\Delta_i(t) + \mathrm{d}L_t^0,$$
$$\mathrm{d}\Delta_i(t) = \sum_{j \in \mathbb{G}} a(i,j)[\Delta_j(t) - \Delta_i(t)]\,\mathrm{d}t + \left[\sqrt{g(x_i^1(t))} - \sqrt{g(x_i^2(t))}\right]\,\mathrm{d}w_i(t) \tag{5.95}$$
$$+ Ke\left[\delta_i(t) - \Delta_i(t)\right]\mathrm{d}t,$$

where L_t^0 is the local time of $\Delta_i(t)$ at 0 (see [RW00, Section IV.43]). Next, we use that $\Delta_i(t)$ has zero local time at $x = 0$ because g is Lipschitz (see [RW00, Proposition V.39.3]). Taking expectation, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{\mathbb{E}}\big[|\Delta_i(t)|\big] = \sum_{j \in \mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\Big[\mathrm{sgn}\,\Delta_i(t)\,[\Delta_j(t) - \Delta_i(t)]\Big] + Ke\,\hat{\mathbb{E}}\Big[\mathrm{sgn}\,\Delta_i(t)\,[\delta_i(t) - \Delta_i(t)]\Big]. \tag{5.96}$$

Similarly, we have

$$\mathrm{d}|\delta_i(t)| = \mathrm{sgn}\,\delta_i(t)\,\mathrm{d}\delta_i(t),$$
$$\mathrm{d}\delta_i(t) = e\left[\Delta_i(t) - \delta_i(t)\right]\mathrm{d}t. \tag{5.97}$$

Taking expectation, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{\mathbb{E}}\big[|\delta_i(t)|\big] = e\,\hat{\mathbb{E}}\Big[\mathrm{sgn}\,\delta_i(t)\,[\Delta_i(t) - \delta_i(t)]\Big]. \tag{5.98}$$

Combining (5.96) and (5.98), we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big] = \sum_{j \in \mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\Big[\mathrm{sgn}\,\Delta_i(t)\,[\Delta_j(t) - \Delta_i(t)]\Big]$$
$$+ K\,e\,\hat{\mathbb{E}}\Big[[\mathrm{sgn}\,\Delta_i(t) - \mathrm{sgn}\,\delta_i(t)]\,[\delta_i(t) - \Delta_i(t)]\Big]. \tag{5.99}$$

Note that

$$\mathrm{sgn}\,\Delta_i(t)\,[\Delta_j(t) - \Delta_i(t)] = |\Delta_j(t)| - |\Delta_i(t)| - 2\,|\Delta_j(t)|\,1_{\{\mathrm{sgn}\,\Delta_i(t) \neq \mathrm{sgn}\,\Delta_j(t)\}}. \tag{5.100}$$

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

$$-2 \sum_{j \in \mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\Big[|\Delta_j(t)|\,1_{\{\mathrm{sgn}\,\Delta_i(t) \neq \mathrm{sgn}\,\Delta_j(t)\}}\Big]. \tag{5.101}$$

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

$$-2Ke\,\hat{\mathbb{E}}\left[\big(|\Delta_i(t)| + |\delta_i(t)|\big)\,1_{\{\mathrm{sgn}\,\Delta_i(t) \neq \mathrm{sgn}\,\delta_i(t)\}}\right]. \tag{5.102}$$

Combining (5.99) and (5.101)–(5.102), we get the claim. $\qquad\square$

Lemma 5.9 tells us that $t \mapsto \hat{\mathbb{E}}[|\Delta_i(t)| + K|\delta_i(t)|]$ is non-increasing, i.e., is a Lyapunov function. Therefore $\lim_{t \to \infty} \hat{\mathbb{E}}[|\Delta_i(t)| + K|\delta_i(t)|] = c_i \in [0, 1+K]$ exists. To show that the coupling is successful we need the following lemma.

Lemma 5.10 (Uniqueness of equilibrium). *If $a(\cdot,\cdot)$ is transient, then $c_i = 0$ for all $i \in \mathbb{G}$, and so the coupling is successful, i.e.,*

$$\lim_{t\to\infty} \hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big] = 0. \tag{5.103}$$

Proof. Write $-h_i(t)$ to denote the right-hand side of (5.94). We begin with the observation that $t \mapsto h_i(t)$ has the following properties:

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

Property (a) is evident. Property (b) follows from integration of (5.94):

$$\int_0^t \mathrm{d}s\, h_i(s) = \hat{\mathbb{E}}\big[|\Delta_i(0)| + K|\delta_i(0)|\big] - \hat{\mathbb{E}}\big[|\Delta_i(t)| + K|\delta_i(t)|\big]. \tag{5.104}$$

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

$$\begin{aligned}
&\forall\, i,j \in \mathbb{G} \text{ with } a(i,j) > 0: \\
&\quad \lim_{t\to\infty} \hat{\mathbb{P}}\Big(\{\Delta_i(t) < -\epsilon,\, \Delta_j(t) > \epsilon\} \cup \{\Delta_i(t) > \epsilon,\, \Delta_j(t) < -\epsilon\}\Big) = 0, \\
&\forall\, i \in \mathbb{G}: \\
&\quad \lim_{t\to\infty} \hat{\mathbb{P}}\Big(\{\Delta_i(t) < -\epsilon,\, \delta_i(t) > \epsilon\} \cup \{\Delta_i(t) > \epsilon,\, \delta_i(t) < -\epsilon\}\Big) = 0.
\end{aligned} \tag{5.105}$$

In Appendix C we will prove the following lemma:

Lemma 5.11 (Successful coupling). *For all $i,j \in \mathbb{G}$ and $\epsilon > 0$,*

$$\lim_{t\to\infty} \hat{\mathbb{P}}\Big(\{\Delta_i(t) < -\epsilon,\, \Delta_j(t) > \epsilon\} \cup \{\Delta_i(t) > \epsilon,\, \Delta_j(t) < -\epsilon\}\Big) = 0. \tag{5.106}$$

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

$$\begin{aligned}
E_0 \times E_0 = &\left\{ \bar{z} \in E \times E \colon z^1_{(i,R_i)} \geq z^2_{(i,R_i)}\ \forall (i, R_i) \in \mathbb{G} \times \{A, D\} \right\} \\
&\cup \left\{ \bar{z} \in E \times E \colon z^2_{(i,R_i)} \geq z^1_{(i,R_i)}\ \forall (i, R_i) \in \mathbb{G} \times \{A, D\} \right\}.
\end{aligned} \tag{5.107}$$

Then Lemma 5.11 together with (5.105) imply that $\lim_{t\to\infty} \hat{\mathbb{P}}\left(\bar{z}(t) \in E_0 \times E_0\right) = 1$, which we express by saying that "one diffusion lies on top of the other".

Using Lemma 5.11 we can complete the proof of the successful coupling. Let $t_n \to \infty$ as $n \to \infty$ and suppose, by possibly going to further subsequences, that $\lim_{n\to\infty} \mu^1(t_n) = \nu^1_\theta$ and $\lim_{n\to\infty} \mu^2(t_n) = \nu^2_\theta$ for some laws ν^1_θ and ν^2_θ on E. Let $\bar{\nu}_\theta$ be the law on $E \times E$ given by $\bar{\nu}_\theta = \nu^1_\theta \times \nu^2_\theta$. Using dominated convergence, invoking the preservation of translation invariance, and using the limiting distribution of $b_t^{(1)}(\cdot,\cdot)$ on

$\{A, D\}$, we find

$$
\int_{E \times E} \mathrm{d}\bar{\nu}_\theta \, |\Delta_i| + K|\delta_i|
$$

$$
= (1 + K) \int_{E_0 \times E_0} \mathrm{d}\bar{\nu}_\theta \lim_{n \to \infty} \sum_{j \in \mathbb{G}} \left[b_{t_n}^{(1)}\big((i, R_i), (j, A)\big) |x_i^1 - x_i^2| + b_{t_n}^{(1)}\big((i, R_i), (j, D)\big) |y_i^1 - y_i^2| \right]
$$

$$
= \lim_{n \to \infty} (1 + K) \int_{E_0 \times E_0} \mathrm{d}\bar{\nu}_\theta \left| \sum_{j \in \mathbb{G} \times \{A, D\}} b_{t_n}^{(1)}\big((i, R_i), (j, R_j)\big) \big(z_{(j, R_j)}^1 - z_{(j, R_j)}^2\big) \right|
$$

$$
\leq \lim_{n \to \infty} (1 + K) \int_E \mathrm{d}\nu_\theta^1 \left| \sum_{j \in \mathbb{G} \times \{A, D\}} b_{t_n}^{(1)}\big((i, R_i), (j, R_j)\big) z_{(j, R_j)}^1 - \theta \right|
$$

$$
+ \lim_{n \to \infty} (1 + K) \int_E \mathrm{d}\nu_\theta^2 \left| \sum_{i \in \mathbb{G} \times \{A, D\}} b_{t_n}^{(1)}\big((i, R_i), (j, R_j)\big) z_{(j, R_j)}^2 - \theta \right| = 0.
$$

(5.108)

Here, the last equality follows because both ν_θ^1 and ν_θ^2 are in $\mathcal{R}_\theta^{(1)}$ by Lemma 5.8. Thus, we see that $\bar{\nu}_\theta$ concentrates on the diagonal. Suppose now that there exists a sequence $(t_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} \mathbb{E}[|\Delta_i(t_n)| + K|\delta_i(t_n)|] = \delta > 0$. Since $\{\mathcal{L}(\bar{Z}(t_n))\}_{n \in \mathbb{N}}$ is tight (recall (5.93)), by Prokhorov's theorem there exists a converging subsequence $\{\mathcal{L}(\bar{Z}(t_{n_k}))\}_{k \in \mathbb{N}}$. Let $\bar{\nu}_\theta$ denote the limiting measure. Then, by Lemma 5.8 and (5.108),

$$
\delta = \lim_{k \to \infty} \mathbb{E}[|\Delta_i(t_{n_k})| + K|\delta_i(t_{n_k})|] = \int_{E \times E} \mathrm{d}\bar{\nu}_\theta \, [|\Delta_i| + K|\delta_i|] = 0. \tag{5.109}
$$

Thus, $\lim_{t \to \infty} \mathbb{E}[|\Delta_i(t)| + K|\delta_i(t)|] = 0$, and we conclude that the coupling is successful. Hence, given the initial average density θ in (3.4), the equilibrium measure is unique if it exists. $\qquad \square$

3. Stationarity of ν_θ and convergence to ν_θ.

Lemma 5.12 (Existence of equilibrium)**.** *Let $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$. Then $\lim_{t \to \infty} \mu(t) = \nu_\theta$ for some invariant measure ν_θ.*

Proof. To prove that the limit is an invariant measure, suppose that $\mu(0) = \delta_{\underline{\theta}}$, where $\underline{\theta}$ is the state with all components equal to θ. Since the state space of $(Z(t))_{t \geq 0}$ is compact, each sequence $\{\mathcal{L}(Z(t_n))\}_{n \in \mathbb{N}}$ is tight. Hence, by Prokhorov's theorem, there exists a sequence $(t_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} \delta_{\underline{\theta}} S_{t_n} = \nu_\theta$. Since $\delta_{\underline{\theta}} \in \mathcal{R}_\theta^{(1)}$, Lemma 5.8 tells us that $\lim_{n \to \infty} \delta_{\underline{\theta}} S_{t_n} \in \mathcal{R}_\theta^{(1)}$.

To prove that ν_θ is invariant, fix any $s_0 \geq 0$ and put $\mu(0) = \delta_{\underline{\theta}} S_{s_0}$. Then, by Lemma 5.8, $\mu(0) \in \mathcal{R}_\theta^{(1)}$ and, by Lemma 5.11, we can find a further subsequence $(t_{n_k})_{k \in \mathbb{N}}$ such that $\lim_{k \to \infty} \mu(0) S_{t_{n_k}} = \nu_\theta$. By the Feller property of the SSDE in (2.4)–(2.5), we obtain

$$
\nu_\theta S_{s_0} = \lim_{n \to \infty} \delta_{\underline{\theta}} S_{t_n} S_{s_0} = \lim_{k \to \infty} \delta_{\underline{\theta}} S_{s_0} S_{t_{n_k}} = \lim_{k \to \infty} \mu(0) S_{t_{n_k}} = \nu_\theta. \tag{5.110}
$$

Hence, ν_θ is indeed an invariant measure.

Suppose that $\mu(0) \neq \delta_{\underline{\theta}}$, but $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$. Then, by Lemma 5.7, we can couple the two processes starting from $\mu(0)$ and $\delta_{\underline{\theta}}$.

To prove the convergence of $\mu(t)$ to ν_θ, note that $\nu_\theta \in \mathcal{R}_\theta^{(1)}$ by Lemma 5.8. By the invariance of ν_θ, we have $\lim_{t \to \infty} \nu_\theta S_t = \nu_\theta$. By Lemma 5.10, we have $\lim_{t \to \infty} \mu(0) S_t = \lim_{t \to \infty} \nu_\theta S_t = \nu_\theta$ for all $\mu(0) \in \mathcal{R}_\theta^{(1)}$. $\qquad \square$

4. Ergodicity, mixing and associatedness.

Lemma 5.13 (Properties of equilibrium). *Let $\mu(0) \in \mathcal{R}_\theta^{(1)}$ be ergodic under translations. Then $\nu_\theta = \lim_{t \to \infty} \mu(t)$ is ergodic and mixing under translations, and is associated.*

Proof. After a standard approximation argument, [HP91, Corollary1.5 and subsequent discussion] tells us that associatedness is preserved over time. Note that δ_θ is an associated measure and lies in $\mathcal{R}_\theta^{(1)}$. Hence, by Lemma 5.12, $\nu_\theta = \lim_{t \to \infty} \delta_{\underline{\theta}} S_t$ and therefore ν_θ is associated.

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

$$Y_0 = \sum_{j \in B} c_j z_{(j, R_j)}, \qquad Y_k = \sum_{i \in B'} d_i z_{(i+k, R_{i+k})}. \tag{5.111}$$

Note that both Y_0 and Y_k are associated under ν_θ because $(z_{(i, R_i)})_{(i, R_i) \in \mathbb{G} \times \{A, D\}}$ are associated. Therefore, by [NW81, Eq.(2.2)], it follows that for $s, t \in \mathbb{R}$,

$$\left| \mathbb{E}_{\nu_\theta}[e^{\mathrm{i}(sY_0 + tY_n)}] - \mathbb{E}_{\nu_\theta}[e^{\mathrm{i}sY_0}] \, \mathbb{E}_{\nu_\theta}[e^{\mathrm{i}tY_n}] \right| \le |st| \, \mathrm{Cov}_{\nu_\theta}(Y_0, Y_n). \tag{5.112}$$

Since $\mu \in \mathcal{R}_\theta^{(1)}$ by Lemma 5.1,

$$\mathrm{Cov}_{\nu_\theta}(Y_0, Y_k) = \sum_{j \in B} \sum_{i \in B'} c_j d_i \lim_{t \to \infty} \mathrm{Cov}_\mu(z_{(j, R_j)}(t), z_{(i+k, R_{i+k})}(t)) \tag{5.113}$$

$$\le 2\|g\| \sum_{j \in B} \sum_{i \in B'} c_j d_i \int_0^\infty \mathrm{d}r \sum_{(l, R_l) \in \mathbb{G} \times \{A\}} b_r^{(1)}\big((j, R_j), (l, A)\big) \, b_r^{(1)}\big((i+k, R_{i+k}), (l, A)\big).$$

The last integral gives the expected total time for two independent random walks, starting in (j, R_j) and $(i+k, R_{i+k})$, to be active at the same site. To show that this integral converges to 0 as $\|k\| \to \infty$, we rewrite the sum as (recall (5.27)–(5.28))

$$\mathbb{E}_{(i+k, R_{i+k}),(j, R_j)}\left[\left(\sum_{l \in \mathbb{G}} a_{T(r)}(j, l) \, a_{T'(r)}(i+k, l) \right) 1_{\mathcal{E}(r)} \, 1_{\mathcal{E}'(r)} \right]$$

$$= \mathbb{E}_{(i+k, R_{i+k}),(j, R_j)}\left[\left(\sum_{l' \in \mathbb{G}} \hat{a}_{2M(r)}(i+k-j, l') \, a_{\Delta(r)}(l', 0) \right) 1_{\mathcal{E}(r)} \, 1_{\mathcal{E}'(r)} \right]$$

$$\le \mathbb{E}_{(i+k, R_{i+k}),(j, R_j)}\left[\left(\sum_{l' \in \mathbb{G}} \hat{a}_{2M(r)}(i+k-j, l') \big[a_{\Delta(r)}(l', 0) + a_{\Delta(r)}(0, l') \big] \right) 1_{\mathcal{E}(r)} \, 1_{\mathcal{E}'(r)} \right]$$

$$= \mathbb{E}_{(i+k, R_{i+k}),(j, R_j)}\left[\hat{a}_{2M(r)+2\Delta(r)}(i+k-j, 0) 1_{\mathcal{E}(r)} \, 1_{\mathcal{E}'(r)} \right]. \tag{5.114}$$

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

$$T(t) + T'(t) \le 2M(r) + 2\Delta(r) \le 2 \left(T(t) + T'(t) \right), \tag{5.115}$$

the Fourier transform in (5.37)–(5.38) implies that

$$\int_0^\infty \mathrm{d}r \, \mathbb{E}_{(i+k, R_{i+k}),(j, R_j)}\left[\hat{a}_{2M(r)+2\Delta(r)}(0, 0) 1_{\mathcal{E}(r)} \, 1_{\mathcal{E}'(r)} \right] < \infty \tag{5.116}$$

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

5.4 Proof of the dichotomy

Theorem 3.1(a) follows from Lemmas 5.7, 5.12 and 5.13. The equality $\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_0] = \theta$ follows from the evolution equations in (2.4)–(2.5), the fact that ν_θ is an equilibrium measure, and the preservation of θ (see (3.5)). Theorem 3.1(b) follows from Lemma 5.5. The fact that $\theta \mapsto \nu_\theta$ is continuous in the weak topology, as claimed in Remark 3.2, follows from the fact that $\|\nu_\theta(z_i \in \cdot) - \nu_{\theta'}(z_i \in \cdot)\|_{tv} \leq (1 + K)|\theta - \theta'|$ for all $i \in \mathbb{G}$, with $z_i = (z_{(i,R_i)})_{R_i \in \{A,D\}}$. The latter follows from the coupling argument in Lemmas 5.9–5.12. (Pick $\mu(0)$ equal to $\delta_{\underline{\theta}}, \delta_{\underline{\theta'}}$ and use the contraction via the Lyapunov function.)

6 Proofs: long-time behaviour for Model 2

In Sections 6.1–6.4 we show that the results proved in Sections 5.1–5.4 carry over from Model 1 to Model 2. In Section 6.5 we show that symmetry of $a(\cdot, \cdot)$ is needed. In Section 6.6 we show what happens when for infinite seed-bank the fat-tailed wake-up time is modulated by a slowly varying function.

6.1 Moment relations

Like in Model 1, we start by relating the first and second moments of the system in (2.12)–(2.13) to the random walk that evolves according to the transition kernel $b^{(2)}(\cdot, \cdot)$ on $\mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$ given by (2.44). Also here these moment relations hold for all $g \in \mathcal{G}$. Moreover these moment relations holds for $\rho < \infty$ as well as for $\rho = \infty$. Below we write $\mathbb{E}_z$ for $\mathbb{E}_{\delta_z}$, the expectation when the process starts from the initial measure δ_z, $z \in E$.

Lemma 6.1 (First and second moment). *For $z \in E$, $t \geq 0$ and $(i, R_i), (j, R_j) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$,*

$$\mathbb{E}_z[z_{(i,R_i)}(t)] = \sum_{(k,R_k) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}\big((i, R_i), (k, R_k)\big) z_{(k,R_k)} \tag{6.1}$$

and

$$\begin{aligned}
&\mathbb{E}_z[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\
&= \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}\big((i, R_i), (k, R_k)\big) b_t^{(2)}\big((j, R_j), (l, R_l)\big) z_{(k,R_k)} z_{(l,R_l)} \\
&\quad + 2 \int_0^t \mathrm{d}s \sum_{k \in \mathbb{G}} b_{(t-s)}^{(2)}((i, R_i), (k, A)) b_{(t-s)}^{(2)}((j, R_j), (k, A)) \mathbb{E}_z[g(x_k(s))].
\end{aligned} \tag{6.2}$$

Proof. The proof follows from that of Lemma 5.1 after we replace $b^{(1)}(\cdot, \cdot)$ by $b^{(2)}(\cdot, \cdot)$ and use (2.12)–(2.13) instead of (2.4)–(2.5). $\qquad\square$

Remark 6.2 (Density). From Lemma 6.1 we obtain that if μ is invariant under translations with $\mathbb{E}_\mu[x_0(0)] = \theta_x$ and $\mathbb{E}_\mu[y_{0,m}(0)] = \theta_{y,m}$ for all $m \in \mathbb{N}_0$, then

$$\begin{aligned}
\mathbb{E}_\mu[z_{(i,R_i)}(t)] = {} & \theta_x \sum_{(k,R_k) \in \mathbb{G} \times \{A\}} b_t^{(2)}\big((i, R_i), (k, R_k)\big) \\
& + \sum_{m \in \mathbb{N}_0} \theta_{y,m} \sum_{(k,R_k) \in \mathbb{G} \times \{D_m\}} b_t^{(2)}\big((i, R_i), (k, R_k)\big)
\end{aligned} \tag{6.3}$$

and

$$
\begin{aligned}
&\mathbb{E}_\mu[z_{(i,R_i)}(t) z_{(j,R_j)}(t)] \\
&= \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A,(D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big) \, b_t^{(2)}\big((j,R_j),(l,R_l)\big) \, \mathbb{E}_\mu[z_{(k,R_k)} z_{(l,R_l)}] \\
&\quad + 2 \int_0^t \mathrm{d}s \sum_{k \in \mathbb{G}} b_{t-s}^{(2)}\big((i,R_i),(k,A)\big) \, b_{t-s}^{(2)}\big((j,R_j),(k,A)\big) \, \mathbb{E}_\mu[g(x_i(s))].
\end{aligned}
\tag{6.4}
$$

- For $\rho < \infty$, $b^{(2)}(\cdot,\cdot)$ projected onto $\{A,(D_m)_{m \in \mathbb{N}_0}\}$ is a recurrent Markov chain. Therefore, by translation invariance in $\mathbb{G}$, we have

$$
\lim_{t \to \infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \frac{\theta_x + \sum_{m \in \mathbb{N}_0} K_m \theta_{y,m}}{1 + \sum_{m \in \mathbb{N}_0} K_m} = \theta.
\tag{6.5}
$$

- For $\rho = \infty$, $b^{(2)}(\cdot,\cdot)$ viewed as a kernel on $\{A,(D_m)_{m \in \mathbb{N}_0}\}$ is a null-recurrent Markov chain. Hence, for all (i,R_i),

$$
\begin{aligned}
&\lim_{t \to \infty} \sum_{k \in \mathbb{G}} b_t^{(2)}((i,R_i),(k,A)) = 0, \\
&\lim_{t \to \infty} \sum_{k \in \mathbb{G}} b_t^{(2)}((i,R_i),(k,D_m)) = 0 \quad \forall\, m \in \mathbb{N}_0, \\
&\lim_{t \to \infty} \sum_{m \geq M} \sum_{k \in \mathbb{G}} b_t^{(2)}((i,R_i),(k,D_m)) = 1 \quad \forall\, M \in \mathbb{N}_0,
\end{aligned}
\tag{6.6}
$$

which is the analogue of (4.31). Since for $\rho = \infty$ we assume that $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$, it follows that, for all $M \in \mathbb{N}_0$,

$$
\begin{aligned}
\lim_{t \to \infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] &= \lim_{t \to \infty} \theta_x \sum_{k \in \mathbb{G}} b_t^{(2)}\big((i,R_i),(k,A)\big) + \sum_{m \in \mathbb{N}_0} \theta_{y,m} \sum_{k \in \mathbb{G}} b_t^{(2)}\big((i,R_i),(k,D_m)\big) \\
&= \lim_{t \to \infty} \sum_{m \geq M} \theta_{y,m} \sum_{k \in \mathbb{G}} b_t^{(2)}\big((i,R_i),(k,D_m)\big).
\end{aligned}
\tag{6.7}
$$

Letting $M \to \infty$, using (6.6) and noting that $\lim_{m \to \infty} \theta_{y,m} = \theta$ by colour regularity, we obtain

$$
\lim_{t \to \infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta.
\tag{6.8}
$$

$\square$

6.2 The clustering case

In this section we prove convergence to a trivial equilibrium when $\rho < \infty$ and $I_{\hat{a}} = \infty$ and when $\rho = \infty$ and $I_{\hat{a},\gamma} = \infty$. The proof follows along the same lines as in Section 5.2. Therefore we again first consider $g = dg_{\mathrm{FW}}$, and subsequently use a duality comparison argument to show that the results hold for $g \neq dg_{\mathrm{FW}}$ as well.

Case $g = dg_{\mathrm{FW}}$. We start by proving the equivalent of Lemma 5.4, which is Lemma 6.3 below.

Lemma 6.3 (Clustering). *Suppose that $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$ and $g = dg_{\mathrm{FW}}$. Let $\mu(t)$ be the law at time t of the system defined in (2.12)–(2.13). Then the following two statements hold:*

- If $\rho < \infty$ and $I_{\hat{a}} = \infty$, *i.e., $\hat{a}(\cdot, \cdot)$ is recurrent, then*

$$\lim_{t \to \infty} \mu(t) = \theta \, [\delta_{(1,1^{\mathbb{N}_0})}]^{\otimes \mathbb{G}} + (1 - \theta) \, [\delta_{(0,0^{\mathbb{N}_0})}]^{\otimes \mathbb{G}}. \tag{6.9}$$

- If $\rho = \infty$ and $I_{\hat{a}, \gamma} = \infty$ then

$$\lim_{t \to \infty} \mu(t) = \theta \, [\delta_{(1,1^{\mathbb{N}_0})}]^{\otimes \mathbb{G}} + (1 - \theta) \, [\delta_{(0,0^{\mathbb{N}_0})}]^{\otimes \mathbb{G}}. \tag{6.10}$$

Proof. We distinguish between $\rho < \infty$ and $\rho = \infty$, which exhibit different behaviour.

Case $\rho < \infty$. The same dichotomy as for Model 1 holds when the average wake-up time is finite (recall (2.23)–(2.24), (2.53)). Indeed, the argument in (5.32)–(5.41) can be copied with Ke, e replaced by $\chi, \chi/\rho$ and A, B by $\chi/(1+\rho), 1/(1+\rho)$. Under the *symmetry assumption* in (3.14) we have $\tilde{a}(\phi) = 0$. Hence only the law of large numbers in (5.33) is needed, not the central limit theorem in (5.35), which may fail (see Section 6.5).

Case $\rho = \infty$. When the average wake-up time is infinite, we need the assumptions in (3.2) and (3.17). By the standard law of large numbers for stable random variables (see e.g. [Fel71, Section XIII.6]), we have

$$\lim_{k \to \infty} \frac{1}{k} \sum_{\ell=1}^{k} \sigma_\ell = \frac{1}{\chi} \quad \mathbb{P}\text{-a.s.}, \qquad \lim_{k \to \infty} \frac{1}{k^{1/\gamma}} \sum_{\ell=1}^{k} \tau_\ell = W \quad \text{in } \mathbb{P}\text{-probability}, \tag{6.11}$$

with W a stable law random variable on $(0, \infty)$ with exponent γ. Therefore

$$\lim_{t \to \infty} \frac{1}{t^\gamma} N(t) = \lim_{t \to \infty} \frac{1}{t^\gamma} N'(t) = W^{-\gamma} \quad \text{in } \mathbb{P}\text{-probability},$$

$$\lim_{t \to \infty} \frac{1}{t^\gamma} T(t) = \lim_{t \to \infty} \frac{1}{t^\gamma} T'(t) = \frac{1}{\chi} W^{-\gamma} \quad \text{in } \mathbb{P}\text{-probability}, \tag{6.12}$$

$$\lim_{t \to \infty} t^{1-\gamma} \, \mathbb{P}\big(\mathcal{E}(t)\big) = \lim_{t \to \infty} t^{1-\gamma} \, \mathbb{P}\big(\mathcal{E}'(t)\big) = \frac{1}{\chi} \, \mathbb{E}[W^{-\gamma}], \qquad t \to \infty.$$

For the last statement to make sense, we must check the following.

Lemma 6.4 (Finite limits). $\mathbb{E}[W^{-\gamma}] < \infty$.

Proof. Let $W_k = k^{-1/\gamma} \sum_{l=1}^{k} \tau_l$. Then $W_k^{-\gamma} \leq k(\max_{1 \leq i \leq k} \tau_i^\gamma)^{-1}$ and, since τ_i are i.i.d. random variables,

$$\mathbb{E}[W_k^{-\gamma}] \leq \int_0^\infty \mathrm{d}x \, \mathbb{P}\left(k \left(\max_{1 \leq i \leq k} \tau_i^\gamma \right)^{-1} > x \right) = \int_0^\infty \mathrm{d}x \, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x} \right)^k. \tag{6.13}$$

To estimate the integral in the right-hand side of (6.13), we introduce three constants, T, C_1, C_2. Let $\epsilon \in (0, 1)$ and choose $T \in \mathbb{R}_+$ such that, for all $t > T$, $|[\mathbb{P}(\tau > t)/(Ct^{-\gamma})] - 1| < \epsilon$. Since $\mathbb{P}(\tau \leq t) = 1 - \chi^{-1} \sum_{m \in \mathbb{N}_0} K_m e_m \, \mathrm{e}^{-e_m t}$, we note that, under assumption (3.17), τ admits a continuous bounded density. Hence there exists a $C_1 \in \mathbb{R}_+$ such that $\mathbb{P}(\tau \leq t) < C_1 t$. Finally, choose $C_2 \in \mathbb{R}_+$ such that $C_2 > \max(1, C_1^\gamma)$. Split

$$\int_0^\infty \mathrm{d}x \, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x} \right)^k = \int_0^{k/T} \mathrm{d}x \, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x} \right)^k + \int_{k/T}^{kC_2} \mathrm{d}x \, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x} \right)^k + \int_{kC_2}^\infty \mathrm{d}x \, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x} \right)^k. \tag{6.14}$$

We estimate each of the three integrals separately. For the first integral, we use the estimate $(1 - \mathbb{P}(\tau_1^\gamma \geq \frac{k}{x}))^k \leq \exp[-k\mathbb{P}(\tau_1^\gamma \geq \frac{k}{x})]$ to obtain

$$\int_0^{k/T} \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k = \int_0^{k/T} \mathrm{d}x\, \exp\left[-k\,\mathbb{P}\left(\tau_1 \geq \left(\tfrac{k}{x}\right)^{1/\gamma}\right)\right] \leq \int_0^{k/T} \mathrm{d}x\, \mathrm{e}^{-(1-\epsilon)Cx} \leq \frac{1}{(1-\epsilon)C}. \tag{6.15}$$

For the second integral, we note that $t \mapsto t\,\mathbb{P}(\tau_1^\gamma > t)$ is a continuous function on $[\frac{1}{C_2}, T]$, and hence attains a minimum value $C_3 \in \mathbb{R}_+$ on $[\frac{1}{C_2}, T]$. Therefore

$$\int_{k/T}^{kC_2} \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k = \int_{k/T}^{kC_2} \mathrm{d}x\, \left[1 - \mathbb{P}\left(\tau_1^\gamma \geq \tfrac{k}{x}\right)\right]^k$$
$$\leq \int_{k/T}^{kC_2} \mathrm{d}x\, \exp\left[-x\left(\tfrac{k}{x}\mathbb{P}(\tau_1^\gamma \geq \tfrac{k}{x})\right)\right] \leq \frac{1}{C_3}. \tag{6.16}$$

For the third integral, we compute

$$\int_{kC_2}^\infty \mathrm{d}x\, \mathbb{P}\left(\tau_1^\gamma < \tfrac{k}{x}\right)^k \leq \int_{kC_2}^\infty \mathrm{d}x\, \left(C_1^\gamma \tfrac{k}{x}\right)^{\frac{k}{\gamma}} = \int_0^{1/C_2} \mathrm{d}v\, \tfrac{k}{v^2}(C_1^\gamma v)^{\frac{k}{\gamma}} = \frac{C_1^\gamma k}{\frac{k}{\gamma}-1}\left(\frac{C_1^\gamma}{C_2}\right)^{\frac{k}{\gamma}-1}, \tag{6.17}$$

where in the first equality we substitute $v = \frac{k}{x}$. Since $C_2 > C_1^\gamma$, we see that the right-hand side tends to zero as $k \to \infty$. Hence

$$\mathbb{E}[W_k^{-\gamma}] \leq \frac{1}{(1-\epsilon)(C/\gamma)} + \frac{1}{C_3} + \frac{C_1^\gamma k}{\frac{k}{\gamma}-1}\left(\frac{C_1^\gamma}{C_2}\right)^{k-1}, \tag{6.18}$$

and by dominated convergence it follows that $\mathbb{E}[W^{-\gamma}] = \lim_{k\to\infty} \mathbb{E}[W_k^{-\gamma}] < \infty$. $\qquad\square$

By (3.14), we have $\hat{a}(\phi) = a(\phi)$ and $\tilde{a}(\phi) = 0$ in (5.29), and so (5.37) becomes, with the help of (6.12),

$$\mathbb{E}_{(0,A),(0,A)}\left[\left(\sum_{i\in\mathbb{G}} a_{T(t)}(0,i)\,a_{T'(t)}(0,i)\right)\mathbb{1}_{\mathcal{E}(t)}\mathbb{1}_{\mathcal{E}'(t)}\right] \asymp t^{-2(1-\gamma)}f(t), \qquad t \to \infty, \tag{6.19}$$

with (recall (5.31))

$$f(t) = \hat{a}_{ct^\gamma}(0,0) \tag{6.20}$$

for some $c \in (0,\infty)$. Here we use that deviations of $T(t)/t^\gamma$ and $T'(t)/t^\gamma$ away from order 1 are stretched exponentially costly in t [EJU19], and therefore are negligible. Since $t \mapsto \hat{a}_t(0,0)$ is regularly varying at infinity (recall (3.2)), it follows that

$$\hat{a}_{ct^\gamma}(0,0) \asymp \hat{a}_{t^\gamma}(0,0), \qquad t \to \infty. \tag{6.21}$$

Combining (5.26) and (6.19)–(6.21), we get

$$I = \infty \iff I_{\hat{a},\gamma} = \infty \tag{6.22}$$

with $I_{\hat{a},\gamma} = \int_1^\infty \mathrm{d}t\, t^{-2(1-\gamma)}\,\hat{a}_{t^\gamma}(0,0)$. Putting $s = t^\gamma$, we have

$$I_{\hat{a},\gamma} = \int_1^\infty \mathrm{d}s\, s^{-(1-\gamma)/\gamma}\hat{a}_s(0,0), \tag{6.23}$$

which is precisely the integral defined in (3.21). $\qquad\square$

Case $g \neq dg_{\mathrm{FW}}$. To prove that the dichotomy criterion of Lemma 6.3 holds for general $g \in \mathcal{G}$ we need the equivalent of Lemma 5.5. Replacing (2.4)–(2.5) by (2.12)–(2.13), replacing $b^{(1)}$ by $b^{(2)}$ in the proof of Lemma 5.5, and using the moment relations in Lemma 6.1 instead of the moment relations in Lemma 5.1, we see that Lemma 6.3 also holds for $g \in \mathcal{G}$.

6.3 The coexistence case

In this section we prove the coexistence results stated in Theorem 3.3. Like for Model 1 the proofs hold for general $g \in \mathcal{G}$ and we need not distinguish between $g = dg_{\mathrm{FW}}$ and $g \neq dg_{\mathrm{FW}}$. For $\rho < \infty$, the argument is given in Section 6.3.1 and proceeds as in Section 5.3. It is organised along the same 4 Steps as the argument for Model 1, plus an extra Step 5 that settles the statement in (3.28). For $\rho = \infty$, the argument is given in Section 6.3.2 and is also organised along 5 Steps, but structured differently. In Step 1 we define a set of measures that is preserved under the evolution. In Step 2 we use a coupling argument to show the existence of invariant measures. In Step 3 we show that these invariant measures have vanishing covariances in the seed-bank direction. In Step 4 we use the vanishing covariances to show uniqueness of the invariant measure by coupling. Finally, in Step 5 we show that the unique equilibrium measure is invariant, ergodic and mixing under translations, and is associated.

6.3.1 Proof of coexistence for finite seed-bank

1. Properties of measures preserved under the evolution. For Model 2 with $\rho < \infty$, the class of preserved measures is equivalent to $\mathcal{R}_\theta^{(1)}$ for Model 1 and is now defined as follows.

Definition 6.5 (Preserved class of measure). Let $\mathcal{R}_\theta^{(2)}$ denote the set of measures $\mu \in \mathcal{T}$ satisfying, for all $(i, R_i), (j, R_j) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$,

(1)
$$\lim_{t \to \infty} \mathbb{E}_\mu[z_{(i,R_i)}(t)] = \theta, \tag{6.24}$$

(2)
$$\lim_{t \to \infty} \sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}\big((i,R_i),(k,R_k)\big)\, b_t^{(2)}\big((j,R_j),(l,R_l)\big)$$
$$\times \mathbb{E}_\mu[z_{(k,R_k)} z_{(l,R_l)}] = \theta^2. \tag{6.25}$$

$\square$

We begin by stating the following analogues.

Lemma 6.6. *Lemmas 5.7 and 5.8 carry over after* $b^{(1)}(\cdot, \cdot)$ *is replaced by* $b^{(2)}(\cdot, \cdot)$ *and* $\mathcal{R}_\theta^{(1)}$ *by* $\mathcal{R}_\theta^{(2)}$.

Proof. Like for Model 1, properties (1) and (2) of Definition 6.5 hold if and only if

$$\lim_{t \to \infty} \mathbb{E}_\mu\left[\left(\sum_{(k,R_k) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}((i,R_i),(k,R_k))\, z_{(k,R_k)} - \theta\right)^2\right] = 0 \tag{6.26}$$

for some $(i, R_i) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$.

(By irreducibility, (6.26) holds either for all (i, R_i) or for no (i, R_i).) Also for Model 2 with $\rho < \infty$ we have $\mathcal{T}_\theta^{\mathrm{erg}} \subset \mathcal{R}_\theta^{(2)}$. To see why, note for all $t > 0$ and $m \in \mathbb{N}_0$, $(x_i(t))_{i \in \mathbb{G}}$ and

$(y_{i,m}(t))_{i\in\mathbb{G}}$ still are stationary in $\mathbb{G}$. Hence with the help of the Herglotz theorem we can define spectral measures λ_A, λ_{D_m} for $m \in \mathbb{N}_0$ as in (5.72). Let $(RW_t)_{t\geq 0}$ be the random walk evolving according to $b^{(2)}(\cdot,\cdot)$. Introduce the sets

$$
\begin{aligned}
\mathcal{E}(t) &= \big\{\text{at time } t \text{ the random walk is active}\big\}, \\
\mathcal{E}_m(t) &= \big\{\text{at time } t \text{ the random walk is dormant with colour } m\big\}.
\end{aligned}
\tag{6.27}
$$

Note that (pick $(i, R_i) = (0, A)$)

$$
\lim_{t\to\infty} \mathbb{E}_\mu \left[\left(\sum_{(k,R_k)\in\mathbb{G}\times\{A,(D_m)_{m\in\mathbb{N}_0}\}} b_t^{(2)}((0,A),(k,R_k))\, z_{(k,R_k)} - \theta \right)^2 \right]
$$

$$
\leq \lim_{t\to\infty} \mathbb{P}_{(i,R_i)}(\mathcal{E}(t))\, \mathbb{E}_\mu \left[\left(\sum_{k\in\mathbb{G}} \frac{b_t^{(2)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))}\, x_k - \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))} \frac{\theta_x}{1+\rho} \right)^2 \right]
$$

$$
+ \sum_{m\in\mathbb{N}_0} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t))\, \mathbb{E}_\mu \left[\left(\sum_{k\in\mathbb{G}} \frac{b_t^{(2)}((0,A),(k,(D_m)))}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))}\, y_{k,m} - \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))} \frac{K_m\theta_{y,m}}{1+\rho} \right)^2 \right].
\tag{6.28}
$$

Hence we can use the same argument as in the proof of Lemma 5.7 to get the claim. Also the proof of Lemma 5.8 carries over. $\qquad\square$

2. Uniqueness of the equilibrium. To prove uniqueness of the equilibrium for given θ, we use a similar coupling as for Model 1 in Section 5.3 in Step 3. Consider two copies of the system in (2.12)–(2.13) *coupled via their Brownian motions*:

$$
\mathrm{d}x_i^k(t) = \sum_{j\in\mathbb{G}} a(i,j)\left[x_j^k(t)-x_i^k(t)\right]\mathrm{d}t + \sqrt{g(x_i^k(t))}\,\mathrm{d}w_i(t) + \sum_{m\in\mathbb{N}_0} K_m e_m \left[y_{i,m}^k(t)-x_i^k(t)\right]\mathrm{d}t,
\tag{6.29}
$$

$$
\mathrm{d}y_{i,m}^k(t) = e_m\left[x_i^k(t)-y_{i,m}^k(t)\right]\mathrm{d}t, \qquad m\in\mathbb{N}_0, \qquad k\in\{1,2\}.
\tag{6.30}
$$

Here, k labels the copy, and the two copies are driven by the same Brownian motions $(w_i(t))_{t\geq 0}$, $i\in\mathbb{G}$. As initial measures we choose $\mu^1(0), \mu^2(0) \in \mathcal{T}_\theta^{\mathrm{erg}}$.

Let

$$
\bar{z}_i(t) = \big(z_i^1(t), z_i^2(t)\big), \qquad z_i^k(t) = \big(x_i^k(t), (y_{i,m}^k(t))_{m\in\mathbb{N}_0}\big), \quad k\in\{1,2\}.
\tag{6.31}
$$

By [SS80, Theorem 3.2], the coupled system $(\bar{z}_i(t))_{i\in\mathbb{G}}$ has a unique strong solution whose marginals are the single-component systems. Write $\hat{\mathbb{P}}$ to denote the law of the coupled system, and let $\Delta_i(t) = x_i^1(t) - x_i^2(t)$ and $\delta_{i,m}(t) = y_{i,m}^1(t) - y_{i,m}^2(t)$, $m\in\mathbb{N}_0$. The analogue of Lemma 5.9 reads:

Lemma 6.7 (Coupling dynamics $\rho < \infty$). *For every $t \geq 0$,*

$$
\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\mathbb{E}}\left[|\Delta_i(t)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_i(t)|\right]
$$

$$
= -2\sum_{j\in\mathbb{G}} a(i,j)\,\hat{\mathbb{E}}\left[|\Delta_j(t)|\,1_{\{\mathrm{sgn}\,\Delta_i(t)\neq\mathrm{sgn}\,\Delta_j(t)\}}\right]
$$

$$
- 2\sum_{m\in\mathbb{N}_0} K_m e_m\,\hat{\mathbb{E}}\left[\big(|\Delta_i(t)| + |\delta_{i,m}(t)|\big)\,1_{\{\mathrm{sgn}\,\Delta_i(t)\neq\mathrm{sgn}\,\delta_{i,m}(t)\}}\right].
\tag{6.32}
$$

Proof. Note that the left-hand side of (6.32) is well defined because $\rho < \infty$. The proof of Lemma 6.7 carries over from that of Lemma 5.9 after replacing (2.4)–(2.5) by (2.12)–(2.13). $\qquad\square$

The analogue of Lemma 5.10 reads as follows.

Lemma 6.8 (Succesfull coupling $\rho < \infty$). *If $a(\cdot,\cdot)$ is transient, then the coupling is successful, i.e.,*

$$\lim_{t\to\infty} \hat{\mathbb{E}}\left[|\Delta_i(t)| + \sum_{m\in\mathbb{N}_0} K_m |\delta_{i,m}(t)| \right] = 0 \qquad \forall i \in \mathbb{G}. \tag{6.33}$$

Proof. This follows in the same way as in the proof of Lemma 5.10, by defining $-h_i(t)$ as in the right-hand side of (6.32). Using that the second line of (5.105) now holds for $\delta_{i,m}(t)$ and all $m \in \mathbb{N}_0$, we can finish the proof after replacing $b_t^{(1)}(\cdot,\cdot)$ in (5.108) by $b_t^{(2)}(\cdot,\cdot)$ and summing over the seed-banks D_m, $m \in \mathbb{N}_0$. $\qquad\square$

3. Stationarity of the equilibrium ν_θ and convergence to ν_θ. Lemma 5.12 holds also for $\mu \in \mathcal{R}_\theta^{(2)}$. This follows after replacing $\mu \in \mathcal{R}_\theta^{(1)}$ by $\mu \in \mathcal{R}_\theta^{(2)}$ in the proof of Lemma 5.12, using the equivalent of Lemma 5.8 and invoking Lemma 6.8 instead of Lemma 5.10.

4. Ergodicity, mixing and associatedness. Also Lemma 5.13 holds, after replacing $b^{(1)}(\cdot,\cdot)$ by $b^{(2)}(\cdot,\cdot)$. The proof even simplifies, since we can invoke the symmetry of $a(\cdot,\cdot)$ in (5.114).

5. Variances under the equilibrium measure ν_θ. If $\limsup_{m\to\infty} e_m = 0$, then the claim in (3.28) is a direct consequence of the proof of Lemma 6.12 for $\rho = \infty$. If $\liminf_{m\to\infty} e_m > 0$, then the claim follows from the fact that $\mu \in \mathcal{R}_\theta^{(2)}$ and

$$\mathrm{Var}_{\nu_\theta}(y_{0,m}) = \lim_{t\to\infty} 2 \int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} b_{t-s}^{(2)}\big((0,D_m),(k,A)\big)\, b_{t-s}^{(2)}\big((0,D_m),(k,A)\big)\, \mathbb{E}_\mu[g(x_i(s))]. \tag{6.34}$$

Since $e_m > 0$ for all $m \in \mathbb{N}_0$ and $\liminf_{m\to\infty} e_m > 0$, there is a positive probability that after the first steps the two random walks are both active at 0, i.e., are both in state $(0,A)$. Hence, for all $m \in \mathbb{N}_0$ there exists a constant $c > 0$ such that

$$\mathrm{Var}_{\nu_\theta}(y_{0,m}) \geq c\,\mathrm{Var}_{\nu_\theta}(x_0). \tag{6.35}$$

Since ν_θ is a non-trivial equilibrium, we have $\mathrm{Var}_{\nu_\theta}(x_0) > 0$.

6.3.2 Proof of coexistence for infinite seed-bank

1. Properties of measures preserved under the evolution. For $\rho = \infty$, the class of preserved measures is given by

$$\mathcal{R}_\theta^{(2),\bullet} = \big\{ \mu \in \mathcal{R}_\theta^{(2)} : \mu \text{ is colour regular} \big\}. \tag{6.36}$$

(Recall Definition 6.5.) We begin by stating the following analogues.

Lemma 6.9. *Lemmas 5.7 and carry over after $\mathcal{T}_\theta^{\mathrm{erg}}, \mathcal{R}_\theta^{(2)}$ are replaced by $\mathcal{T}_\theta^{\mathrm{erg},\bullet}, \mathcal{R}_\theta^{(2),\bullet}$.*

Proof. Let the sets $\mathcal{E}_m(t)$, $t > 0$, $m \in \mathbb{N}_0$, be defined as in (6.27), and define λ_A and λ_{D_m} analogously to (5.72), like for $\rho < \infty$. The equivalent of (5.74) is

$$\mathbb{E}_\mu \left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))} \sum_{k \in \mathbb{G}} b_t^{(2)}((0,A),(k,A)) x_k - \theta_x \right)^2 \right]$$
$$= \frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))^2} \int_{[-\pi,\pi]^d} \mathbb{E}_{(0,A),(0,A)} \left[\mathrm{e}^{-T(t)(1-a(\phi))} 1_{\mathcal{E}(t)} \mathrm{e}^{-T'(t)(1-\bar{a}(\phi))} 1_{\mathcal{E}'(t)} \right] \mathrm{d}\lambda_A. \tag{6.37}$$

Using that $T(t), T'(t) \to \infty$ as $t \to \infty$ (see (6.12)), that $T(t)$, $T'(t)$, $\mathcal{E}(t)$, $\mathcal{E}'(t)$ are asymptotically independent and that $a(\cdot, \cdot)$ is irreducible, we still find

$$\lim_{t \to \infty} \mathbb{E}_\mu \left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))} \sum_{k \in \mathbb{G}} b_t^{(2)}((0,A),(k,A)) x_k - \theta_x \right)^2 \right] = \lambda_A(\{0\}) \tag{6.38}$$

and, similarly,

$$\lim_{t \to \infty} \mathbb{E}_\mu \left[\left(\frac{1}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))} \sum_{k \in \mathbb{G}} b_t^{(2)}((0,A),(k,A)) y_{k,m} - \theta_{y,m} \right)^2 \right] = \lambda_{D_m}(\{0\}). \tag{6.39}$$

Since $\mu \in \mathcal{T}_\theta^{\mathrm{erg}}$, we have $\lambda_A(\{0\}) = 0$ and $\lambda_{D_m}(\{0\}) = 0$ for all $m \in \mathbb{N}_0$ (recall (5.79)). By the colour regularity,

$$\lim_{t \to \infty} \theta_x \mathbb{P}_{(0,A)}(\mathcal{E}(t)) + \sum_{m \in \mathbb{N}_0} \theta_{y,m} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t)) = \theta. \tag{6.40}$$

Therefore we can rewrite (6.28) as

$$\lim_{t \to \infty} \mathbb{E}_\mu \left[\left(\sum_{(k,R_k),(l,R_l) \in \mathbb{G} \times \{A,(D_m)_{m \in \mathbb{N}_0}\}} b_t^{(2)}((0,A),(k,R_k)) z_{(k,R_k)} - \theta \right)^2 \right]$$
$$\leq \lim_{t \to \infty} \mathbb{P}_{(0,A)}(\mathcal{E}(t)) \mathbb{E}_\mu \left[\left(\sum_{k \in \mathbb{G}} \frac{b_t^{(2)}((0,A),(k,A))}{\mathbb{P}_{(0,A)}(\mathcal{E}(t))} (x_k - \theta_x) \right)^2 \right]$$
$$+ \sum_{m \in \mathbb{N}_0} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t)) \mathbb{E}_\mu \left[\left(\sum_{k \in \mathbb{G}} \frac{b_t^{(2)}((0,A),(k,(D_m)))}{\mathbb{P}_{(0,A)}(\mathcal{E}_m(t))} (y_{k,m} - \theta_{y,m}) \right)^2 \right]$$
$$= \lim_{t \to \infty} \mathbb{P}_{(0,A)}(\mathcal{E}(t)) \lambda_A(\{0\}) + \sum_{m \in \mathbb{N}_0} \mathbb{P}_{(0,A)}(\mathcal{E}_m(t)) \lambda_{D_m}(\{0\}) = 0. \tag{6.41}$$

We conclude that indeed $\mu \in \mathcal{R}_\theta^{(2),\bullet}$. Also the proof of Lemma 5.8 carries over. $\square$

Like for $\rho < \infty$, Lemma 5.8 carries over after we replace $b^{(1)}(\cdot, \cdot)$ by $b^{(2)}(\cdot, \cdot)$ and $\mathcal{R}_\theta^{(1)}$ by $\mathcal{R}_\theta^{(2),\bullet}$.

2. Existence of invariant measures ν_θ for $\rho = \infty$. Since the dynamics for $\rho = \infty$ and $\rho < \infty$ are the same, we can still use the coupling in (6.29)–(6.30). Also Lemma 6.7 holds for $\rho = \infty$, but if $\rho = \infty$, then the left-hand side of (6.32) can become infinite. Therefore we cannot use the line of argument used for Model 1 to show that the coupling is successful for arbitrary initial measures μ^1, $\mu^2 \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$. However, we can prove the following lemma.

Lemma 6.10 (Successful coupling). *If μ_1, $\mu_2 \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ satisfy*

$$\hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m \in \mathbb{N}_0} K_m |\delta_i(0)|\right] < \infty, \tag{6.42}$$

then the coupling in (6.29)–(6.30) *is successful.*

Proof. We proceed similarly as in Step 3 for $\rho < \infty$. Note, in particular, that $\int_0^\infty \mathrm{d}t\, h_i(t)$ (recall (6.32)) is bounded from above by $\hat{\mathbb{E}}[|\Delta_i(0)| + \sum_{m \in \mathbb{N}_0} K_m |\delta_i(0)|]$ (compare with (5.104)). Also for $\rho = \infty$ we obtain Lemma 5.11. Like for Model 1, if we define

$$\begin{aligned}
E_0 \times E_0 = &\left\{ \bar{z} \in E \times E \colon z^1_{(i,R_i)} \geq z^2_{(i,R_i)} \; \forall (i,R_i) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\} \right\} \\
&\cup \left\{ \bar{z} \in E \times E \colon z^2_{(i,R_i)} \geq z^1_{(i,R_i)} \; \forall (i,R_i) \in \mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\} \right\},
\end{aligned} \tag{6.43}$$

then we find $\lim_{t \to \infty} \mathbb{P}(E_0 \times E_0) = 1$ and hence the coupled diffusions $(Z^1(t))_{t \geq 0}$ and $(Z^2(t))_{t \geq 0}$ lay on top of each other as $t \to \infty$. However, in (5.108) the limiting distribution of $b^{(1)}_{t_n}(\cdot, \cdot)$ was used "to compensate" the factors K_m in $|\Delta_i| + \sum_{m \in \mathbb{N}_0} K_m |\delta_{i,m}|$. Since, for $\rho = \infty$, $b^{(1)}_{t_n}(\cdot, \cdot)$ does not have a well-defined limiting distribution for the projection on $\{A, (D_m)_{m \in \mathbb{N}_0}\}$, we need a different strategy.

To obtain a successful coupling, as before, let $(t_n)_{n \in \mathbb{N}}$ be a subsequence such that $\nu^1_\theta = \lim_{n \to \infty} \mathcal{L}(Z^1(t_n))$ with $\mathcal{L}(Z^1(0)) = \mu^1$ and $\nu^2_\theta = \lim_{n \to \infty} \mathcal{L}(Z^2(t_n))$ with $\mathcal{L}(Z^2(0)) = \mu^2$. For $\mathbb{G} = \mathbb{Z}^d$, let $\Lambda_N = [0,N)^d \cap \mathbb{Z}^d$, $N \in \mathbb{N}$. (As noted before, for amenable groups $\mathbb{G}$, $(\Lambda_N)_{N \in \mathbb{N}}$ must be replaced by a so-called Følner sequence.) Note that

$$\mathbb{E}_{\nu^1_\theta}\left[\left(\frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} x_j - \theta\right)^2\right] = \frac{1}{|\Lambda_N|^2} \sum_{i,j \in \Lambda_N} \mathrm{Cov}_{\nu^1_\theta}(x_i, x_j). \tag{6.44}$$

Since $\mu^1 \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$, we have $\mu^1 \in \mathcal{R}_\theta^{(2)}$. Hence, by Lemma 6.1 and Remark 6.2,

$$\begin{aligned}
\mathrm{Cov}_{\nu^1_\theta}(x_i, x_j) &= \lim_{n \to \infty} \mathrm{Cov}_{\mu^1}(x_i(t_n), x_j(t_n)) \\
&\leq \lim_{n \to \infty} 2\|g\| \int_0^{t_n} \mathrm{d}s \sum_{k \in \mathbb{G}} b^{(2)}_{(t_n - s)}((i,A),(k,A))\, b^{(2)}_{(t_n - s)}((j,A),(k,A)) \\
&\leq 2\|g\| \int_0^\infty \mathrm{d}s \sum_{k \in \mathbb{G}} \mathbb{E}_{(i,A),(j,A)}\left[a_{T(s)}(i,k)\, \mathbb{1}_{\mathcal{E}(s)}\, a_{T'(s)}(j,k)\, \mathbb{1}_{\mathcal{E}'(s)}\right] \\
&\leq 2\|g\| \int_0^\infty \mathrm{d}s\, \mathbb{E}_{(i,A),(j,A)}\left[\hat{a}_{T(s)+T'(s)}(i-j,0)\, \mathbb{1}_{\mathcal{E}(s)}\, \mathbb{1}_{\mathcal{E}'(s)}\right].
\end{aligned} \tag{6.45}$$

Since $I_{\alpha,\gamma} < \infty$, we see that the last integral is finite. Since $\lim_{\|i-j\| \to \infty} \hat{a}_t(i-j,0) = 0$ for all $t > 0$, it follows by transience and dominated convergence that $\lim_{\|i-j\| \to \infty} \mathrm{Cov}_{\nu^1_\theta}(x_i, x_j) = 0$. Since $\mathrm{Cov}_{\nu^1_\theta}(x_i, x_j) \leq 1$ for all $i,j \in \mathbb{G}$, for all $\epsilon > 0$ there exists an $L \in \mathbb{N}$ such that

$$\begin{aligned}
\lim_{N \to \infty} \mathbb{E}_{\nu^1_\theta}\left[\left(\frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} x_j - \theta\right)^2\right] &= \lim_{N \to \infty} \frac{1}{|\Lambda_N|^2} \sum_{i,j \in \Lambda_N} \mathrm{Cov}_{\nu^1_\theta}(x_i, x_j) \\
&= \lim_{N \to \infty} \frac{1}{|\Lambda_N|^2} \sum_{\substack{i,j \in \Lambda_N \\ \|i-j\| \leq L}} \mathrm{Cov}_{\nu^1_\theta}(x_i, x_j) + \frac{1}{|\Lambda_N|^2} \sum_{\substack{i,j \in \Lambda_N \\ \|i-j\| > L}} \mathrm{Cov}_{\nu^1_\theta}(x_i, x_j) \\
&\leq \lim_{N \to \infty} \frac{|\{i,j \in \Lambda_N \colon \|i-j\| \leq L\}|}{|\Lambda_N|^2} + \epsilon \lim_{N \to \infty} \frac{|\{i,j \in \Lambda_N \colon \|i-j\| > L\}|}{|\Lambda_N|^2} < \epsilon.
\end{aligned} \tag{6.46}$$

We conclude that

$$\lim_{N \to \infty} \mathbb{E}_{\nu_\theta^1} \left[\left(\frac{1}{\Lambda_N} \sum_{j \in \Lambda_N} x_j - \theta \right)^2 \right] = 0, \tag{6.47}$$

and the same holds for ν_2^θ. Let $\lim_{n \to \infty} \mathcal{L}(\bar{Z}(t_n)) = \bar{\nu}_\theta$ such that $\lim_{n \to \infty} \mathcal{L}(Z^1(t_n)) = \nu_\theta^1$ and $\lim_{n \to \infty} \mathcal{L}(Z^2(t_n)) = \nu_\theta^2$. Then by translation invariance of $\bar{\nu}_\theta$ and the fact that $\bar{\nu}_\theta(E_0 \times E_0) = 1$, we find

$$\int_{E \times E} d\bar{\nu}_\theta |\Delta_i| = \int_{E_0 \times E_0} d\bar{\nu}_\theta \frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} |x_j^1 - x_j^2|$$

$$\leq \int_{E_0} d\nu_\theta^1 \left| \frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} x_j^1 - \theta \right| + \int_{E_0} d\nu_\theta^2 \left| \frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} x_j^2 - \theta \right|. \tag{6.48}$$

Letting $N \to \infty$, we see by translation invariance of $\bar{\nu}_\theta$ that $\mathbb{E}_{\bar{\nu}_\theta}[|\Delta_i|] = 0$ for all $i \in \mathbb{G}$.

The result in (6.47) holds also for x_i replaced by $y_{i,m}$, $m \in \mathbb{N}_0$, since the integral in (6.45) can only become smaller when we start from a dormant site. Replacing $|\Delta_i|$ in (6.48) by $|\delta_{i,m}|$, we obtain, for all $m \in \mathbb{N}_0$,

$$\mathbb{E}_{\bar{\nu}_\theta}[|\delta_{i,m}|] = 0 \qquad \forall\, m \in \mathbb{N}_0. \tag{6.49}$$

We conclude that the coupling is successful. $\qquad \square$

Let $(S_t)_{t \geq 0}$ denote the semigroup associated with (2.12)–(2.13). To prove the existence of an invariant measure, note that $E \times E$ is a compact space (in the product topology). Hence, if $t_n \to \infty$, then the sequence μS_{t_n} has a convergent subsequence. In Lemma 6.11 below we show that each weak limit point of the sequence μS_{t_n} is invariant under the evolution of (2.12)–(2.13).

Lemma 6.11 (Invariant measure). *Suppose that* $\mu \in \mathcal{R}_\theta^{(2),\bullet}$. *If* $t_n \to \infty$ *and* $\mu S_{t_n} \to \nu_\theta$, *then* ν_θ *is an invariant measure under the evolution in* (2.12)–(2.13).

Proof. Fix $s > 0$. Let $\mu^1 = \mu$ and $\mu^2 = \mu S_s$. We couple the evolutions starting from μ^1 and μ^2 via their Brownian motions (see (6.29)–(6.30)). Note that, by the SSDE in (2.12)–(2.13),

$$\hat{\mathbb{E}} \left[|\Delta_i(0)| + \sum_{m \in \mathbb{N}_0} K_m |\delta_{i,m}(0)| \right] = \mathbb{E} \left[|x_i(0) - x_i(s)| + \sum_{m \in \mathbb{N}_0} K_m |y_{i,m}(0) - y_{i,m}(s)| \right]$$

$$= \mathbb{E} \Bigg[\Bigg| \int_0^s \sum_{j \in \mathbb{G}} a(i,j) \left[x_j(r) - x_i(r) \right] dr$$

$$+ \int_0^s \sqrt{g(x_i(r))}\, dw_i(r)$$

$$+ \int_0^s \sum_{m \in \mathbb{N}_0} K_m e_m \left[y_{i,m}(r) - x_i(r) \right] dr \Bigg|$$

$$+ \sum_{m \in \mathbb{N}_0} K_m \int_0^s |e_m[y_{i,m}(r) - x_i(r)]|\, dr \Bigg]. \tag{6.50}$$

Using that all rates are finite and that, by Knight's theorem (see [RY99, Theorem V.1.9 p.183]), we can write the Brownian integral as a time-transformed Brownian motion, we see that $\hat{\mathbb{E}}[|\Delta_i(0)| + \sum_{m \in \mathbb{N}_0} K_m|\delta_{i,m}(0)|] < \infty$. Hence, by (6.42), we can successfully

couple the evolutions starting from μ^1 and μ^2, and so $\lim_{n\to\infty} \mu^2 S_{t_n} = \lim_{n\to\infty} \mu S_s S_{t_n} = \nu_\theta$. By the Feller property of the SSDE in (2.12)–(2.13), it follows that

$$\nu_\theta S_s = \lim_{n\to\infty} \mu(t_n) S_s = \lim_{n\to\infty} \mu S_{t_n} S_s = \lim_{n\to\infty} \mu S_s S_{t_n} = \nu_\theta. \tag{6.51}$$

We conclude that ν_θ is indeed an invariant measure for the SSDE in (2.12)–(2.13). $\qquad\square$

3. Invariant measures have vanishing covariances in the seed-bank direction for $\rho = \infty$. In this step we prove that an invariant measure ν_θ has vanishing variances in the seed-bank direction. In Step 5 we use this property to successfully couple any two invariant measures.

Lemma 6.12 (Deterministic deep seed-banks)**.** *If* $\nu_\theta = \lim_{n\to\infty} \mu S_{t_n}$ *for some* $\mu \in \mathcal{R}_\theta^{(2),\bullet}$ *and* $t_n \to \infty$, *then*

$$\lim_{m\to\infty} \mathrm{Var}_{\nu_\theta}(y_{i,m}) = 0 \qquad \forall i \in \mathbb{G}. \tag{6.52}$$

Proof. Since ν_θ is translation invariant, it is enough to show that $\lim_{m\to\infty} \mathrm{Var}_{\nu_\theta}(y_{0,m}) = 0$. Since $\mu(0) \in \mathcal{R}_\theta^{(2),\bullet}$, it follows from Lemma 6.1 that

$$\lim_{m\to\infty} \mathrm{Var}_{\nu_\theta}[y_{0,m}]$$
$$= \lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}_\mu \left[\left(y_{0,m}(t_n) - \mathbb{E}_\mu[y_{0,m}(t_n)] \right)^2 \right]$$
$$= \lim_{m\to\infty} \lim_{n\to\infty} 2 \int_0^{t_n} \mathrm{d}s \sum_{k\in\mathbb{G}} b_{(t_n-s)}^{(2)}((0,D_m),(k,A))\, b_{(t_{n_k}-s)}^{(2)}((0,D_m),(k,A))\, \mathbb{E}_z[g(x_k(s))]. \tag{6.53}$$

Since g is positive and bounded, it is therefore enough to prove that

$$\lim_{m\to\infty} \lim_{n\to\infty} \int_0^{t_n} \mathrm{d}u \sum_{k\in\mathbb{G}} b_u^{(2)}((0,D_m),(k,A))\, b_u^{(2)}((0,D_m),(k,A)) = 0. \tag{6.54}$$

Recall (see e.g. (6.45)) that $b_u^{(2)}((0,D_m),(k,A))\, b_u^{(2)}((0,D_m),(k,A))$ is the probability that two independent random walks, denoted by RW and RW' and moving according to $b^{(2)}(\cdot,\cdot)$, are at time u at the same site k and both active. Define

$$\tau = \left\{ t \geq 0 : RW(t) = RW'(t) = (i,A) \text{ for some } i \in \mathbb{G} \right\}. \tag{6.55}$$

Then we can rewrite the left-hand side of (6.54) as

$$\lim_{m\to\infty} \lim_{n\to\infty} \int_0^{t_n} \mathrm{d}u \sum_{k\in\mathbb{G}} b_u^{(2)}((0,D_m),(k,A))\, b_u^{(2)}((0,D_m),(k,A))$$
$$= \lim_{m\to\infty} \lim_{n\to\infty} \int_0^{t_n} \mathrm{d}u\, \mathbb{E}_{(0,D_m),(0,D_m)} \left[\sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}}\, 1_{\{RW'(u)=k\}}\, 1_{\mathcal{E}(u)}\, 1_{\mathcal{E}'(u)} \right]$$
$$= \lim_{m\to\infty} \lim_{n\to\infty}$$
$$\int_0^{t_n} \mathrm{d}u\, \mathbb{E}_{(0,D_m),(0,D_m)} \left[\sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}}\, 1_{\{RW'(u)=k\}}\, 1_{\mathcal{E}(t)}\, 1_{\mathcal{E}'(t)} \left(1_{\{\tau<\infty\}} + 1_{\{\tau=\infty\}} \right) \right]$$
$$= \lim_{m\to\infty} \lim_{n\to\infty}$$
$$\mathbb{E}_{(0,D_m),(0,D_m)} \left[1_{\{\tau<\infty\}} \mathbb{E}_{(0,D_m),(0,D_m)} \left[\int_0^{t_n} \mathrm{d}u \sum_{k\in\mathbb{G}} 1_{\{RW(u)=k\}} 1_{\{RW'(u)=k\}} 1_{\mathcal{E}(u)} 1_{\mathcal{E}'(u)} \mid \mathcal{F}_\tau \right] \right]$$
$$= \lim_{m\to\infty} \lim_{n\to\infty}$$

$$\mathbb{E}_{(0,D_m),(0,D_m)}\left[1_{\{\tau<\infty\}}\mathbb{E}_{(0,A),(0,A)}\left[\int_0^{t_n-\tau}\mathrm{d}u\sum_{k\in\mathbb{G}}1_{\{RW(u)=k\}}1_{\{RW'(u)=k\}}1_{\mathcal{E}(u)}1_{\mathcal{E}'(u)}\right]\right]$$

$$\leq\lim_{m\to\infty}\mathbb{E}_{(0,D_m),(0,D_m)}\left[1_{\{\tau<\infty\}}\mathbb{E}_{(0,A),(0,A)}\left[\int_0^\infty\mathrm{d}u\sum_{k\in\mathbb{G}}1_{\{RW(u)=k\}}1_{\{RW'(u)=k\}}1_{\mathcal{E}(u)}1_{\mathcal{E}'(u)}\right]\right]$$

$$=\lim_{m\to\infty}\mathbb{P}_{(0,D_m),(0,D_m)}\left(\tau<\infty\right)I_{\hat{a},\gamma}, \tag{6.56}$$

where we use that $I_{\hat{a},\gamma}<\infty$, the strong Markov property, and the fact that for $\tau=\infty$ the product of the indicators equals 0 for all $u\in\mathbb{R}_{\geq0}$. Therefore (6.52) holds if

$$\lim_{m\to\infty}\mathbb{P}_{(0,D_m),(0,D_m)}\left(\tau<\infty\right)=0. \tag{6.57}$$

Define

$$\tau^*=\inf\left\{t\geq0:\text{ both }RW\text{ and }RW'\text{ are active at time }t\right\}. \tag{6.58}$$

Note that $\tau^*\leq\tau$. Theorefore we can write (recall that in Model 2 the random walk kernel $a(\cdot,\cdot)$ is assumed to be symmetric),

$$\lim_{m\to\infty}\mathbb{P}_{(0,D_m),(0,D_m)}\left(\tau<\infty\right)$$
$$=\lim_{m\to\infty}\mathbb{E}_{(0,D_m),(0,D_m)}[1_{\{\tau<\infty\}}]$$
$$=\lim_{m\to\infty}\mathbb{E}_{(0,D_m),(0,D_m)}\left[1_{\{\tau^*<\infty\}}\mathbb{E}_{(0,D_m)^2}\left[1_{\{\tau<\infty\}}\mid\mathcal{F}_{\tau^*}\right]\right]$$
$$=\lim_{m\to\infty}\mathbb{E}_{(0,D_m),(0,D_m)}\left[\mathbb{E}_{RW(\tau^*),RW'(\tau^*)}\left[1_{\{\tau<\infty\}}\right]\right]$$
$$=\lim_{m\to\infty}\sum_{k,l\in\mathbb{G}}\mathbb{P}_{(0,D_m),(0,D_m)}\left(RW(\tau^*)=(k,A),RW'(\tau^*)=(l,A)\right)\mathbb{E}_{(k,A),(l,A)}\left[1_{\{\tau<\infty\}}\right]$$
$$=\lim_{m\to\infty}\sum_{k,l\in\mathbb{G}}\mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)}(0,k)\,\hat{a}_{T'(\tau^*)}(0,l)]\,\mathbb{E}_{(0,A),(l-k,A)}\left[1_{\{\tau<\infty\}}\right]$$
$$=\lim_{m\to\infty}\sum_{k,l\in\mathbb{G}}\mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)}(0,k)\,\hat{a}_{T'(\tau^*)}(-k,l-k)]\,\mathbb{E}_{(0,A),(l-k,A)}\left[1_{\{\tau<\infty\}}\right]$$
$$=\lim_{m\to\infty}\sum_{k,j\in\mathbb{G}}\mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)}(0,-k)\,\hat{a}_{T'(\tau^*)}(-k,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]$$
$$=\lim_{m\to\infty}\sum_{j\in\mathbb{G}}\mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]$$
$$=\lim_{m\to\infty}\sum_{\substack{j\in\mathbb{G}\\\|j\|\leq L}}\mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]$$
$$+\lim_{m\to\infty}\sum_{\substack{j\in\mathbb{G}\\\|j\|>L}}\mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right].$$
$$\tag{6.59}$$

To prove that the expression in the right-hand side tends to zero, we fix $\epsilon>0$ and prove that there exists an $L\in\mathbb{N}$ such that both sums are smaller that $\frac{\epsilon}{2}$.

Claim 1: There exists an L such that
$\lim_{m\to\infty}\sum_{j\in\mathbb{G},\|j\|>L}\mathbb{E}_{(0,D_m)^2}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\,\mathbb{E}_{(0,A),(j,A)}[1_{\{\tau<\infty\}}]<\frac{\epsilon}{2}$. Using the sym-

metry of the kernel $a(\cdot, \cdot)$ in Model 2, we find

$$
\begin{aligned}
\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right] &= \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s\, 1_{\{\tau\in\mathrm{d}s\}}\right] \\
&\leq \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s \sum_{k\in\mathbb{G}} 1_{\mathcal{E}(s)} 1_{\mathcal{E}'(s)} 1_{\{RW=k\}} 1_{\{RW'=k\}}\right] \\
&\leq \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s \sum_{k\in\mathbb{G}} \hat{a}_{T(s)}(0,k)\, \hat{a}_{T'(s)}(j,k)\, 1_{\mathcal{E}(s)}\, 1_{\mathcal{E}'(s)}\right] \\
&\leq \mathbb{E}_{(0,A),(j,A)}\left[\int_0^\infty \mathrm{d}s\, \hat{a}_{T(s)+T'(s)}(j,0) 1_{\mathcal{E}(s)} 1_{\mathcal{E}'(s)}\right].
\end{aligned}
\tag{6.60}
$$

The last integral in the right-hand side is dominated by $I_{\hat{a},\gamma}$ (recall (3.42)). Since, for all $t\in\mathbb{R}_{\geq 0}$,

$$
\lim_{\|j\|\to\infty} \hat{a}_t(0,j) = 0,
\tag{6.61}
$$

it follows by dominated convergence that for each $\epsilon > 0$ we can find an L such that, for all $\|j\| > L$,

$$
\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right] < \tfrac{\epsilon}{2}.
\tag{6.62}
$$

Hence, for L sufficiently large, we find

$$
\begin{aligned}
&\lim_{m\to\infty} \sum_{j\in\mathbb{G},\|j\|>L} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\left[\mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right]\right] \\
&\leq \lim_{m\to\infty} \tfrac{\epsilon}{2} \sum_{j\in\mathbb{G},\|j\|>L} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)] \leq \tfrac{\epsilon}{2}.
\end{aligned}
\tag{6.63}
$$

Claim 2: For L given as in Claim 1,
$\lim_{m\to\infty} \sum_{j\in\mathbb{G},\|j\|\leq L} \mathbb{E}_{(0,D_m)^2}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)]\, \mathbb{E}^{(0,A),(j,A)}[1_{\{\tau<\infty\}}] < \tfrac{\epsilon}{2}.$ For the first sum, note that

$$
\begin{aligned}
&\lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\\|j\|\leq L}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{(T(\tau^*)+T'(\tau^*))}(0,j)]\, \mathbb{E}_{(0,A),(j,A)}\left[1_{\{\tau<\infty\}}\right] \\
&\leq \lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\\|j\|\leq L}} \mathbb{E}_{(0,D_m),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)] \\
&= \lim_{m\to\infty} \sum_{\substack{j\in\mathbb{G}\\\|j\|\leq L}} \mathbb{E}_{(0,A),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)],
\end{aligned}
\tag{6.64}
$$

where in the last equality we condition on the first time one of the two random walks wakes up, and use the strong Markov property. We will show that the right-hand side tends to zero as $m\to\infty$. Recall that we assumed (3.17): $e_m \sim Bm^{-\beta}$ for $\beta > 0$. Note that, in order for the random walks to be both active at the same time, the random walk starting in $(0,D_m)$ has to become active at least once. Hence, for all $t\geq 0$, we have

$$
\lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,A)}(\tau^* \leq t) \leq \lim_{m\to\infty} 1 - e^{-e_m t} = 0.
\tag{6.65}
$$

By (6.12) and [EJU19], we also have for the random walk starting in $(0,A)$ that

$$
\lim_{t\to\infty} T(t) \sim ct^\gamma.
\tag{6.66}
$$

Fix $\epsilon > 0$. Since $\lim_{t\to\infty} \hat{a}_t(0,j) = 0$ for all $j \in \mathbb{G}$, we can find a $T^\star$ such that, for all $t > T^\star$,

$$\sum_{\substack{j \in \mathbb{G} \\ \|j\| \leq L}} \hat{a}_t(0,j) < \tfrac{\epsilon}{6}. \tag{6.67}$$

By (6.66), we can find a $\tilde{t} \in \mathbb{R}_{\geq 0}$ such that $\mathbb{P}_{(0,A)}(T(\tilde{t}) > T^\star) \geq 1 - \tfrac{\epsilon}{6}$. By (6.65), we can find an $M \in \mathbb{N}_0$ such that for all $m > M$,

$$\lim_{m\to\infty} \mathbb{P}_{(0,D_m),(0,A)}(\tau^* \leq \tilde{t}) < \tfrac{\epsilon}{6}, \tag{6.68}$$

and hence

$$\lim_{m\to\infty} \sum_{\substack{j \in \mathbb{G} \\ \|j\| \leq L}} \mathbb{E}_{(0,A),(0,D_m)}[\hat{a}_{T(\tau^*)+T'(\tau^*)}(0,j)] < \tfrac{\epsilon}{6} + \tfrac{\epsilon}{6} + \tfrac{\epsilon}{6} = \tfrac{\epsilon}{2}. \quad \square \tag{6.69}$$

4. Uniqueness of the invariant measure ν_θ when $\rho = \infty$.

Lemma 6.13 (Uniqueness of and convergence to ν_θ). *For all $\theta \in (0,1)$ there exists a unique invariant measure ν_θ such that $\lim_{t\to\infty} \mu(t) = \nu_\theta$ for all $\mu(0) \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$.*

Proof. Suppose that ν_θ^1 and ν_θ^2 and are two different weak limit points of $\mu(t_n)$ as $t_n \to \infty$, and that $\mu(0) \in \mathcal{R}_\theta^{(2),\bullet}$. Let $(\bar{Z}(t))_{t\geq 0} = (Z^1(t), Z^2(t))_{t\geq 0}$ be the coupled process from (6.29)–(6.30) with $\mathcal{L}(\bar{Z}(0)) = \bar{\nu}_\theta$, and marginals $\mathcal{L}(Z^1(0)) = \nu_\theta^1$, $\mathcal{L}(Z^2(0)) = \nu_\theta^2$. Define the process Y^1 by

$$\begin{aligned} Y^1 &= \big(Y^1(m)\big)_{m\in\{-1\}\cup\mathbb{N}_0}, \\ Y^1(-1) &= (x_i^1(0))_{i\in\mathbb{G}}, \qquad Y^1(m) = (y_{i,m}^1(0))_{i\in\mathbb{G}}, \quad m \in \mathbb{N}_0. \end{aligned} \tag{6.70}$$

Thus, Y^1 has state space $([0,1]^{\mathbb{G}})^{\{-1\}\cup\mathbb{N}_0}$ and $\mathcal{L}(Y^1) = \mathcal{L}(Z^1(0)) = \nu_\theta^1$. We can interpret Y^1 as a process that describes the states of the population *in the seed-bank direction*. Similarly, define the process Y^2 by

$$\begin{aligned} Y^2 &= \big(Y^2(m)\big)_{m\in\{-1\}\cup\mathbb{N}_0}, \\ Y^2(-1) &= (x_i^2(0))_{i\in\mathbb{G}}, \qquad Y^2(m) = (y_{i,m}^2(0))_{i\in\mathbb{G}}, \quad m \in \mathbb{N}_0. \end{aligned} \tag{6.71}$$

Thus, Y^2 has state space $([0,1]^{\mathbb{G}})^{\{-1\}\cup\mathbb{N}_0}$ and $\mathcal{L}(Y^2) = \mathcal{L}(Z^2(0)) = \nu_\theta^2$.

Define the σ-algebras $\mathcal{B}_M^1$ and $\mathcal{B}^1$, respectively, $\mathcal{B}_M^2$ and $\mathcal{B}^2$ by

$$\mathcal{B}^k = \cap_{M\in\mathbb{N}_0} \mathcal{B}_M^k, \qquad \mathcal{B}_M^k = \sigma\big(y_{i,m}^k : i \in \mathbb{G}, \, m \geq M \big), \quad k \in \{1,2\}. \tag{6.72}$$

Here, $\mathcal{B}^1$ and $\mathcal{B}^2$ are *the tail-σ-algebras in the seed-bank direction*. By Lemma 6.12, we have

$$\lim_{m\to\infty} \mathcal{L}_{\nu_\theta^1}(y_{i,m}) = \lim_{m\to\infty} \mathcal{L}_{\nu_\theta^2}(y_{i,m}) = \delta_\theta. \tag{6.73}$$

Hence, $\mathcal{B}^1 = \mathcal{B}^2$, both are trivial, and ν_θ^1 and ν_θ^2 agree on them. Therefore Goldstein's Theorem [Gol79] implies that there exists a successful coupling of Y^1 and Y^2. Consequently, there exists a random variable $T^{\mathrm{coup}} \in \{-1\} \cup \mathbb{N}_0$ such that, for all $m \geq T^{\mathrm{coup}}$, $Y^1(m) = Y^2(m)$, i.e., $|\delta_{i,m}(0)| = 0$ for all $i \in \mathbb{G}$ and $\mathbb{P}(T^{\mathrm{coup}} < \infty) = 1$. Hence

$$\hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_i(0)|\right] = \hat{\mathbb{E}}\left[|\Delta_i(0)| + \sum_{m=0}^{T^{\mathrm{coup}}} K_m|\delta_i(0)|\right]. \tag{6.74}$$

However, we cannot conclude that the left-hand side of (6.74) is finite. Therefore, let $\bar{\nu}_\theta|_{\{T^{\mathrm{coup}}<T\}}$ denote the restriction of the measure $\bar{\nu}_\theta$ to the set $\{T^{\mathrm{coup}} < T\}$. Since

$\{T^{\mathrm{coup}} < T\}$ is a translation-invariant event in the spatial direction, the measure $\bar{\nu}_\theta|_{\{T^{\mathrm{coup}}<T\}}$ is translation invariant. Moreover,

$$\hat{\mathbb{E}}_{\bar{\nu}_\theta|_{\{T^{\mathrm{coup}}<T\}}}\left[|\Delta_i(0)| + \sum_{m\in\mathbb{N}_0} K_m|\delta_i(0)|\right] = \hat{\mathbb{E}}_{\bar{\nu}_\theta|_{\{T^{\mathrm{coup}}<T\}}}\left[|\Delta_i(0)| + \sum_{m=0}^T K_m|\delta_i(0)|\right] < \infty.$$
(6.75)

Therefore we can use the dynamics in (6.32) and conclude that, for all $T \in \mathbb{N}$, $\hat{\mathbb{P}}_{\bar{\nu}_\theta|_{\{T^{\mathrm{coup}}<T\}}}(E_0 \times E_0) = 1$ (recall (6.43)). Since $\lim_{T\to\infty} \bar{\nu}_\theta|_{\{T^{\mathrm{coup}}<T\}} = \bar{\nu}_\theta$, it follows that

$$\hat{\mathbb{P}}_{\bar{\nu}_\theta}(E_0 \times E_0) = 1.$$
(6.76)

By (6.48) and (6.49), we conclude that $\nu_\theta^1 = \nu_\theta^2$ and hence that all weak limit points of $(\mu(t))_{t\geq 0}$ are the same. Suppose now that $\mu^1(0), \mu^2(0) \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ are two different initial measures. By the above argument, we know that $\lim_{t\to\infty} \mu^1(t) = \nu_\theta^1$ and $\lim_{t\to\infty} \mu^2(t) = \nu_\theta^2$. By Lemma 6.12, we know that ν_θ^1 and ν_θ^2 have the same trivial tail-σ-algebras in the seed-bank direction. Hence, repeating the above argument, we find that $\nu_\theta^1 = \nu_\theta^2$. We conclude that for each initial measure $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$ the SSDE in (2.12)–(2.13) converges to a unique non-trivial equilibrium measure ν_θ. $\qquad\square$

5. Ergodicity, mixing and associatedness. The equivalent of Lemma 5.13 for $\rho = \infty$ follows in the same way as for $\rho < \infty$.

6.4 Proof of the dichotomy

Theorem 3.3(I)(a) follows from Lemma (6.8) and Steps 3-5 in Section 6.3.1. The equality $\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta$, $m \in \mathbb{N}_0$, follows from (2.12)–(2.13), the fact that ν_θ is an equilibrium measure, and the preservation of θ (see Section 3.2). Theorem 3.3(I)(b) follows by combining Lemma 6.3 with the analogue of Lemma 5.5. Theorem 3.3(II) follows from Lemmas 6.3, 6.12, 6.13, the analogue of Lemma 5.5, and Step 6 in Section 6.3.2. The equality $\mathbb{E}_{\nu_\theta}[x_0] = \mathbb{E}_{\nu_\theta}[y_{0,m}] = \theta$, $m \in \mathbb{N}_0$, follows from (6.8) in Step 1 of Section 6.3.2.

The fact that, for $\rho < \infty$, $\theta \mapsto \nu_\theta$ is continuous in the weak topology, as claimed in Remark 3.4, follows from the fact that $\|\nu_\theta(z_i \in \cdot) - \nu_{\theta'}(z_i \in \cdot)\|_{tv} \leq (1+\rho)|\theta - \theta'|$ for all $i \in \mathbb{G}$, with $z_i = (z_{(i,R_i)})_{R_i\in\{A,(D_m)_{m\in\mathbb{N}_0}\}}$, as follows from the coupling argument in Lemmas 6.7–6.8. (Pick $\mu(0)$ equal to $\delta_\theta, \delta_{\underline{\theta'}}$ and use the contraction via the Lyapunov function.) The fact that, for $\rho = \infty$, $\theta \mapsto \nu_\theta$ is continuous in the uniform weak topology, as claimed in Remark 3.4, uses the coupling argument in Lemma 6.10–6.11 together with the fact that the deep seed-banks are deterministic stated in Lemma 6.12. Lemma 6.9 guarantees that the limit in (3.31) equals θ μ-a.s. for every $\mu \in \mathcal{T}_\theta^{\mathrm{erg},\bullet}$.

Corollary 3.5(1) corresponds to $\gamma \in (1,\infty)$ and $\rho < \infty$, and migration dominates. Corollary 3.5(2) corresponds to $\gamma \in [\frac{1}{2}, 1]$ and $\rho = \infty$, and $I_{\hat{a},\gamma}$ shows in interplay between migration and seed-bank. Corollary 3.5(3) corresponds to $\gamma \in (0, \frac{1}{2}, 1)$ and $\rho = \infty$, and the seed-bank dominates: $I_{\hat{a},\gamma} < \infty$ because $\hat{a}_t(0,0) \leq 1$.

6.5 Different dichotomy for asymmetric migration

It remains to explain how the counterexample below Theorem 3.3 arises. We focus on the case when $\rho < \infty$, which implies $\mathbb{E}(\tau) < \infty$, but we assume $\mathbb{E}(\tau^2) = \infty$. Therefore the central limit theorem does not hold for $T(t)$, $T'(t)$, and $\Delta(t) \gg \sqrt{M(t)}$. Hence (5.38) must be replaced by

$$f(t) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \mathrm{d}\phi\, \mathrm{e}^{-[1+o(1)]\,2Bt\,[1-\hat{a}(\phi)]}\,\mathbb{E}\left[\cos\left(\Delta(t)\tilde{a}(\phi)\right)\right].$$
(6.77)

The key observation is that if $\tilde{a}(\phi) \neq 0$ (due to the asymmetry of $a(\cdot, \cdot)$; recall (5.29)), then the expectation in (6.77) can change the integrability properties of $f(t)$.

Under the assumption that τ has a *one-sided stable distribution* with parameter $\gamma \in (1,2)$, we have (5.33) with $A = \chi/(1+\rho)$ and $B = 1/(1+\rho)$, while there exists a constant $C \in (0, \infty)$ such that (see [Fel71, Chapter XVII])

$$\mathbb{E}[\cos(\Delta(t)\tilde{a}(\phi))] = e^{-[1+o(t)]\,At|C\tilde{a}(\phi)|^{\gamma}}. \tag{6.78}$$

Substituting (6.78) into (6.77), we see that for large t the contribution to $f(t)$ comes from ϕ such that $\hat{a}(\phi) \to 1$ and $\tilde{a}(\phi) \to 0$. By our choice of the migration kernel in (3.32), this holds as $\phi = (\phi_1, \phi_2) \to (0,0)$. Using that $1 - \hat{a}(\phi) \sim \frac{1}{2}(\phi_1^2 + \phi_2^2)$ and $\tilde{a}(\phi) \sim \frac{1}{2}\eta(\phi_1 + \phi_2)$ for $(\phi_1, \phi_2) \to (0,0)$, we find that (6.77) equals

$$f(t) = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} d\phi \, e^{-[1+o(1)]\,\{Bt(\phi_1^2+\phi_2^2)+At[|\frac{1}{2}C\eta(\phi_1+\phi_2)|]^{\gamma}\}}, \qquad t \to \infty. \tag{6.79}$$

Hence the integral in (6.79) is determined by ϕ such that

$$B(\phi_1^2 + \phi_2^2) + A\left[|\tfrac{1}{2}C\eta(\phi_1 + \phi_2)|\right]^{\gamma} \leq \frac{c}{t}. \tag{6.80}$$

for c a positive constant, and we find that $f(t) \asymp t^{-\left(\frac{1}{\gamma}+\frac{1}{2}\right)}$. Since $\gamma \in (1,2)$, $f(t)$ is much smaller than $\hat{a}_t(0,0) \asymp 1/t$, valid for two-dimensional simple random walk. Thus we see that $t \mapsto f(t)$ is integrable, while $t \mapsto \hat{a}_t(0,0)$ is not.

6.6 Modulation of the law of the wake-up times by a slowly varying function

The integral in (3.38) is the *total hazard of coalescence* of two dual lineages:

- If $\gamma \in (0,1)$, then the probability for each of the lineages to be active at time s decays like $\asymp \varphi(s)^{-1}s^{-(1-\gamma)}$ [AB16]. Hence the expected total time they are active up to time s is $\asymp \varphi(s)^{-1}s^{\gamma}$. Because the lineages only move when they are active, the probability that the two lineages meet at time s is $\asymp a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. Hence the total hazard is $\asymp \int_1^{\infty} ds\, \varphi(s)^{-2}s^{-2(1-\gamma)}\, a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. After the transformation $t = t(s) = \varphi(s)^{-1}s^{\gamma}$, we get the integral in (3.38), modulo a constant. (When carrying out this transformation, we need that $\lim_{s\to\infty} s\varphi'(s)/\varphi(s) = 0$, which is immediate from (3.37), and $\varphi(t(s))/\varphi(s) \asymp 1$ as $s \to \infty$, which is immediate from the bound we imposed on ψ together with the fact that $\lim_{s\to\infty} \log\varphi(s)/\log s = 0$.)
- If $\gamma = 1$, then the probability for each of the lineages to be active at time s decays like $\hat{\varphi}(s)^{-1}$ [AB16]. Hence the expected total time they are active up to time s is $\asymp s\hat{\varphi}(s)^{-1}$. Hence the total hazard is $\asymp \int_1^{\infty} ds\, \hat{\varphi}(s)^{-2}\, a^{(N)}_{\hat{\varphi}(s)^{-1}s}(0,0)$. After the transformation $t = t(s) = \hat{\varphi}(s)^{-1}s$, we get the integral in (3.38), modulo a constant.

7 Proofs: long-time behaviour for Model 3

The arguments for Model 2 in Section 6 all carry over with minor adaptations. The only difference is that for $\rho = \infty$ the clustering criterion changes. In this section we prove the new clustering criterion and comment on the modifications needed in the corresponding proofs for Model 2 in Section 6.

7.1 Moment relations

Like in Model 1 and 2, we can relate the first and second moments of the system in (2.18)–(2.19) to the random walk that evolves according to the transition kernel $b^{(3)}(\cdot, \cdot)$ on $\mathbb{G} \times \{A, (D_m)_{m \in \mathbb{N}_0}\}$ given by (2.56). Replacing in Lemma 6.1 the kernel $b^{(2)}(\cdot, \cdot)$ by $b^{(3)}(\cdot, \cdot)$, we find the moment relation for Model 3. Also here these moment relations hold for all $g \in \mathcal{G}$. Moreover these moment relations holds for $\rho < \infty$ as well as for $\rho = \infty$.

7.2 The clustering case

To obtain the equivalent of Lemma 6.3, we need to replace the kernel $\hat{a}(\cdot,\cdot)$ by the convoluted kernel $(\hat{a}*\hat{a}^\dagger)(\cdot,\cdot)$. Each time one of the two copies of the random walk with migration kernel $a(\cdot,\cdot)$ moves from the active state to the dormant state, it makes a transition according to the displacement kernel $a^\dagger(\cdot,\cdot)$ (recall (3.39)). Therefore the expression in (5.23) needs to be replaced by

$$I = \int_0^\infty \mathrm{d}t \sum_{k,k'\in\mathbb{N}} \sum_{i,i'\in\mathbb{G}} \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,A)}\Big[\hat{a}_{T(k,t)}(0,i)\,\hat{a}_{T'(k',t)}(0,i')\,\hat{a}_k^\dagger(i,j)\,\hat{a}_{k'}^\dagger(i',j)\,1_{\mathcal{E}(k,t)}\,1_{\mathcal{E}'(k',t)}\Big],$$
(7.1)

where $\hat{a}_k^\dagger(\cdot,\cdot)$ is the step-k transition kernel of the random walk with displacement kernel $\hat{a}^\dagger(\cdot,\cdot)$. Using the *symmetry* of both kernels, we can carry out the sum over j,i' and write

$$\begin{aligned}
I &= \int_0^\infty \mathrm{d}t \sum_{k,k'\in\mathbb{N}} \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,A)}\Big[\hat{a}_{T(k,t)+T'(k',t)}(0,j)\,\hat{a}_{k+k'}^\dagger(0,j)\,1_{\mathcal{E}(k,t)}\,1_{\mathcal{E}'(k',t)}\Big] \\
&= \int_0^\infty \mathrm{d}t \sum_{j\in\mathbb{G}} \mathbb{E}_{(0,A)}\Big[\hat{a}_{T(t)+T'(t)}(0,j)\,\hat{a}_{N(t)+N'(t)}^\dagger(0,j)\,1_{\mathcal{E}(t)}\,1_{\mathcal{E}'(t)}\Big] \\
&= \int_0^\infty \mathrm{d}t\, \mathbb{E}_{(0,A)}\Big[\big(\hat{a}_{T(t)+T'(t)}*\hat{a}_{N(t)+N'(t)}^\dagger\big)(0,0)\,1_{\mathcal{E}(t)}\,1_{\mathcal{E}'(t)}\Big].
\end{aligned}$$
(7.2)

The last expression is the analogue of (5.26).

For $\rho < \infty$, following the same line of argument as for Model 2, we find with the help of (3.40) that

$$I \asymp \int_1^\infty \mathrm{d}t\,(\hat{a}_t*\hat{a}_t^\dagger)(0,0).$$
(7.3)

For $\rho = \infty$, with the help of the Fourier transform we compute

$$\begin{aligned}
&\mathbb{E}_{(0,A)}\Big[\big(a_{T(t)+T'(t)}*a_{N(t)+N'(t)}^\dagger\big)(0,0)\Big] \\
&= \mathbb{E}_{(0,A)}\left[\frac{1}{(2\pi)^d}\int_{(-\pi,\pi]^d} \mathrm{d}\phi\,\mathrm{e}^{-(T(t)+T'(t))[1-\hat{a}(\phi)]}\,\hat{a}^\dagger(\phi)^{N(t)+N'(t)}\right] \\
&= \frac{1}{(2\pi)^d}\int_{(-\pi,\pi]^d} \mathrm{d}\phi\,\mathrm{e}^{-[1+o(1)]\,2ct^{-\gamma}\,[1-\hat{a}(\phi)]}\,\mathrm{e}^{-[1+o(1)]\,2t^{-\gamma}[1-\hat{a}^\dagger(\phi)]} \\
&\asymp (\hat{a}_{ct^{-\gamma}}*\hat{a}_{t^{-\gamma}}^\dagger)(0,0) \asymp (\hat{a}_{t^{-\gamma}}*\hat{a}_{t^{-\gamma}}^\dagger)(0,0),
\end{aligned}$$
(7.4)

where we use (3.40), (6.12) and the fact that deviations of $T(t)/t^\gamma$ and $T'(t)/t^\gamma$ away from order 1 are stretched exponentially costly in t [EJU19]. Hence

$$I \asymp \int_1^\infty \mathrm{d}t\,t^{-2(1-\gamma)}(\hat{a}_{t^\gamma}*\hat{a}_{t^\gamma}^\dagger)(0,0).$$
(7.5)

Putting $s = t^\gamma$ we obtain, instead of (6.22),

$$I = \infty \quad\Longleftrightarrow\quad I_{\hat{a}*\hat{a}^\dagger,\gamma} = \infty$$
(7.6)

with

$$I_{\hat{a}*\hat{a}^\dagger,\gamma} = \int_1^\infty \mathrm{d}s\,s^{-(1-\gamma)/\gamma}\,(\hat{a}_s*\hat{a}_s^\dagger)(0,0),$$
(7.7)

which is precisely the integral in (3.42).

7.3 The coexistence case

The coexistence results in Theorem 3.8 follow for both $\rho < \infty$ and $\rho = \infty$ by the same type of argument as the one we used for Model 2 in Section 6.3. We replace (2.12)–(2.13) by (2.18)–(2.19), replace $b^{(2)}(\cdot, \cdot)$ (see (2.44)) by $b^{(3)}(\cdot, \cdot)$ (see (2.56)), and use the Fourier transform of $\hat{a} * \hat{a}^{\dagger}(\cdot, \cdot)$ instead of $\hat{a}(\cdot, \cdot)$. The key of the argument is that, in the coexistence case, for $\rho < \infty$ we have $I_{\hat{a}*\hat{a}^{\dagger}} < \infty$, while for $\rho = \infty$ we have $I_{\hat{a}*\hat{a}^{\dagger}, \gamma} < \infty$.

7.4 Proof of the dichotomy

This follows in exactly the same way as for Model 2.

A Derivation of continuum model from individual-based model

Model 1. We give the derivation of (2.4)–(2.5) as the continuum limit of an individual-based model when the size of the colonies tends to infinity. We start with the continuum limit of the Fisher-Wright model with (strong) seed-bank for a *single-colony* model as defined in [BGCKWB16]. Subsequently we show how the limit extends to a *multi-colony* model with seed-bank.

Single-colony model. The Fisher-Wright model with (strong) seed-bank defined in [BGCKWB16] consists of a *single colony* with $N \in \mathbb{N}$ active individuals and $M \in \mathbb{N}$ dormant individuals. Each individual can carry one of two types: $\heartsuit$ or $\diamondsuit$. Let $\epsilon \in [0, 1]$ be such that ϵN is integer and $\epsilon N \leq M$. Put $\delta = \frac{\epsilon N}{M}$. The evolution of the population is described by a discrete-time Markov chain that undergoes four independent transitions per step:

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

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

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

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

Note that the total sizes of the active and the dormant population remain fixed. During the evolution the dormant and active population *exchange* individuals. We are interested in the fractions of individuals of type $\heartsuit$ in the active and the dormant population.

Let $c = \epsilon N = \delta M$, i.e., c is the number of pairs of individuals that change state. Label the N active individuals from 1 to N and the M dormant individuals from 1 up to M. We denote by $[N] = \{1, \ldots, N\}$ and by $[M] = \{1, \ldots, M\}$. Let $\xi(k) = (\xi_j(k))_{j \in [N]} \in \{0, 1\}^{[N]}$ be the random vector where $\xi_j(k) = 1$ if the j-th active individual is of type $\heartsuit$ at time k and $\xi_j(k) = 0$ if the j-th active individual is of type $\diamondsuit$ at time k. Similarly, we let $\eta(k) = (\eta_j(k))_{j \in [M]} \in \{0, 1\}^{[M]}$ be the random vector where $\eta_j(k) = 1$ if the j-th dormant individual is of type $\heartsuit$ at time k and $\eta_j(k) = 0$ if the j-th dormant individual is of type $\diamondsuit$ at
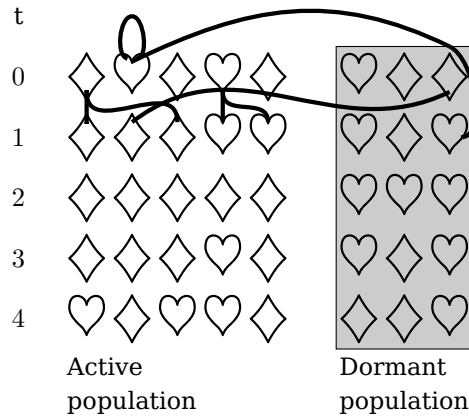
Figure 6: Example of the evolution for a population with $N = 5$ active individuals and $M = 3$ dormant individuals. The solid lines within the active population represent resampling, those between the active and the dormant population represent exchange with the seed-bank. Only 1 active individual and 1 dormant individual exchange places per unit of time, which corresponds to $\epsilon = \frac{1}{5}$ and $\delta = \frac{1}{3}$. The relative size of the dormant and the active population is $K = \frac{3}{5}$. Note that the genetic diversity in the active population is lost in generation $t = 2$, but returns in generation $t = 3$ via the seed-bank.

time k. Let $I^N = \{0, \frac{1}{N}, \frac{2}{N}, \frac{3}{N} \ldots, 1\}$ and $I^M = \{0, \frac{1}{M}, \frac{2}{M}, \frac{3}{M} \ldots, 1\}$. Define the variables

$$X^N(k) = \frac{1}{N} \sum_{j \in [N]} \mathbf{1}_{\{\xi_j(k)=1\}} \quad \text{on } I^N,$$

$$Y^N(k) = \frac{1}{N} \sum_{j \in [N]} \mathbf{1}_{\{\eta_j(k)=1\}} \quad \text{on } I^M. \tag{A.1}$$

Let $\mathbb{P}_{x,y}$ denote the law of the Markov process

$$(X^N, Y^N) = (X^N(k), Y^N(k))_{k \in \mathbb{N}_0} \tag{A.2}$$

given that $(X^N(0), Y^N(0)) = (x, y) \in I^N \times I^M$. Then, as shown in [BGCKWB16],

$$
\begin{aligned}
p_{x,y}(\bar{x}, \bar{y}) &= \mathbb{P}_{x,y}(X_1^N = \bar{x}, Y_1^N = \bar{y}) \\
&= \sum_{c'=0}^{c} \mathbb{P}_{x,y}(Z = c') \, \mathbb{P}_{x,y}(U = \bar{x}N - c') \, \mathbb{P}_{x,y}(V = \bar{y}M - yM + c').
\end{aligned} \tag{A.3}
$$

Here, Z denotes the number of dormant $\heartsuit$-individuals in generation 0 that become active in generation 1 ($\mathcal{L}_{x,y}(Z) = \text{Hyp}_{M,c,yM}$), U denotes the number of active individuals in generation 1 that are offspring of active $\heartsuit$-individuals in generation 0 ($\mathcal{L}_{x,y}(U) = \text{Bin}_{N-c,x}$), and V denotes the number of active individuals in generation 0 that become dormant $\heartsuit$-individuals in generation 1 ($\mathcal{L}_{x,y}(V) = \text{Bin}_{c,x}$).

Speed up time by a factor N. The discrete-time generator G^N for the process $((X^N(\lfloor Nk \rfloor), Y^N(\lfloor Nk \rfloor)))_{k \in \mathbb{N}_0}$ equals

$$
\begin{aligned}
(G^N f)(x, y) &= N \, \mathbb{E}_{x,y}\big[f(X^N(1), Y^N(1)) - f(x, y) \big], \\
(x, y) &\in I^N \times I^M,
\end{aligned} \tag{A.4}
$$

where the prefactor N appears because one step of the Markov chain takes time $\frac{1}{N}$. Inserting the Taylor expansion for f (which we assume to be smooth), using that $X^N(1) =$

$\frac{U+Z}{N}$ and $Y^N(1) = \frac{yM+V-U}{M}$ and letting $N \to \infty$, we end up with the limiting continuous-time generator G given by

$$(Gf)(x,y) = c(y-x)\frac{\partial f}{\partial x}(x,y) + \frac{c}{K}(x-y)\frac{\partial f}{\partial y}(x,y) + \tfrac{1}{2}x(1-x)\frac{\partial^2 f}{\partial x^2}(x,y), \tag{A.5}$$
$$(x,y) \in [0,1] \times [0,1],$$

where $K = \frac{M}{N}$ is the relative size of the dormant population compared to the active population. This is the generator of the Markov process in the continuum limit [EK86, Section 7.8]. It follows from the form of G that this limit is described by the system of coupled stochastic differential equations

$$\mathrm{d}x(t) = c\,[y(t) - x(t)]\,\mathrm{d}t + \sqrt{x(t)(1-x(t))}\,\mathrm{d}w(t),$$
$$\mathrm{d}y(t) = \frac{c}{K}\,[x(t) - y(t)]\,\mathrm{d}t. \tag{A.6}$$

This is the version of (2.4)–(2.5) for a single colony (no migration) and exchange rate

$$e = \frac{c}{K}. \tag{A.7}$$

Multi-colony model. First fix a number $L \in \mathbb{N}$ and consider $|\mathbb{G}| = L$ colonies. The *multi-colony* version with migration is obtained by allowing the $(1-\epsilon)N$ selected active individuals to undergo a migration in step (1):

(1) Each active individual at colony $i \in \mathbb{G}$ chooses colony $j \in \mathbb{G}$ with probability $\frac{1}{N}a(i,j)$ and adopts the type of a parent chosen from colony j. If an active individual does not migrate, it adopts the type of a parent chosen from its own population.

Using the same strategy as in the single-colony model, this results in (2.4)–(2.5), for $|\mathbb{G}| = L$. Subsequently we can let $L \to \infty$ and use convergence of generators to obtain (2.4)–(2.5) for countable $\mathbb{G}$.

Model 2. The same argument works for (2.12)–(2.13). Steps (1)-(4) are extended by considering a seed-bank with colours labelled by $\mathbb{N}_0$. First we consider the truncation where only finitely many colours are allowed, for which the argument carries through with minor adaptations. Afterwards, we pass to the limit of infinitely many colours, which is straightforward for a finite time horizon because large colours are only seen after large times. See also [Moj18].

Model 3. To get (2.18)–(2.19), also extend Step (3) by adding a displacement via the kernel $a^\dagger(\cdot,\cdot)$ for each transition into the seed-bank.

B Moran versions of Models 1–2

In this appendix we consider the Moran version of the single-colony individual-based model in Appendix A, i.e., we allow individuals to become active or dormant *without exchange*. Our goal is to show that, in the large-colony-size limit, this modification has no effect on the long-time behaviour: after an initial transient period the evolution converges to that of the model *with exchange*, after an appropriate tuning of parameters. The brief exposition below is based on [Moj18].

In the Moran version each active individual resamples at rate 1 and becomes dormant at a certain rate, while each dormant individual does not resample and becomes active at a certain rate. Since switches between active and dormant are done independently,

the sizes of the active and the dormant population are *no longer* fixed and individuals *change* state without the necessity to *exchange* state. In Model 1 there are two Poisson clocks, in Model 2 there are two sequences of Poisson clocks, namely, two for each colour. In Appendices B.1–B.2 we compute the scaling limit for the case where the number of colours is $\mathfrak{m} = 1$ and $\mathfrak{m} = 2$, respectively. The extension to $\mathfrak{m} \geq 3$ is given in Appendix B.3.

Migration can be added in the same way as is done in Appendix A. We leave this extension to the reader.

B.1   Moran version of Model 1

To describe the Moran version of Model 1 we need the following variables.

- Total number of individuals: $N \in \mathbb{N}$.
- Two types: $\heartsuit$ and $\diamondsuit$.
- $X(t)$ is the number of $\heartsuit$-individuals in the active population at time t.
- $Y(t)$ is the number of $\heartsuit$-individuals in the dormant population at time t.
- $Z(t)$ is the number of individuals in the active population at time t (either $\heartsuit$ or $\diamondsuit$).

In the Moran model with seed-bank each active individual resamples at rate 1, each active individual becomes dormant at rate ϵ and each dormant individual becomes active at rate δ. Hence the transition rates for $(X(t), Y(t), Z(t))$ are:

- $(i, j, k) \to (i+1, j, k)$ at rate $(k-i)\frac{i}{k}$.
- $(i, j, k) \to (i-1, j, k)$ at rate $i\frac{(k-i)}{k}$.
- $(i, j, k) \to (i-1, j+1, k-1)$ at rate ϵi.
- $(i, j, k) \to (i+1, j-1, k+1)$ at rate δj.
- $(i, j, k) \to (i, j, k-1)$ at rate $\epsilon \frac{k-i}{N}$.
- $(i, j, k) \to (i, j, k+1)$ at rate $\delta \frac{N-k-j}{N}$.

For the scaling limit we consider the variables

$$\bar{X}(t) = \frac{1}{N}X(Nt), \quad \bar{Y}(t) = \frac{1}{N}Y(Nt), \quad \bar{Z}(t) = \frac{1}{N}Z(Nt). \tag{B.1}$$

Hence

$$(\bar{X}(t), \bar{Y}(t), \bar{Z}(t)) \in I^N \times I^N \times I^N, \qquad I^N = \left\{0, \tfrac{1}{N}, \tfrac{2}{N}, \ldots, \tfrac{N-1}{N}, 1\right\}. \tag{B.2}$$

Since in (B.1) we speed up time by a factor N, we must also speed up the transition rates by a factor N. To get a meaningful scaling limit, we assume that there exist $c^A, c^D \in (0, \infty)$ such that (see [BGCKWB16, p. 8])

$$N\epsilon = c^A, \qquad N\delta = c^D, \qquad N \in \mathbb{N}. \tag{B.3}$$

We can then write down the generator G^N:

$$
\begin{aligned}
(G^N f)\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right) = {}& N(k-i)\frac{i}{k}\left[f\left(\frac{i+1}{N}, \frac{j}{N}, \frac{k}{N}\right) - f\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)\right] \\
& + Ni\frac{k-i}{k}\left[f\left(\frac{i-1}{N}, \frac{j}{N}, \frac{k}{N}\right) - f\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)\right] \\
& + c^A i\left[f\left(\frac{i-1}{N}, \frac{j+1}{N}, \frac{k-1}{N}\right) - f\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)\right] \\
& + c^D j\left[f\left(\frac{i+1}{N}, \frac{j-1}{N}, \frac{k+1}{N}\right) - f\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)\right] \\
& + c^A(k-i)\left[f\left(\frac{i}{N}, \frac{j}{N}, \frac{k-1}{N}\right) - f\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)\right] \\
& + c^D(N-k-j)\left[f\left(\frac{i}{N}, \frac{j}{N}, \frac{k+1}{N}\right) - f\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)\right]
\end{aligned}
\tag{B.4}
$$

Assuming that f is smooth and Taylor expanding f around $\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right)$, we get

$$
\begin{aligned}
(G^N f)\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right) = {} & \frac{i(k-i)}{k}\left[\left(\frac{1}{N}\right)\frac{\partial^2 f}{\partial x^2} + \mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right] \\
& + c^A i\left[\left(\frac{-1}{N}\right)\frac{\partial f}{\partial x} + \left(\frac{1}{N}\right)\frac{\partial f}{\partial y} + \left(\frac{-1}{N}\right)\frac{\partial f}{\partial z} + \mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right] \\
& + c^D j\left[\left(\frac{1}{N}\right)\frac{\partial f}{\partial x} + \left(\frac{-1}{N}\right)\frac{\partial f}{\partial y} + \left(\frac{1}{N}\right)\frac{\partial f}{\partial z} + \mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right] \\
& + c^A(k-i)\left[\left(\frac{-1}{N}\right)\frac{\partial f}{\partial z} + \mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right] \\
& + c^D(N-k-j)\left[\left(\frac{1}{N}\right)\frac{\partial f}{\partial z} + \mathcal{O}\left(\left(\frac{1}{N}\right)^2\right)\right].
\end{aligned}
\tag{B.5}
$$

Next, suppose that

$$
\lim_{N\to\infty}\frac{i}{N} = x, \qquad \lim_{N\to\infty}\frac{j}{N} = y, \qquad \lim_{N\to\infty}\frac{k}{N} = z.
\tag{B.6}
$$

Letting $N\to\infty$ in (B.5), we obtain the limiting generator G:

$$
\begin{aligned}
(Gf)(x,y,z) = {} & z\frac{x}{z}\left(1 - \frac{x}{z}\right)\left(\frac{\partial^2 f}{\partial x^2}\right) + [c^D y - c^A x]\frac{\partial f}{\partial x} \\
& + [c^A x - c^D y]\frac{\partial f}{\partial y} + [c^D(1-z) - c^A z]\frac{\partial f}{\partial z}.
\end{aligned}
\tag{B.7}
$$

Therefore the continuum limit equals

$$
\begin{aligned}
\mathrm{d}x(t) &= \sqrt{z(t)\frac{x(t)}{z(t)}\left(1 - \frac{x(t)}{z(t)}\right)}\,\mathrm{d}w(t) + \left[c^D y(t) - c^A x(t)\right]\mathrm{d}t, \\
\mathrm{d}y(t) &= \left[c^A x(t) - c^D y(t)\right]\mathrm{d}t, \\
\mathrm{d}z(t) &= \left[c^D(1 - z(t)) - c^A z(t)\right]\mathrm{d}t.
\end{aligned}
\tag{B.8}
$$

Since $z(t)$ is the fraction of active individuals in the population, $1 - z(t)$ is the fraction of dormant individuals in the population. Therefore the equivalent of the parameter K in Appendix A is $K(t) = (1 - z(t))/z(t)$. Moreover, $x(t)/z(t)$ is the fraction of $\heartsuit$-individuals in the active population at time t and $y(t)/(1 - z(t))$ is the fraction of $\heartsuit$-individuals in the dormant population at time t. The last line of (B.8) is an autonomous differential equation whose solution converges to

$$
z^* = \frac{1}{1 + \frac{c^A}{c^D}}
\tag{B.9}
$$

exponentially fast. After this transition period we can replace $z(t)$ by z^*, and we see that $K^* = c^A/c^D$.

Time is to be scaled by the total number of active *and* dormant individuals, instead of the total number of active individuals only:

$$
\begin{aligned}
x(t) &= \frac{\text{number of active individuals of type } \heartsuit}{\text{total number of individuals}}, \\
y(t) &= \frac{\text{number of dormant individuals of type } \heartsuit}{\text{total number of individuals}}.
\end{aligned}
\tag{B.10}
$$

To compare the Moran model with a 1-colour seed-bank with the Fisher-Wright model with a 1-colour seed-bank, we look at the variables

$$\bar{x}(t) = \left(1 + \frac{c^A}{c^D}\right) x \left(\frac{t}{1 + \frac{c^A}{c^D}}\right), \qquad \bar{y}(t) = \left(1 + \frac{c^A}{c^D}\right) \left(\frac{c^D}{c^A}\right) y \left(\frac{t}{1 + \frac{c^A}{c^D}}\right). \qquad \text{(B.11)}$$

After a short transition period in which $z(t)$ tends to z^*, we see that by setting

$$K = K^* = \frac{c^A}{c^D}, \qquad e = \frac{c^D}{c^A} \frac{1}{\frac{1}{c^A} + \frac{1}{c^D}}, \qquad \text{(B.12)}$$

we obtain

$$\begin{aligned} \mathrm{d}\bar{x}(t) &= \sqrt{\bar{x}(t)(1 - \bar{x}(t))} \, \mathrm{d}w(t) + Ke\left[\bar{y}(t) - \bar{x}(t)\right] \mathrm{d}t, \\ \mathrm{d}\bar{y}(t) &= e\left[\bar{x}(t) - \bar{y}(t)\right] \mathrm{d}t, \end{aligned} \qquad \text{(B.13)}$$

which is the single-colony version of (2.4)–(2.5) but without migration. Migration can be added in the same way as was done in Appendix A.

B.2 Moran version of Model 2: Two colours

We consider the following system:

- Total number of individuals: $N \in \mathbb{N}$.
- Two types: $\heartsuit$ and $\diamondsuit$.
- $X(t)$ is the number of $\heartsuit$-individuals in the active population at time t.
- $Y_1(t)$ is the number of $\heartsuit$-individuals of colour 1 in the dormant population at time t.
- $Y_2(t)$ is the number of $\heartsuit$-individuals of colour 2 in the dormant population at time t.
- $Z_{D_1}(t)$ is the number of dormant individuals of colour 1 at time t (either $\heartsuit$ or $\diamondsuit$).
- $Z_{D_2}(t)$ is the number of dormant individuals of colour 2 at time t. (either $\heartsuit$ or $\diamondsuit$).

Note that the number of active individuals at time t (either $\heartsuit$ or $\diamondsuit$) is given by $Z_A(t) = N - Z_{D_1}(t) - Z_{D_2}(t)$. Since the number of individuals N is constant during the evolution, $Z_A(t)$ can be derived from $Z_{D_1}(t)$ and $Z_{D_2}(t)$. Each active individual resamples at rate 1, and becomes dormant at rate ϵ. When an individual becomes dormant, it gets either colour 1 with probability p_1 or colour 2 with probability p_2, where $p_1, p_2 \in (0, 1)$ and $p_1 + p_2 = 1$. For ease of notation, we denote the rate to become dormant with colour 1 by $\epsilon_1 = \epsilon \cdot p_1$ and the rate to become dormant with colour 2 by $\epsilon_2 = \epsilon \cdot p_2$. A dormant individual with colour 1 becomes active at rate δ_1, a dormant individual with colour 2 becomes active at rate δ_2. Thus, the transition rates for $(X(t), Y_1(t), Y_2(t), Z_{D_1}(t), Z_{D_2}(t))$ are:

- $(i, j, k, l, m) \to (i + 1, j, k, l, m)$ at rate $(N - l - m - i)\frac{i}{N-l-m}$.
- $(i, j, k, l, m) \to (i - 1, j, k, l, m)$ at rate $i\frac{(N-l-m-i)}{N-l-m}$.
- $(i, j, k, l, m) \to (i - 1, j + 1, k, l + 1, m)$ at rate $\epsilon_1 i$.
- $(i, j, k, l, m) \to (i + 1, j - 1, k, l - 1, m)$ at rate $\delta_1 j$.
- $(i, j, k, l, m) \to (i - 1, j, k + 1, l, m + 1)$ at rate $\epsilon_2 i$.
- $(i, j, k, l, m) \to (i + 1, j, k - 1, l, m - 1)$ at rate $\delta_2 k$.
- $(i, j, k, l, m) \to (i, j, k, l + 1, m)$ at rate $\epsilon_1 (N - l - m - i)$.
- $(i, j, k, l, m) \to (i, j, k, l, m + 1)$ at rate $\epsilon_2 (N - l - m - i)$.
- $(i, j, k, l, m) \to (i, j, k, l - 1, m)$ at rate $\delta_1 (l - j)$.
- $(i, j, k, l, m) \to (i, j, k, l, m - 1)$ at rate $\delta_2 (m - k)$.

Proceeding in the same way as for the 1-colour seed-bank, we define the scaled variables

$$\bar{X}(t) = \frac{1}{N} X(Nt), \quad \bar{Y}_1(t) = \frac{1}{N} Y_1(Nt), \quad \bar{Y}_2(t) = \frac{1}{N} Y_2(Nt),$$

$$\bar{Z}_A(t) = \frac{1}{N} Z_A(Nt), \quad \bar{Z}_{D_1}(t) = \frac{1}{N} Z_{D_1}(Nt), \quad \bar{Z}_{D_2}(t) = \frac{1}{N} Z_{D_1}(Nt).$$

(B.14)

We assume that there exist $c_1^A, c_2^A, c_1^D, c_2^D \in (0, \infty)$ such that

$$N\epsilon_1 = c_1^A, \qquad N\epsilon_2 = c_2^A, \qquad N\delta_1 = c_1^D, \qquad N\delta_2 = c_2^D, \qquad N \in \mathbb{N}, \tag{B.15}$$

and further assume for the starting values that

$$\lim_{N\to\infty} \frac{i}{N} = x, \quad \lim_{N\to\infty} \frac{j}{N} = y_1, \quad \lim_{N\to\infty} \frac{k}{N} = y_2,$$

$$\lim_{N\to\infty} \frac{N-l-m}{N} = z_A \quad \lim_{N\to\infty} \frac{N-l-m}{N} = z_{D_2}, \quad \lim_{N\to\infty} \frac{N-l-m}{N} = z_{D_1}.$$

(B.16)

Using the same method of converging generators as for Model 1, we obtain the following continuum limit:

$$
\begin{aligned}
dx(t) &= \sqrt{z_A(t) \frac{z_A - x(t)}{z_A(t)} \frac{x(t)}{z_A(t)}} \, dw(t) \\
&\quad + \left[c_1^D y_1(t) - c_1^A x(t) \right] dt + \left[c_2^D y_2(t) - c_2^A x(t) \right] dt, \\
dy_1(t) &= \left[c_1^A x(t) - c_1^D y_1(t) \right] dt, \\
dy_2(t) &= \left[c_2^A x(t) - c_2^D y_2(t) \right] dt, \\
dz_A(t) &= \left[c_1^D z_{D_1}(t) - c_1^A z_A(t) + c_2^D z_{D_2}(t) - c_2^A z_A(t) \right] dt, \\
dz_{D_1}(t) &= \left[c_1^A z_A(t) - c_1^D z_{D_1}(t) \right] dt, \\
dz_{D_2}(t) &= \left[c_2^A z_A(t) - c_2^D z_{D_2}(t) \right] dt.
\end{aligned}
\tag{B.17}
$$

Note that the equation for $z_A(t) = 1 - z_{D_1}(t) - z_{D_2}(t)$ follows directly from the equations from $z_{D_1}(t)$ and $z_{D_2}(t)$. It is therefore redundant, but we use it for notational reasons. Again, we see that $z(t) = (z_A(t), z_{D_1}(t), z_{D_2}(t))$ is governed by an autonomous system of differential equations. Solving this system, we see that

$$\lim_{t\to\infty} z_A(t) = \frac{1}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \qquad a.s.,$$

$$\lim_{t\to\infty} z_{D_1}(t) = \frac{\frac{c_1^A}{c_1^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \qquad a.s.,$$

$$\lim_{t\to\infty} z_{D_2}(t)) = \frac{\frac{c_2^A}{c_2^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \qquad a.s.$$

(B.18)

To compare the Moran model with a 2-colour seed-bank with the Fisher-Wright model

with a 2-colour seed-bank, we look at the variables

$$\bar{x}(t) = \left(1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}\right) x\left(\frac{t}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}\right),$$

$$\bar{y}_1(t) = \left(1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}\right)\left(\frac{c_1^D}{c_1^A}\right) y_1\left(\frac{t}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}\right), \tag{B.19}$$

$$\bar{y}_2(t) = \left(1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}\right)\left(\frac{c_2^D}{c_2^A}\right) y_2\left(\frac{t}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}\right).$$

Defining

$$K_m = \frac{c_m^A}{c_m^D}, \quad e_m = \frac{c_m^D}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}, \qquad m \in \{1, 2\}, \tag{B.20}$$

we see that, after a short transition period, the system becomes

$$\begin{aligned} \mathrm{d}\bar{x}(t) &= \sqrt{\bar{x}(t)(1 - \bar{x}(t))}\,\mathrm{d}w(t) + K_1 e_1\left[\bar{y}_2(t) - \bar{x}(t)\right]\mathrm{d}t + K_2 e_2\left[\bar{y}_1(t) - \bar{x}(t)\right]\mathrm{d}t, \\ \mathrm{d}\bar{y}_1(t) &= e_1\left[\bar{x}(t) - \bar{y}_1(t)\right]\mathrm{d}t, \\ \mathrm{d}\bar{y}_2(t) &= e_2\left[\bar{x}(t) - \bar{y}_2(t)\right]\mathrm{d}t, \end{aligned} \tag{B.21}$$

which is the single-colony version of (2.12)–(2.13) with 2 colours and without migration. Note, in particular, that after $z(t)$ reaches the equilibrium point in (B.18), we have

$$K_m = \frac{\text{number of dormant individuals with colour } m}{\text{number of active individuals}}, \quad m \in \{1, 2\}. \tag{B.22}$$

It is instructive to show how the above result can also be derived with the help of *duality*. The argument that follows easily extends to an n-coloured seed-bank for any $n \in \mathbb{N}$ finite, to be considered in Appendix B.3. Recall from (B.17) that

$$\begin{aligned} \mathrm{d}z_A(t) &= \left[c_1^D z_{D_1}(t) - c_1^A z_A(t) + c_2^D z_{D_2}(t) - c_2^A z_A(t)\right]\mathrm{d}t, \\ \mathrm{d}z_{D_1}(t) &= \left[c_1^A z_A(t) - c_1^D z_{D_1}(t)\right]\mathrm{d}t, \\ \mathrm{d}z_{D_2}(t) &= \left[c_2^A z_A(t) - c_2^D z_{D_2}(t)\right]\mathrm{d}t. \end{aligned} \tag{B.23}$$

Let

$$\bar{z}_A(t) = \left(1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}\right) z_A(t),$$

$$\bar{z}_{D_1}(t) = \left(1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}\right)\left(\frac{c_1^D}{c_1^A}\right) z_{D_1}(t), \tag{B.24}$$

$$\bar{z}_{D_2}(t) = \left(1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}\right)\left(\frac{c_2^D}{c_2^A}\right) z_{D_2}(t).$$

Substitute (B.24) into (B.23), to obtain

$$\begin{aligned} \mathrm{d}\bar{z}_A(t) &= c_1^A\left[\bar{z}_{D_1}(t) - \bar{z}_A(t)\right] + c_2^A\left[\bar{z}_{D_2}(t) - \bar{z}_A(t)\right]\mathrm{d}t, \\ \mathrm{d}\bar{z}_{D_1}(t) &= c_1^D\left[\bar{z}_A(t) - \bar{z}_{D_1}(t)\right]\mathrm{d}t, \\ \mathrm{d}\bar{z}_{D_2}(t) &= c_2^D\left[\bar{z}_A(t) - \bar{z}_{D_2}(t)\right]\mathrm{d}t. \end{aligned} \tag{B.25}$$

To define a dual for the process $(\bar{z}_A(t), \bar{z}_{D_1}(t), \bar{z}_{D_2}(t)))_{t \geq 0}$, let $(M(t))_{t \geq 0}$ be the continuous-time Markov chain on $\{A, D_1, D_2\}$ with transition rates

$$\begin{aligned} A &\to D_m \text{ at rate } c_m^A, \quad m \in \{1, 2\}, \\ D_m &\to A \text{ at rate } c_m^D, \quad m \in \{1, 2\}. \end{aligned} \tag{B.26}$$

Consider l independent copies of $(M(t))_{t \geq 0}$, evolving on the same state space $\{A, D_1, D_2\}$. Let $(L(t))_{t \geq 0} = (L_A(t), L_{D_1}(t), L_{D_2}(t))_{t \geq 0}$ be the process that counts how many copies of $M(t)$ are on site $\{A\}$, $\{D_1\}$ and $\{D_2\}$ at time t. Let $l = m + n_1 + n_2$. Then $(L(t))_{t \geq 0}$ is the Markov process on $\mathbb{N}_0^3$ with transition rates

$$(m, n_1, n_2) \to \begin{cases} (m-1, n_1+1, n_2) & \text{at rate } mc_1^A, \\ (m-1, n_1, n_2+1) & \text{at rate } mc_2^A, \\ (m+1, n_1-1, n_2) & \text{at rate } n_1 c_1^D, \\ (m+1, n_1, n_2-1) & \text{at rate } n_2 c_2^D. \end{cases} \tag{B.27}$$

Note that $L_A(t) + L_{D_1}(t) + L_{D_2}(t) = L_A(0) + L_{D_1}(0) + L_{D_2}(0) = m + n_1 + n_2 = l$. Define $H \colon \mathbb{R}^3 \times \mathbb{N}_0^3 \to \mathbb{R}$ by

$$H((\bar{z}_A, \bar{z}_{D_1}, \bar{z}_{D_2}), (m, n_1, n_2)) := \bar{z}_A^m \bar{z}_{D_1}^{n_1} \bar{z}_{D_2}^{n_2} \tag{B.28}$$

Using the generator criterion [JK14, Proposition 1.2], we see that, for all $t \geq 0$,

$$\begin{aligned} &\mathbb{E}\left[H((\bar{z}_A(t), \bar{z}_{D_1}(t), \bar{z}_{D_2}(t)), (m(0), n_1(0), n_2(0)))\right] \\ &= \mathbb{E}\left[H((\bar{z}_A(0), \bar{z}_{D_1}(0), \bar{z}_{D_2}(0)), (m(t), n_1(t), n_2(t)))\right]. \end{aligned} \tag{B.29}$$

Therefore $(L(t))_{t \geq 0}$ and $(\bar{z}(t))_{t \geq 0}$ are dual to each other with duality function H.

Since $(M(t))_{t \geq 0}$ is a irreducible and recurrent, we can define

$$\begin{aligned} \pi_A &= \lim_{t \to \infty} \mathbb{P}(M(t) = A) = \frac{1}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}, \\ \pi_{D_1} &= \lim_{t \to \infty} \mathbb{P}(M(t) = D_1) = \frac{\frac{c_1^A}{c_1^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}, \\ \pi_{D_2} &= \lim_{t \to \infty} \mathbb{P}(M(t) = D_2) = \frac{\frac{c_2^A}{c_2^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}}. \end{aligned} \tag{B.30}$$

Using the duality relation in (B.29) together with (B.30) and (B.24), we find

$$\begin{aligned} \lim_{t \to \infty} \mathbb{E}[\bar{z}_A(t)] &= \pi_A \bar{z}_A(0) + \pi_{D_1} \bar{z}_{D_1}(0) + \pi_{D_2} \bar{z}_{D_2}(0) \\ &= \frac{1}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \bar{z}_A(0) + \frac{\frac{c_1^A}{c_1^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \bar{z}_{D_1}(0) + \frac{\frac{c_2^A}{c_2^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \bar{z}_{D_2}(0) \\ &= z_A(0) + z_{D_1}(0) + z_{D_2}(0) = 1. \end{aligned} \tag{B.31}$$

Using the duality relation in (B.29) once more, we get

$$\lim_{t \to \infty} \mathbb{E}[\bar{z}_A(t)] = \lim_{t \to \infty} \mathbb{E}[\bar{z}_{D_1}(t)] = \lim_{t \to \infty} \mathbb{E}[\bar{z}_{D_2}(t)] = 1. \tag{B.32}$$

Computing the limiting second moment $\lim_{t \to \infty} \mathbb{E}[\bar{z}_A(t)^2]$ by duality, we obtain

$$\begin{aligned} \lim_{t \to \infty} \mathbb{E}[\bar{z}_A(t)^2] &= \lim_{t \to \infty} \sum_{\substack{i,j \in \\ \{A, D_1, D_2\}}} \mathbb{P}(M_t^1 = i) \, \bar{z}_i(0) \, \mathbb{P}(M_t^2 = j) \, \bar{z}_j(0) \\ &= \sum_{i \in \{A, D_1, D_2\}} \pi_i \bar{z}_i(0) \sum_{j \in \{A, D_1, D_2\}} \pi_j \bar{z}_j(0) = 1. \end{aligned} \tag{B.33}$$

Similarly, we find $\lim_{t \to \infty} \mathbb{E}[\bar{z}_{D_1}(t)^2] = 1$ and $\lim_{t \to \infty} \mathbb{E}[\bar{z}_{D_2}(t)^2] = 1$. Combining (B.32) and (B.33),

$$\lim_{t \to \infty} \bar{z}_A(t) = \lim_{t \to \infty} \bar{z}_{D_1}(t) = \lim_{t \to \infty} \bar{z}_{D_2}(t) = 1 \qquad a.s. \tag{B.34}$$

Hence we conclude that

$$\lim_{t \to \infty} z_A(t) = \frac{1}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \quad a.s.,$$

$$\lim_{t \to \infty} z_{D_1}(t) = \frac{\frac{c_1^A}{c_1^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \quad a.s., \tag{B.35}$$

$$\lim_{t \to \infty} z_{D_2}(t) = \frac{\frac{c_2^A}{c_2^D}}{1 + \frac{c_1^A}{c_1^D} + \frac{c_2^A}{c_2^D}} \qquad a.s.$$

Continuing as in (B.19), we again find the single-colony version of (2.12)-(2.13) with 2 colours and no migration.

B.3 Moran version of Model 2: Three or more colours

The argument in Appendix B.2 can be extended to an $\mathfrak{m} \in \mathbb{N}$-colour seed-bank, by introducing sequences of variables $(Y_m(t))_{m=0}^{\mathfrak{m}}$ and $(Z_m(t))_{m=0}^{\mathfrak{m}}$ that count the number of $\heartsuit$-individuals in the colour-m seed-bank at time t, respectively, the total number of individuals in the colour-m seed-bank at time t. Let $\epsilon > 0$ be the total rate at which an active individual becomes dormant, and define a probability vector $(p_m)_{m=0}^{\mathfrak{m}}$ such that $\epsilon_m = \epsilon p_m$ is the rate at which an active individual becomes dormant with colour m. Let δ_m be the rate at which m-dormant individuals become active. Via the same line of argument as in Appendix B.2, we see that the equivalent of (B.17) reads

$$\mathrm{d}x(t) = \sqrt{z_A(t) \frac{z_A - x(t)}{z_A(t)} \frac{x(t)}{z_A(t)}} \, \mathrm{d}w(t) + \sum_{m=0}^{\mathfrak{m}} \left[c_m^D y_m(t) - c_m^A x(t) \right] \mathrm{d}t,$$

$$\mathrm{d}y_m(t) = \left[c_m^A x(t) - c_m^D y_m(t) \right] \mathrm{d}t,$$

$$\mathrm{d}z_A(t) = \sum_{m=0}^{\mathfrak{m}} \left[c_m^D z_{D_m}(t) - c_m^A z_A(t) \right] \mathrm{d}t, \tag{B.36}$$

$$\mathrm{d}z_{D_m}(t) = \left[c_m^A z_A(t) - c_m^D z_{D_m}(t) \right] \mathrm{d}t, \qquad 0 \le m \le \mathfrak{m}.$$

Solving the autonomous system describing $z(t) = (z_A(t), (z_{D_m}(t))_{m=0}^{N})$ via duality, and subsequently substituting into (B.36) the variables

$$\bar{x}(t) = \left(1 + \sum_{n=0}^{\mathfrak{m}} \frac{c_n^A}{c_n^D} \right) x \left(\frac{t}{1 + \sum_{n=0}^{\mathfrak{m}} \frac{c_n^A}{c_n^D}} \right),$$

$$\bar{y}_m(t) = \left(1 + \sum_{n=0}^{\mathfrak{m}} \frac{c_n^A}{c_n^D} \right) \left(\frac{c_m^D}{c_m^A} \right) y_m \left(\frac{t}{1 + \sum_{n=0}^{\mathfrak{m}} \frac{c_n^A}{c_n^D}} \right), \qquad 0 \le m \le \mathfrak{m}, \tag{B.37}$$

we find the single-colony version of (2.12)–(2.13) with N-colours and no migration. Migration can be added as in Appendix A.

It is straightforward to derive the version (2.12)–(2.13) with N-colours and M colonies. Afterwards we can let $N, M \to \infty$ and use convergence of generators, to find (2.12)–(2.13). The limit is unproblematic because we are interested in finite time horizons only.

C Successful coupling

To prove Lemma 5.11 we proceed as in [CG94], with minor adaptations. The notation used in this appendix is the same as in Section 5.3. For Model 1 we write down the full proof. The proof works for Model 2 and 3 by invoking the colours $m \in \mathbb{N}_0$ and the SSDE in (2.12)–(2.13), respectively, (2.18)–(2.19).

Proof of Lemma 5.11. The proof consists of 5 steps.

Step 1. If $z \in E$ with $x_i = 0$ and $x_k > 0$ for some $k \neq i$, then

$$\mathbb{P}_z \left(\exists\, t^* > 0 \text{ such that } x_i(t) = 0\ \forall\, t \in [0, t^*]\right) = 0. \tag{C.1}$$

Proof. Suppose that z is such that $x_i = 0$, but $x_k > 0$ for some $i, k \in \mathbb{G}$. By (2.4),

$$x_i(t) = \int_0^t \mathrm{d}s \sum_{j \in \mathbb{G}} a(i,j)\,[x_j(s) - x_i(s)] + \int_0^t \mathrm{d}s\, Ke\,[y_i(s) - x_i(s)] + \int_0^t \mathrm{d}w_i(s)\,\sqrt{g(x_i(s))}. \tag{C.2}$$

Suppose that there exists a $T > 0$ such that $x_i(t) = 0$ for all $t \in [0, T]$, and therefore $g(x_i(t)) = 0$. Then we obtain for all $t \in [0, T]$ that

$$\int_0^t \mathrm{d}s \sum_{j \in \mathbb{G}} a(i,j)x_j(s) + \int_0^t \mathrm{d}s\, Key_i(s) = 0. \tag{C.3}$$

Hence, by path continuity of $(Z(t))_{t \geq 0}$, we see that $y_i(t) = 0$ for all $t \in [0, T]$, as well as $x_j(t) = 0$ for all $j \in \mathbb{G}$ such that $a(i,j) > 0$. Repeating this argument, we obtain by irreducibility of $a(\cdot, \cdot)$ that $x_k(t) = 0$ for all $k \in \mathbb{G}$ and hence $y_k(t) = 0$ for all $k \in \mathbb{G}$. By path continuity, this contradicts the assumption that $x_k(0) > 0$. We conclude that (C.1) holds. $\qquad\square$

Step 2. If $\bar{z} \in E \times E$ and $g(x_i^1) \neq g(x_i^2)$, then for all j,

$$\hat{\mathbb{P}}_{\bar{z}} \left(\exists\, t^* > 0 \text{ such that } \Delta_j(t) = 0\ \forall\, t \in [0, t^*]\right) = 0. \tag{C.4}$$

Proof. Note that the SSDE in (2.4)–(2.5) can be rewritten as

$$\begin{aligned}
\mathrm{d}z_{(i,R_i)}(t) &= \sum_{(j,R_j) \in \mathbb{G} \times \{A,D\}} b^{(1)}((i,R_i),(j,R_j))\,[z_{(j,R_j)}(t) - z_{(i,R_i)}(t)]\,\mathrm{d}t \\
&\quad + \sqrt{g(z_{(i,R_i)}(t))}\,1_{\{R_i = A\}}\,\mathrm{d}w_i(t), \\
&\quad \forall\,(i,R_i) \in \mathbb{G} \times \{A,D\},
\end{aligned} \tag{C.5}$$

with $b^{(1)}(\cdot, \cdot)$ defined as in (2.34).

Suppose that $\bar{z}$ is such that $g(x_i^1) \neq g(x_i^2)$. Suppose there exist a $T > 0$ such that $\Delta_j(t) = 0$ for all $t \in [0, T]$. Then also $\sqrt{g(x_j^1(t))} - \sqrt{g(x_j^2(t))} = 0$ for all $t \in [0, T]$. Using (C.5) on $\Delta_j(t) = z_{(j,A)}^1(t) - z_{(j,A)}^2(t)$, we obtain

$$0 = \int_0^t \mathrm{d}s$$
$$\sum_{(k,R_k) \in \mathbb{G} \times \{A,D\}} b^{(1)}((j,A),(k,R_k)) \left[\left(z_{(k,R_k)}^1(s) - z_{(k,R_k)}^2(s)\right) - \left(z_{(j,R_j)}^1(s) - z_{(j,R_j)}^2(s)\right)\right]. \tag{C.6}$$

Hence

$$
\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \left[\left(z^1_{(k,R_k)}(t) - z^2_{(k,R_k)}(t) \right) - \left(z^1_{(j,R_j)}(t) - z^2_{(j,R_j)}(t) \right) \right]
$$
$$
= 0 \qquad \forall\, t \in [0,T].
\tag{C.7}
$$

Using (C.5), we can write the SDE for

$$
\sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \left[\left(z^1_{(j,R_j)}(t) - z^2_{(j,R_j)}(t) \right) - \left(z^1_{(i,R_i)}(t) - z^2_{(i,R_i)}(t) \right) \right],
\tag{C.8}
$$

which yields that, for all $t \in [0,T]$,

$$
-\int_0^t \mathrm{d}w_k(s) \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1)}((j,A),(k,R_k)) \left(\sqrt{g(z^1_{k,R_k}(s))} - \sqrt{g(z^2_{k,R_k}(s))} \right) 1_{\{R_k = A\}}
$$
$$
= \int_0^t \mathrm{d}s \sum_{(k,R_k)\in\mathbb{G}\times\{A,D\}} b^{(1),2}((j,A),(l,R_l))
$$
$$
\times \left[\left(z^1_{(j,R_j)}(t) - z^2_{(j,R_j)}(t) \right) - \left(z^1_{(i,R_i)}(t) - z^2_{(i,R_i)}(t) \right) \right],
\tag{C.9}
$$

where $b^{(1),2}(\cdot,\cdot)$ is the 2-step kernel of $b^{(1)}(\cdot,\cdot)$.

The two process in the right-hand side form a process of bounded variation, while the process in the left-hand side is a continuous square-integrable martingale, whose quadratic variation is given by

$$
\int_0^t \mathrm{d}s \sum_{k\in\mathbb{G}} a(j,k)^2 \left(\sqrt{g(x^1_k(s))} - \sqrt{g(x^2_k(s))} \right)^2.
\tag{C.10}
$$

Since a square-integrable martingale of bounded variation is constant, it follows that (C.10) equals 0. Hence, for all k such that $a(j,k) > 0$, it follows that $g(x^1_k(t)) = g(x^2_k(t))$ for all $t \in [0,T]$. Moreover, the right-hand side of (C.9) is equal to 0. Iterating the right-hand side of (C.9) further, we find by the irreducibility of $a(\cdot,\cdot)$ that $g(x^1_i(t)) = g(x^2_i(t))$ for all $t \in [0,T]$, which contradicts the assumption on $\bar{z}$ that $g(x^1_i(0)) \neq g(x^2_i(0))$. Hence we find that there does not exist a $T > 0$ such that $\Delta_j(t) = 0$ for all $t \in [0,T]$. $\qquad\square$

Step 3. If $\bar{z} \in E \times E$, $i,k \in \mathbb{G}$ and $g(x^1_i) = g(x^2_i)$ with $\Delta_i < 0$ and $\Delta_k > 0$ for some $k \neq i$, then

$$
\hat{\mathbb{P}}_{\bar{z}} \left(\exists\, t^* \in [0,\tfrac{1}{2}]\colon \Delta_i(t^*) < 0,\ \Delta_k(t^*) > 0,\ g(x^1_i(t^*)) \neq g(x^2_i(t^*)) \right) > 0.
\tag{C.11}
$$

Proof. Note that by assumption we have $x^1_i < 1$ and $x^1_k > 0$. Let $t_0 \in [0,\tfrac{1}{4}]$. If $x^1_i > 0$, then set $t_0 = 0$. Otherwise, by Step 1 and path continuity, we find with probability 1 a $t_0 \in [0,\tfrac{1}{4}]$ such that $x^1_i(t_0) > 0$, $\Delta_i(t_0) < 0$ and $\Delta_k(t_0) > 0$. Let $\tilde{z} = \bar{z}(t_0)$. By the existence of t_0 and the Markov property, it is enough to prove that

$$
\hat{\mathbb{P}}_{\tilde{z}} \left(\exists\, t^* \in [0,\tfrac{1}{4}]\colon \Delta_i(t^*) < 0,\ \Delta_k(t^*) > 0,\ g(x^1_i(t^*)) \neq g(x^2_i(t^*)) \right) > 0
\tag{C.12}
$$

in order to prove (C.11). Define the following two martingales:

$$
M_i(t) = \int_0^t \mathrm{d}w_i(s) \sqrt{g(x^1_i(s))},
\tag{C.13}
$$

$$
M_k(t) = \int_0^t \mathrm{d}w_k(s) \left(\sqrt{g(x^1_k(s))} - \sqrt{2g(x^2_k(s))} \right).
\tag{C.14}
$$

Their corresponding quadratic variation processes are given by

$$\langle M_i(t)\rangle = \int_0^t \mathrm{d}s\, g(x_i(s)), \tag{C.15}$$

$$\langle M_k(t)\rangle = \int_0^t \mathrm{d}s\left(\sqrt{g(x_k^1(s))} - \sqrt{2g(x_k^2(s))}\right)^2. \tag{C.16}$$

By Knight's theorem (see [RY99, Theorem V.1.9 p.183]), we can write $M_i(t)$ and $M_k(t)$ as time-transformed Brownian motions:

$$M_i(t) = w_i\left(\langle M_i(t)\rangle\right), \tag{C.17}$$

$$M_k(t) = w_k\left(\langle M_k(t)\rangle\right). \tag{C.18}$$

We may assume that $g(\tilde{x}_i^1) = g(\tilde{x}_i^2)$, otherwise we can set $t^* = 0$. Recall that $0 < \tilde{x}_i^1 < 1$, $\tilde{\Delta}_i < 0$ and $\tilde{\Delta}_k > 0$, and, since $0 < g(\tilde{x}_i^1) = g(\tilde{x}_i^2)$, also $\tilde{x}_i^2 < 1$. Choose an $\epsilon \in (0, \frac{1}{15})$ such that $\tilde{x}_i^1, \tilde{x}_i^2 \in [5\epsilon, 1-5\epsilon]$, $-\tilde{\Delta}_i > 5\epsilon$ and $\tilde{\Delta}_k > 5\epsilon$. Let $\xi \in (0, \epsilon)$ be such that $g(\xi) < \min\{g(u)\colon \epsilon \le u \le 1-\epsilon\}$, and set $c_1 = \min\{g(u)\colon \xi \le u \le 1-\xi\}$ and $c_2 = \|g\|$. Then we can make the following estimates:

$$\langle M_i(t)\rangle \le c_2 t \quad \langle M_k(t)\rangle \le c_2 t,\, t \ge 0, \tag{C.19}$$

$$\langle M_i(t)\rangle \ge c_1 t \quad \text{for } t \ge 0 \text{ such that } x_i(s) \in [\xi, 1-\xi]\, \forall\, s \in [0,t]. \tag{C.20}$$

Define $c_3 = \min\{\frac{\xi}{2Ke}, \frac{\xi}{2}\}$. Fix $T \in [0, c_3]$ and define

$$\Omega_0 = \left\{\min_{t\in[0,c_1 T]} w_i(t) < -1,\ \max_{t\in[0,c_2 T]} w_i(t) < \epsilon,\ \max_{t\in[0,c_2 T]} |w_k(t)| < \epsilon\right\}, \tag{C.21}$$

$$\Omega_1 = \left\{\exists t^* \in [0,1] \text{ such that } \Delta_i(t^*) < 0,\ \Delta_k(t^*) > 0,\ g(x_i^1(t^*)) = g(x_i^2(t^*))\right\} \tag{C.22}$$

Note that $\mathbb{P}(\Omega_0) > 0$. Therefore it suffices that $\Omega_0 \subset \Omega_1$.

We start by checking the conditions Δ_k. Using (2.4), we can write

$$\Delta_k(t) = \Delta_k(0) + \int_0^t \mathrm{d}s \sum_{l\in\mathbb{G}} a(k,l)\left[\Delta_l(s) - \Delta_k(s)\right] + \int_0^t \mathrm{d}s\, Ke\left[\delta_k(s) - \Delta_k(s)\right]$$
$$+ \int_0^t \mathrm{d}w_k(s)\left(\sqrt{g(x_k^1(s))} - \sqrt{2g(x_k^2(s))}\right)^2. \tag{C.23}$$

Since $|\Delta_l(t)| \le 1$, $|\delta_k(t)| \le 1$ for all $t \ge 0$, and $M_k(t) = w_k(\langle M_k(t)\rangle)$ for $t \in [0,T]$, we may estimate

$$\Delta_k(t) > 5\epsilon - 2c_3 - 2Kec_3 - \epsilon = 2\epsilon. \tag{C.24}$$

So, on Ω_0, $\Delta_k(t) > 0$ for all $t \in [0,T]$. By expanding $x_i^1(t)$, we find

$$x_i^1(t) = x_i^1(0) + \int_0^t \mathrm{d}s \sum_{l\in\mathbb{G}} a(i,l)\left[x_l^1(s) - x_i^1(s)\right] + \int_0^t \mathrm{d}s\, Ke\left[y_i^1(s) - x_i^1(s)\right] + M_i(t), \tag{C.25}$$

so that on Ω_0 we have, for $t \in [0,T]$,

$$x_i^1(t) < 1 - 10\epsilon + c_3 + Kec_3 + \epsilon = 1 - 8\epsilon. \tag{C.26}$$

To check the conditions on $x_i^1(t)$ and $\Delta_i(t)$, we define the following random times:

$$\sigma = \inf\{t \ge 0\colon x_i^1(t) = \xi\},$$
$$\tau = \inf\left\{t > 0\colon g(x_i^1(t)) \ne g(x_i^2(t))\right\}. \tag{C.27}$$

We will prove that, on Ω_0, we have $\sigma < \tau$ and $x_i^2(\tau) \geq x_i^1(\tau) + 3\epsilon$. To do so, we first prove that $\sigma < T$. Assume the contrary $\sigma \geq T$. Then by (C.26) we have $x_i^1(t) \in [\xi, 1 - \xi]$ for all $t \in [0, T]$, which implies that $\min_{[0,T]} M_i(t) < -1$. Hence there exists a κ such that, by (C.25),

$$x_i^1(\kappa) < 1 - 10\epsilon + \epsilon - 1 < 0. \qquad \text{(C.28)}$$

However, this contradicts the fact that $x_i^1 > 0$ for all $t \geq 0$. We conclude that $\sigma < T$. Now suppose that $\tau > \sigma$. Expanding Δ_i, we get, for $t < \tau$,

$$\Delta_i(t) = \Delta_i(0) + \int_0^t \mathrm{d}s \sum_{l \in \mathbb{G}} a(i, l)\left[\Delta_l(s) - \Delta_i(s)\right] + \int_0^t \mathrm{d}s\, Ke\left[\delta_i(s) - \Delta_i(s)\right], \qquad \text{(C.29)}$$

which can be rewritten as

$$x_i^2(t) = x_i^1(t) - x_i^1(0) + x_i^2(0) - \int_0^t \mathrm{d}s \sum_{l \in \mathbb{G}} a(i, l)[\Delta_l(s) - \Delta_i(s)] - \int_0^t \mathrm{d}s\, Ke[\delta_i(s) - \Delta_i(s)]. \qquad \text{(C.30)}$$

By (C.30), we obtain, for $t \in [0, \sigma]$,

$$\begin{aligned} x_i^2(t) &\leq 1 - 5\epsilon + 2\epsilon + 2\epsilon = 1 - \epsilon, \\ x_i^2(t) &\geq x_i^1(t) + 5\epsilon - 2\epsilon \geq 3\epsilon, \end{aligned} \qquad \text{(C.31)}$$

so $x_i^2(t) \in [\epsilon, 1 - \epsilon]$ for $t \in [0, \sigma]$. But then $g(x_i^1(\sigma)) = g(\xi) < g(x_i^2(t))$ by the definition of ξ. Hence we obtain a contradiction and conclude that $\tau \leq \sigma$. From (C.31) we obtain that $\Delta_i(t) < 0$ for all $t \in [0, \tau]$, which concludes the proof that $\Omega_0 \subset \Omega_1$. $\qquad \square$

Step 4. If $\bar{z} \in E \times E$ and $\Delta_i < 0, \Delta_j = 0, \Delta_k > 0$ for some i, j, k, then

$$\hat{\mathbb{P}}_{\bar{z}}\left(\exists\, t^* \in [0, 1]\colon \Delta_i(t^*) < 0, \Delta_j(t^*) \neq 0, \Delta_k(t^*) > 0\right) > 0. \qquad \text{(C.32)}$$

Proof. Suppose that $\bar{z}$ satisfies $\Delta_i < 0$, $\Delta_j = 0$, $\Delta_k > 0$. Define

$$\begin{aligned} \Gamma_0 &= \{\bar{z} \in E \times E : \Delta_i < 0, \Delta_j \neq 0, \Delta_k > 0\}, \\ \Gamma_1 &= \{\bar{z} \in E \times E : \Delta_i < 0,\, g(x_i^1) \neq g(x_i^2),\, \Delta_k > 0\}. \end{aligned} \qquad \text{(C.33)}$$

By Step 3 and path continuity, there exists a $T \in [0, \frac{1}{2}]$ such that $\mathbb{P}^{\bar{z}}\left(\bar{z}(T) \in \Gamma_1\right) > 0$. By the Markov property,

$$\hat{\mathbb{P}}_{\bar{z}}\left(\exists\, t^* \in [0, 1]\colon \bar{z}(t^*) \in \Gamma_0\right) \geq \int_{\Gamma_1} \hat{\mathbb{P}}_{\bar{z}}\left(\bar{z}(T) \in d\tilde{z}\right) \hat{\mathbb{P}}_{\tilde{z}}\left(\exists\, t^* \in [0, \tfrac{1}{2}]\colon \bar{z}(t^*) \in \Gamma_0\right). \qquad \text{(C.34)}$$

By path continuity, we can find for $\bar{z} \in \Gamma_1$ a t' such that, for all $t \leq t'$, $\Delta_i(t) < 0$, $\Delta_k(t) > 0$ and $g(x_i^1(t)) \neq g(x_i^2(t))$. By Step 2 there exists a $t^* < t'$ such that $\bar{z}(t^*) \in \Gamma_0$. Hence both probabilities in the integral on the right-hand side of (C.34) are positive. $\qquad \square$

Step 5. Proof of Lemma 5.11.

Proof. Suppose that (5.106) holds for the pair i, j, and $a(j, k) > 0$, but (5.106) fails for the pair i, k. This implies that there exist $\epsilon_0 > 0$, $\delta_0 > 0$ and a positive increasing sequence $(t_n)_{n \in \mathbb{N}}$ of times with $t_n \to \infty$, such that

$$\lim_{t \to \infty} \hat{\mathbb{P}}_{\bar{z}}\left(\{\Delta_i(t) < \epsilon_0, \Delta_k(t) > \epsilon_0\} \cup \{\Delta_i(t) > \epsilon_0, \Delta_k(t) < \epsilon_0\}\right) > \delta_0. \qquad \text{(C.35)}$$

By compactness of $E \times E$, there exists a subsequence t_{n_k} such that $\mathcal{L}(\bar{z}(t_{n_k}))$ converges and (C.35) holds. Let $\bar{\nu} = \lim_{k \to \infty} \mathcal{L}(\bar{z}(t_{n_k}))$. Then

$$
\begin{aligned}
&\bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_j > \epsilon_0\} \cup \{\Delta_i > \epsilon_0, \Delta_j < \epsilon_0\}\right) = 0, \\
&\bar{\nu}\left(\{\Delta_j < \epsilon_0, \Delta_k > \epsilon_0\} \cup \{\Delta_j > \epsilon_0, \Delta_k < \epsilon_0\}\right) = 0, \\
&\bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_k > \epsilon_0\} \cup \{\Delta_i > \epsilon_0, \Delta_k < \epsilon_0\}\right) > \delta_0.
\end{aligned}
\tag{C.36}
$$

Assume without loss of generality that $\bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_k > \epsilon_0\}\right) > 0$. Hence, by (C.36),

$$
\bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_k > \epsilon_0\}\right) = \bar{\nu}\left(\{\Delta_i < \epsilon_0, \Delta_j \in (-\epsilon_0, \epsilon_0), \Delta_k > \epsilon_0\}\right) > 0.
\tag{C.37}
$$

For each $\bar{z} \in \{\Delta_i < \epsilon_0, \Delta_j \in (-\epsilon_0, \epsilon_0), \Delta_k > \epsilon_0\}$, Step 4 implies that

$$
\hat{\mathbb{P}}_{\bar{z}}\left(\exists\, t^* \in [0,1]\colon\ \Delta_i(t^*) < 0, \Delta_j(t^*) \neq 0, \Delta_k(t^*) > 0\right) > 0,
\tag{C.38}
$$

and therefore, by (C.37),

$$
\hat{\mathbb{P}}_{\bar{\nu}}\left(\exists\, t^* \in [0,1]\colon\ \Delta_i(t^*) < 0, \Delta_j(t^*) \neq 0, \Delta_k(t^*) > 0\right) > 0.
\tag{C.39}
$$

By path continuity, we can find $T \in [0,1]$ and $\epsilon > 0$ such that

$$
\hat{\mathbb{P}}_{\bar{\nu}}\left(\Delta_i(T) < -\epsilon,\ |\Delta_j(T)|,\ \Delta_k(T) > \epsilon\right) > 0.
\tag{C.40}
$$

Let $\bar{\mu}(t_n) = \mathcal{L}(\bar{z}(t_n))$. Then, by the Markov property and (C.40),

$$
\begin{aligned}
&\liminf_{n \to \infty} \hat{\mathbb{P}}_{\bar{\mu}(t_n)}\left(\Delta_i(T) < -\epsilon,\ |\Delta_j(T)| > \epsilon,\ \Delta_k(T) > \epsilon\right) \\
&= \liminf_{n \to \infty} \hat{\mathbb{P}}_{\bar{\mu}(0)}\left(\Delta_i(T + t_n) < -\epsilon,\ |\Delta_j(T + t_n)| > \epsilon,\ \Delta_k(T + t_n) > \epsilon\right) > 0.
\end{aligned}
\tag{C.41}
$$

However, this violates (5.106) for either i, j or j, k. We conclude that (C.35) fails and that (C.35) holds for i, k. By irreducibility, (C.35) holds for all $i, k \in \mathbb{G}$. $\qquad \square$

D Bounded derivative of Lyapunov function

Recall from Section 5.3 that

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

In this section we show that $h'(t)$ exists for all $t > 0$ and is bounded. To do so, we need to get rid of the indicator in the expectations.

Let

$$
h_{1,j}(t) = \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\Delta_j(t)\}}\right]
\tag{D.2}
$$

and

$$
h_2(t) = 2Ke\, \hat{\mathbb{E}}\left[\big(|\Delta_i(t)| + |\delta_i(t)|\big)\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\delta_i(t)\}}\right].
\tag{D.3}
$$

Then $h(t) = 2\sum_{j \in \mathbb{G}} a(i,j)h_{1,j}(t) + h_2(t)$. We show that $h_{1,j}(t)$ is differentiable with bounded derivative for $j \in \mathbb{G}$. The proof of the differentiability of $h_2(t)$ is similar. Fix $t \geq 0$. Note that

$$
\begin{aligned}
&\hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\Delta_j(t)\}}\right] \\
&= \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\Delta_j(t)\}} \mid |\Delta_i(t)| \neq 0, |\Delta_i(t)| \neq 0\right] \mathbb{P}\left(|\Delta_i(t)| \neq 0, |\Delta_j(t)| \neq 0\right) \\
&\quad + \hat{\mathbb{E}}\left[|\Delta_j(t)|\, 1_{\{\operatorname{sgn}\Delta_i(t) \neq \operatorname{sgn}\Delta_j(t)\}} \mid |\Delta_i(t)| = 0 \text{ or } |\Delta_j(t)| = 0\right] \mathbb{P}\left(|\Delta_i(t)| = 0 \text{ or } |\Delta_j(t)| = 0\right).
\end{aligned}
\tag{D.4}
$$

Since $\Delta_i(t)$ and $\Delta_j(t)$ have zero local time, the second term vanishes and $\mathbb{P}(|\Delta_i(t)| \neq 0, |\Delta_j(t)| \neq 0) = 1$. By continuity of $\Delta_i(t)$ and $\Delta_j(t)$, we can define sets

$$B_n = \{|\Delta_i(r)| > 0 \text{ and } |\Delta_j(r)| > 0, \forall r \in \mathcal{B}(t, \tfrac{1}{n})\}. \tag{D.5}$$

Then

$$\cdots \subset B_n \subset B_{n+1} \subset B_{n+2} \subset \cdots, \tag{D.6}$$

so

$$B_n = \bigcup_{i=0}^{n} B_i \tag{D.7}$$

and we define

$$B := \bigcup_{i=0}^{\infty} B_n = \lim_{n \to \infty} B_n. \tag{D.8}$$

Since $\mathbb{P}(|\Delta_i(t)| \neq 0, |\Delta_j(t)| \neq 0) = 1$, it follows that $\mathbb{P}(B) = 1$.

For each B_n, we have

$$B_n = C_n \cup C_n^c, \qquad C_n = \big\{\omega \in B_n \colon 1_{\{\text{sgn}\,\Delta_i(r) \neq \text{sgn}\,\Delta_j(r)\}} = 1, \ \forall r \in \mathcal{B}(t, \tfrac{1}{n})\big\}, \tag{D.9}$$

and, by the definition of B_n,

$$\cdots \subset C_n \subset C_{n+1} \subset C_{n+2} \subset \cdots \qquad \cdots \subset C_n^c \subset C_{n+1}^c \subset C_{n+2}^c \subset \cdots \tag{D.10}$$

Let $C = \bigcup_{i=0}^{\infty} C_i$ and $C^c = \bigcup_{i=0}^{\infty} C_i^c$ be such that $B = C \cup C^c$. Using (5.95), we obtain

$$\frac{1}{s}\left(h_{1,j}(t+s) - h_{1,j}(t)\right)$$

$$= \frac{1}{s}\left[\hat{\mathbb{E}}\left[|\Delta_j(t)|\,1_{\{\text{sgn}\,\Delta_i(t+s) \neq \text{sgn}\,\Delta_j(t+s)\}}\right] - \hat{\mathbb{E}}\left[|\Delta_j(t)|\,1_{\{\text{sgn}\,\Delta_i(t) \neq \text{sgn}\,\Delta_j(t)\}}\right]\right]$$

$$= \frac{1}{s}\left[\hat{\mathbb{E}}\left[|\Delta_j(t+s)|\,1_{\{\text{sgn}\,\Delta_i(t+s) \neq \text{sgn}\,\Delta_j(t+s)\}} - |\Delta_j(t)|\,1_{\{\text{sgn}\,\Delta_i(t) \neq \text{sgn}\,\Delta_j(t)\}} \mid B\right]\right]$$

$$= \frac{1}{s}\left[\hat{\mathbb{E}}\left[|\Delta_j(t+s)| - |\Delta_j(t)| \mid C\right]\right]\mathbb{P}(C)$$

$$= \frac{1}{s}\hat{\mathbb{E}}\left[\sum_{j \in \mathbb{G}} a(i,j)\int_t^{t+s} \mathrm{d}r\,\text{sgn}\,(\Delta_i(r))[\Delta_j(r) - \Delta_i(r)] \mid C\right]\mathbb{P}(C)$$

$$\quad + \frac{1}{s}\hat{\mathbb{E}}\left[\int_t^{t+s} \mathrm{d}w_i(r)\,\text{sgn}\,(\Delta_i(r))\left[\sqrt{g(x_i^1(r))} - \sqrt{g(x_i^2(r))}\right] \mid C\right]\mathbb{P}(C) \tag{D.11}$$

$$\quad + \frac{1}{s}\hat{\mathbb{E}}\left[Ke\int_t^{t+s} \mathrm{d}r\,\text{sgn}\,(\Delta_i(r))[\delta_i(r) - \Delta_i(r)] \mid C\right]\mathbb{P}(C)$$

$$= \sum_{j \in \mathbb{G}} a(i,j)\hat{\mathbb{E}}\left[\frac{1}{s}\int_t^{t+s} \mathrm{d}r\,\text{sgn}\,(\Delta_i(r))[\Delta_j(r) - \Delta_i(r)] \mid C\right]\mathbb{P}(C)$$

$$\quad + \frac{1}{s}\hat{\mathbb{E}}\left[\int_t^{t+s} \mathrm{d}w_i(r)\,\text{sgn}\,(\Delta_i(r))\left[\sqrt{g(x_i^1(r))} - \sqrt{g(x_i^2(r))}\right] \mid C\right]\mathbb{P}(C)$$

$$\quad + \hat{\mathbb{E}}\left[Ke\,\frac{1}{s}\int_t^{t+s} \mathrm{d}r\,\text{sgn}\,(\Delta_i(r))[\delta_i(r) - \Delta_i(r)] \mid C\right]\mathbb{P}(C).$$

In the last equality, the first and third term are bounded, because $\Delta_i(t), \delta_i(t)$ and $\Delta_j(t)$ are continuous functions of t, and $\text{sgn}\,(\Delta_i)$ is constant since we conditioned on the set C. Therefore, letting $s \to 0$, it follows from the fundamental theorem of calculus that these terms are bounded. The second term is more involved. Since, on the set C,

$$\text{sgn}\,(\Delta_i(r))\left[\sqrt{g(x_i^1(r))} - \sqrt{g(x_i^2(r))}\right] \tag{D.12}$$

is a continuous function, we can rewrite the stochastic integral as a time-transformed Brownian motion:

$$
\begin{aligned}
\frac{1}{s}\hat{\mathbb{E}}&\left[\int_t^{t+s}\mathrm{d}w_i(r)\,\mathrm{sgn}\left(\Delta_i(r)\right)\left[\sqrt{g(x_i^1(r))}-\sqrt{g(x_i^2(r))}\right]\mid C\right]\\
&=\frac{1}{s}\hat{\mathbb{E}}\left[W\left(\int_0^{t+s}\mathrm{d}r\left[\sqrt{g(x_i^1(r))}-\sqrt{g(x_i^2(r))}\right]^2\right)\right.\\
&\qquad\left.-W\left(\int_0^t\mathrm{d}r\left[\sqrt{g(x_i^1(r))}-\sqrt{g(x_i^2(r))}\right]^2\right)\mid C\right].
\end{aligned}
\tag{D.13}
$$

Since the normal distribution is differentiable with respect to its variance, we are done.

References

[AB16]   K. Alexander and Q. Berger. Local asymptotics for the first intersection of two independent renewals. *Electr. J. Probab.*, 21:1–20, 2016. MR3580034

[Asm03]   S. Asmussen. *Applied Probability and Queues*, volume 51 of *Applications of Mathematics*. Springer-Verlag, New York, USA, 2003. MR1978607

[BBGCWB19] J. Blath, E. Buzzoni, A. González Casanova, and M. Wilke-Berenguer. Structural properties of the seed bank and the two island diffusion. *J. Math. Biology*, 79:369–392, 2019. MR3975877

[BGCEK15] J. Blath, A. González Casanova, B. Eldon, and N. Kurt. Genealogy of a Wright-Fisher model with strong seed-bank component. *Birkhäuser Progress in Probability*, 69:81–100, 2015. MR3558137

[BGCKS13] J. Blath, A. González Casanova, N. Kurt, and D. Spano. The ancestral process of long-range seed bank models. *J. Appl. Prob.*, 50:741–759, 2013. MR3102512

[BGCKWB16] J. Blath, A. González Casanova, N. Kurt, and M. Wilke-Berenguer. A new coalescent for seed-bank models. *Ann. Appl. Probab.*, 26:857–891, 2016. MR3476627

[BGT87]   N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1987. MR1015093

[Bre68]   L. Breiman. *Probability*. Addison-Wesley, Reading, Massachusetts, 1968. MR0229267

[CG94]   J.T. Cox and A. Greven. Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.*, 22:833–853, 1994. MR1288134

[Daw93]   D.A. Dawson. *Measure-Valued Markov Processes*, volume 1541 of *Lecture Notes in Mathematics*. Springer-Verlag, New York, 1993. MR1242575

[DG93]   D.A. Dawson and A. Greven. Multiple scale analysis of interacting diffusions. *Probab. Theory Relat. Fields*, 95:467–508, 1993. MR1217447

[DG96]   D.A. Dawson and A. Greven. Multiple space-time scale analysis for interacting branching models. *Electron. J. Probab.*, 1:paper no. 14, 1–84, 1996. MR1423467

[DGdH+08] D.A. Dawson, A. Greven, F. den Hollander, R. Sun, and J.M. Swart. The renormalization transformation of two-type branching models. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44:1038–1077, 2008. MR2469334

[DGV95]   D.A. Dawson, A. Greven, and J. Vaillancourt. Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc.*, 347:2277–2360, 1995. MR1297523

[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. MR2106854

[DGW05]   D.A. Dawson, L. Gorostiza, and A. Wakolbinger. Degrees of transience and recurrence and hierarchical random walk. *Potential Anal.*, 22:305–350, 2005. MR2135263

[dHP17]   F. den Hollander and G. Pederzani. Multi-colony Wright-Fisher with seed-bank. *Indag. Math.*, 28:637–669, 2017. MR3651340

[EF96]   S.N. Evans and K. Fleischmann. Cluster formation in a stepping stone model with continuous, hierarchically structured sites. *Ann. Probab.*, 24:1926–1952, 1996. MR1415234

[EJU19]   N.M. Ercolani, J. Jansen, and D. Ueltschi. Singularity analysis for heavy-tailed random variables. *J. Theor. Probab.*, 32:1–46, 2019. MR3908904

[EK86]   S.N. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. John Wiley, New York, 1986. MR0838085

[Fel71]   W. Feller. *An Introduction to Probability Theory and Its Applications (2nd ed)*. John Wiley, New York, 1971. MR0270403

[FG96]   K. Fleischmann and A. Greven. Time-space analysis of the cluster formation in interacting diffusions. *Electr. J. Probab.*, 1:1–46, 1996. MR1386298

[Fol92]   G.B. Folland. *Fourier Analysis and its Applications*. Wadsworth & Brooks/Cole, Belmont, CA, USA, 1992. MR1145236

[GCAvWE$^+$14] A. González Casanova, E. Aguirre-von Wobeser, G. Espin, N. Servin-Gonzalez, L. Kurt, D. Spano, J. Blath, and G. Soberon-Chavez. Strong seed-bank effects in bacterial evolution. *J. Theor. Biol.*, 356:62–70, 2014.

[GdHKK14]   A. Greven, F. den Hollander, S. Kliem, and A. Klimovsky. Renormalisation of hierarchically interacting Cannings processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11:43–140, 2014. MR3194010

[GHO20a]   A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: cluster formation and genealogy. *in preparation*, 2022.

[GHO20b]   A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: finite-systems scheme. *Preprint*, 2022.

[GHO20c]   A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: renormalisation on the hierarchical group. *Preprint*, 2021.

[Gol79]   S. Goldstein. Maximal coupling. *Z. Wahrsch. Verw. Gebiete*, 46:193–204, 1978/79. MR0516740

[Hol06]   F. den Hollander. Renormalization of interacting diffusions: a program and four examples. In *Partial Differential Equations and Functional Analysis*, volume 168 of *Operator Theory, Advances and Applications*, pages 123–136. Birkhauser, Basel, 2006. MR2240057

[HP91]   I. Herbst and L. Pitt. Diffusion equation techniques in stochastic monotonicity and positive correlations. *Probab. Theory Relat. Fields*, 87:275–312, 1991. MR1084331

[HS98]   F. den Hollander and J. Swart. Renormalization of hierarchically interacting isotropic diffusions. *J. Stat. Phys.*, 93:243–291, 1998. MR1656371

[Hug95]   B.D. Hughes. *Random Walk in Random Environment*, volume I. Clarendon Press, 1995. MR1341369

[JK14]   S. Jansen and N. Kurt. On the notion(s) of duality for Markov processes. *Probab. Surv.*, 11:59–120, 2014. MR3201861

[KKL01]   I. Kaj, S. Krone, and M. Lascoux. Coalescent theory for seed bank models. *J. Appl. Prob.*, 38:285–300, 2001. MR1834743

[Lig85]   T.M. Liggett. *Interacting Particle Systems*. Springer, New York, 1985. MR0776231

[Lin92]   L. Lindvall. *Lectures on the Coupling Method*. Wiley, New York, 1992. MR1180522

[Lin99]   L. Lindenstrauss. Pointwise theorems for amenable groups. *Electron. Res. Announc. AMS*, 5:82–90, 1999. MR1696824

[LJ11]   J.T. Lennon and S.E. Jones. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9:119–130, 2011.

[Moj18]   T. Mojet. *Scaling Limits of Moran and Wright-Fisher Genetic Drift Models with Seedbank Components*. Universiteit Leiden, Bachelor thesis, 2018.

[NW81]   C.M. Newman and A.L. Wright. An invariance principle for certain dependent sequences. *Ann. Probab.*, 9:671–675, 1981. MR0624694

[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. MR0406577

[Rud62] W. Rudin. *Fourier Analysis on Groups*, volume 12 of *Interscience Tracts in Pure and Applied Mathematics*. Interscience Publishers, New York, 1962. MR0152834

[RW00]   L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge University Press, 2000. MR1780932

[RY99]   D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, New York, 1999. MR1725357

[Shi80] T. Shiga. An interacting system in population genetics. *J. Math. Kyoto Univ.*, 20:213–242, 1980. MR0582165

[Smi55] W.L. Smith. Regenerative stochastic processes. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 232:6–31, 1955. MR0073877

[Spi64] F. Spitzer. *Principles of Random Walk*, volume 34. D. Van Nostrand Company, Princeton, NJ, 1964. MR0171290

[SS80]   T. Shiga and A. Shimizu. Infinite-dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.*, 20:395–416, 1980. MR0591802