Spatially inhomogeneous populations with seed-banks: I. Duality, existence and clustering

Hollander, W.T.F. den; Nadan, S.

Journal of Theoretical Probability
https://doi.org/10.1007/s10959-021-01119-z

Spatially Inhomogeneous Populations with Seed-Banks: I. Duality, Existence and Clustering

Frank den Hollander[1] · **Shubhamoy Nandan**[1]

Received: 1 October 2020 / Revised: 7 July 2021 / Accepted: 12 July 2021

Abstract
We consider a system of interacting Moran models with seed-banks. Individuals live in colonies and are subject to resampling and migration as long as they are *active*. Each colony has a seed-bank into which individuals can retreat to become *dormant*, suspending their resampling and migration until they become active again. The colonies are labelled by $\mathbb{Z}^d$, $d \geq 1$, playing the role of a *geographic space*. The sizes of the active and the dormant population are *finite* and depend on the *location* of the colony. Migration is driven by a random walk transition kernel. Our goal is to study the equilibrium behaviour of the system as a function of the underlying model parameters. In the present paper, under a mild condition on the sizes of the active populations, the system is well defined and has a dual. The dual consists of a system of *interacting* coalescing random walks in an *inhomogeneous* environment that switch between an active state and a dormant state. We analyse the dichotomy of *coexistence* (= multi-type equilibria) versus *clustering* (= mono-type equilibria) and show that clustering occurs if and only if two random walks in the dual starting from arbitrary states eventually coalesce with probability one. The presence of the seed-bank *enhances genetic diversity*. In the dual this is reflected by the presence of time lapses during which the random walks are dormant and do not move.

Keywords Moran model · Resampling · Seed-bank · Migration · Duality · Coexistence versus clustering

The research in this paper was supported by the Netherlands Organisation for Scientific Research (NWO) through grant TOP1.17.019. The authors thank Simone Floreani, Cristian Gardinà and Frank Redig for various discussions on duality.

✉ Shubhamoy Nandan
s.nandan@math.leidenuniv.nl

Frank den Hollander
denholla@math.leidenuniv.nl

[1] Mathematisch Instituut, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

Published online: 27 July 2021

Mathematics Subject Classification (2020) Primary 60J70 · 60K35; Secondary 92D25

1 Background, Motivation and Outline

Dormancy is an evolutionary trait observed in plants, bacteria and other microbial populations, where an organism enters a reversible state of low metabolic activity as a response to adverse environmental conditions. The dormant state of an organism in a population is characterized by interruption of basic reproduction and phenotypic development during periods of environmental stress [24,29]. The dormant organisms reside in what is called a *seed-bank* of the population. After a varying and possibly large number of generations, dormant organisms can be resuscitated under more favourable conditions and reprise reproduction after becoming *active* by leaving the seed-bank. This strategy is known to have important implications for the genetic diversity and overall fitness of the underlying population [23,24], since the seed-bank of a population often acts as a *buffer* against evolutionary forces such as genetic drift, selection and environmental variability. The importance of dormancy has led to several attempts to model seed-banks from a mathematical perspective ([1,2]; see also [3] for a broad overview).

In [2] and [1], the Fisher–Wright model with *seed-bank* was introduced and analysed. In the Fisher–Wright model with seed-bank, individuals live in a colony, are subject to *resampling* where they adopt each other's type, and move in and out of the seed-bank where they suspend resampling. The seed-bank acts as a repository for the genetic information of the population. Individuals that reside inside the seed-bank are called *dormant* and those that reside outside are called *active*. Both the long-time behaviour and the genealogy of the population were analysed for the continuum model obtained by letting the size of the colony tend to infinity, called the Fisher–Wright diffusion with seed-bank.

In [13–15], the continuum model was extended to a *spatial* setting in which individuals live in multiple colonies, labelled by a countable Abelian group playing the role of a *geographic space*. In the spatial model with seed-banks, each colony is endowed with its own seed-bank and individuals are allowed to *migrate* between colonies. The goal was to understand the change in behaviour compared to the spatial model without seed-bank.

Most papers on seed-banks deal with the large-colony-size limit, for which the evolution is described by a system of coupled SDE's. In [19], a multi-colony Fisher–Wright model with seed-banks was introduced where the colony sizes are finite. However, this model is restricted to *homogeneous* population sizes and a finite geographic space. The present paper introduces an individual-based spatial model with seed-banks in continuous time where the sizes of the underlying populations are *finite* and *vary* across colonies. The latter make the model more interesting from a biological perspective, but raise extra technical challenges. The key tool that we use to tackle these challenges is *stochastic duality* [4,11]. The spatial model introduced in this paper fits in the realm of interacting particle systems, which often admit additional structures such as duality [25,28]. In particular, our spatial model can be viewed as a *hybrid* of the well-known Voter Model and the *generalized Symmetric Exclusion Process*, $2j$-SEP, $j \in \mathbb{N}/2$

[5,11,26]. Both the Voter Model and the $2j$-SEP enjoy the stochastic duality property, and our system inherits this as well: it is dual to a system consisting of *coalescing* random walks with *repulsive* interactions. The resulting dual process shares striking resemblances with the dual processes of the Voter Model and $2j$-SEP, because the original process is a modified hybrid of them. It has been recognized in the literature [1,2,23,24,31] that qualitatively different behaviour may occur when the exit time of a typical individual from the seed-bank can become large. In the present paper, we are able to model this phenomenon as well, due to the *inhomogeneity* in the seed-bank sizes. Our main goals are the following:

(1) Introduce a model with seed-banks whose size is *finite* and depends on the geographic location of the colony. Prove *existence* and *uniqueness* of the process via well-posedness of an associated martingale problem and duality with a system of interacting coalescing random walks.
(2) Identify a criterion for *coexistence* (= convergence towards multi-type equilibria) and *clustering* (= convergence towards mono-type equilibria). Show that there is a one-parameter family of equilibria controlled by the density of types.
(3) Identify the *domain of attraction* of the equilibria.
(4) Identify the *parameter regime* under which the criterion for clustering is met. In case of clustering, find out how fast the mono-type clusters grow in space-time. In case of coexistence, establish mixing properties of the equilibria.

In the present paper, we settle (1) and (2). In [18], we will address (3) and (4). 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.

The paper is organized as follows: In Sect. 2, we give a quick definition of the model and state our main theorems about the well-posedness, the duality and the clustering criterion. In Sect. 3, we give a more detailed definition of the model, prove that the martingale problem associated with its generator is well-posed, establish duality with an interacting seed-bank coalescent, demonstrate that the system exhibits a dichotomy between clustering and coexistence, and formulate a necessary and sufficient condition for clustering to prevail in terms of the dual, called the *clustering criterion*. Sections 4–6 are devoted to the proof of our main theorems.

2 Main Theorems

In Sect. 2.1, we give a quick definition of the system. In Sect. 2.2, we argue that under mild conditions on the sizes of the active population, the system is well defined and has a dual that consists of finitely many interacting coalescing random walks.

2.1 Quick Definition of the Multi-colony System

Individuals live in colonies labelled by $\mathbb{Z}^d$, $d \geq 1$, which plays the role of a *geographic space*. (In what follows, the geographic space can be any countable Abelian group.)

Each colony has an *active* population and a *dormant* population. Each individual carries one of two *types*: ♡ and ♠. Individuals are subject to:

(1) Active individuals in any colony *resample* with active individuals in *any* colony.
(2) Active individuals in any colony *exchange* with dormant individuals in the *same* colony.

For (1) we assume that each active individual at colony i at rate $a(i, j)$ uniformly draws an active individual at colony j and *adopts its type*. For (2) we assume that each active individual at colony i at rate λ uniformly draws a dormant individual at colony i and the two individuals *trade places while keeping their type* (i.e. the active individual becomes dormant and the dormant individual becomes active). Note that dormant individuals do *not* resample.

At each colony i, we register the pair $(X_i(t), Y_i(t))$, representing the number of active, respectively, dormant individuals of type ♡ at time t at colony i. We write (N_i, M_i) to denote the *size* of the active, respectively, dormant population at colony i. The resulting Markov process is denoted by

$$(Z(t))_{t\geq 0}, \qquad Z(t) = ((X_i(t), Y_i(t))_{i\in\mathbb{Z}^d}, \tag{2.1}$$

and lives on the state space

$$\mathcal{X} := \prod_{i\in\mathbb{Z}^d} [N_i] \times [M_i], \tag{2.2}$$

where $[n] = \{0, 1, \ldots, n\}$, $n \in \mathbb{N}$. In Sect. 3.2, we will show that under mild assumptions on the model parameters, the Markov process in (2.1) is well defined and has a *dual* $(Z^*(t))_{t\geq 0}$. The latter consists of finite collections of particles that perform *interacting coalescing* random walks, with rates that are controlled by the model parameters.

Let $\mathcal{P}$ be the set of probability distributions on $\mathcal{X}$ defined by

$$\mathcal{P} := \big\{\mathcal{P}_\theta :\ \theta \in [0, 1]\big\}, \qquad \mathcal{P}_\theta := \theta \bigotimes_{i\in\mathbb{Z}^d} \delta_{(0,0)} + (1-\theta) \bigotimes_{i\in\mathbb{Z}^d} \delta_{(N_i,M_i)}. \tag{2.3}$$

We say that (2.1) exhibits *clustering* if the distribution of $Z(t)$ converges to a limiting distribution $\mu \in \mathcal{P}$ as $t \to \infty$. Otherwise, we say that it exhibits *coexistence*. In Sect. 3.2, we will show that clustering is equivalent to *coalescence* occurring eventually with probability 1 in the dual consisting of *two* particles. This will be the main route to the dichotomy.

For simplicity, we let the exchange rate $\lambda \in (0, \infty)$ be the same for every colony, and let the *migration kernel* be translation invariant and irreducible.

Assumption 2.1 (Homogeneous migration) The migration kernel $a(\cdot, \cdot)$ satisfies:

- $a(\cdot, \cdot)$ is irreducible in $\mathbb{Z}^d$.
- $a(i, j) = a(0, j - i)$ for all $i, j \in \mathbb{Z}^d$.

- $c := \sum_{i \in \mathbb{Z}^d} a(0,i) < \infty$ and $a(0,0) = \frac{1}{2}$. □

The former of the last two assumptions ensures that the way genetic information moves between colonies is homogeneous in space, while the latter ensures that the total rate of resampling is finite and that resampling is possible also at the same colony. Since it is crucial for our analysis that the population sizes remain constant, we view migration as a change of types without the individuals actually moving themselves. In this way, genetic information moves between colonies, while the individuals themselves stay put.

We write

$$K_i = \frac{N_i}{M_i}, \qquad i \in \mathbb{Z}^d, \tag{2.4}$$

to denote the *ratio* of the size of the active and the dormant population in colony i.

2.2 Well-Posedness and Duality

Theorem 2.2 (Well-posedness and duality) *Suppose that Assumption 2.1 is in force. Then, the Markov process $(Z(t))_{t\geq 0}$ in (2.1) has a factorial moment dual $(Z^*(t))_{t\geq 0}$ living in the state space $\mathcal{X}^* \subset \mathcal{X}$ consisting of all configurations with finite mass, and the martingale problem associated with (2.1) is well posed under either of the two following conditions:*

(a) $\lim_{\|i\|\to\infty} \|i\|^{-1} \log N_i = 0$ *and* $\sum_{i\in\mathbb{Z}^d} \mathrm{e}^{\delta\|i\|} a(0,i) < \infty$ *for some* $\delta > 0$,
(b) $\sup_{i\in\mathbb{Z}^d\setminus\{0\}} \|i\|^{-\gamma} N_i < \infty$ *and* $\sum_{i\in\mathbb{Z}^d} \|i\|^{d+\gamma+\delta} a(0,i) < \infty$ *for* $\gamma > 0$ *and some* $\delta > 0$.

Theorem 2.2 provides us with two sufficient conditions under which the system is well-defined and has a tractable dual. It shows a *trade-off*: the more we restrict the tails of the migration kernel, the less we need to restrict the sizes of the active population. The sizes of the dormant population play no role because all the events (resampling, migration and exchange) in our model are initiated by active individuals and dormant individuals do not feel the spatial extent of the geographic space. Theorem 3.10, Corollary 3.11 and Theorem 3.13 in Sect. 3.2 contain the fine details.

2.3 Equilibrium: Coexistence Versus Clustering

Theorem 2.3 (Equilibrium) *If the initial distribution of the system is such that each active and each dormant individual adopts a type with the same probability independently of other individuals, then the system admits a one-parameter family of equilibria.*

- *The family of equilibria is parameterized by the probability to have one of the two types.*
- *The system converges to a mono-type equilibrium if and only if two random walks in the dual starting from arbitrary states eventually coalesce with probability one.*

Table 1 Scheme of transitions in the single-colony model

Initial state	Event	Final state	Transition rate
(x, y)	Resampling	$(x-1, y)$	$x(N-x)/N$
		$(x+1, y)$	$x(N-x)/N$
	Exchange	$(x-1, y+1)$	$\lambda x(M-y)/M$
		$(x+1, y-1)$	$\lambda(N-x)y/M$

Theorem 2.3 tells us that the system converges to an equilibrium when it is started from a specific class of initial distributions, namely products of binomials. It also provides a *criterion* in terms of the dual that determines whether the equilibrium is mono-type or multi-type. Theorem 3.14, Corollary 3.15 and Theorem 3.17 in Sect. 3.2 contain the fine details.

3 Basic Theorems: Duality, Well-Posedness and Clustering Criterion

In Sect. 3.1, we define and analyse the single-colony model. In Sect. 3.2, we do the same for the multi-colony model. Our focus is on well-posedness, duality and convergence to equilibrium.

3.1 Single-Colony Model

3.1.1 Definition: Resampling and Exchange

Consider two populations, called *active* and *dormant*, consisting of N and M haploid individuals, respectively. Individuals in the population carry one of two genetic types: ♡ and ♠. Dormant individuals reside inside the *seed-bank*, active individuals reside outside. The dynamics of the single-colony Moran model with seed-bank is as follows:

- Each individual in the active population carries a *resampling clock* that rings at rate 1. When the clock rings, the individual randomly chooses an active individual and *adopts* its type.
- Each individual in the active population also carries an *exchange clock* that rings at rate λ. When the clock rings, the individual randomly chooses a dormant individual and exchanges state, i.e. becomes dormant and forces the chosen dormant individual to become active. During the exchange, the two individuals *retain* their type.

Since the sizes of the two populations remain constant, we only need two variables to describe the dynamics of the population, namely the number of a type-♡ individuals in both populations (see Table 1).

Let x and y denote the number of individuals of type ♡ in the active and the dormant population, respectively. After a resampling event, (x, y) can change to $(x-1, y)$ or $(x+1, y)$, while after an exchange event (x, y) can change to $(x-1, y+1)$ or $(x+1, y-1)$. Both changes in the resampling event occur at rate $x\frac{N-x}{N}$. In the

exchange event, however, to see (x, y) change to $(x-1, y+1)$, an exchange clock of a type-♡ individual in the active population has to ring (which happens at rate λx), and that individual has to choose a type-♠ individual in the dormant population (which happens with probability $\frac{M-y}{M}$). Hence, the total rate at which (x, y) changes to $(x-1, y+1)$ is $\lambda x \frac{M-y}{M}$. By the same argument, the total rate at which (x, y) changes to $(x+1, y-1)$ is $\lambda(N-x)\frac{y}{M}$.

For convenience we multiply the rate of resampling by a factor $\frac{1}{2}$, in order to make it compatible with the Fisher–Wright model. Thus, the generator G of the process is given by

$$G = G_{\mathrm{Mor}} + G_{\mathrm{Exc}}, \tag{3.1}$$

where

$$(G_{\mathrm{Mor}} f)(x, y) = \frac{x(N-x)}{2N}\,[f(x-1, y) + f(x+1, y) - 2f(x, y)] \tag{3.2}$$

describes the Moran resampling of active individuals at rate $\frac{1}{2}$ and

$$\begin{aligned}(G_{\mathrm{Exc}} f)(x, y) = {} & \frac{\lambda}{M} x(M-y)\,[f(x-1, y+1) - f(x, y)] \\ & + \frac{\lambda}{M} y(N-x)\,[f(x+1, y-1) - f(x, y)]\end{aligned} \tag{3.3}$$

describes the exchange between active and dormant individuals at rate λ. From here onwards, we denote the Markov process associated with the generator G by

$$Z = (Z(t))_{t\geq 0}, \qquad Z(t) = (X(t), Y(t)), \tag{3.4}$$

where $X(t)$ and $Y(t)$ are the number of type-♡ active and dormant individuals at time t, respectively. The process Z has state space $[N] \times [M]$, where $[N] = \{0, 1, \ldots, N\}$ and $[M] = \{0, 1, \ldots, M\}$. Note that Z is well defined because it is a continuous-time Markov chain with finitely many states.

3.1.2 Duality and Equilibrium

The classical Moran model is known to be dual to the block-counting process of the Kingman coalescent. In this section, we show that the single-colony Moran model with seed-bank also has a coalescent dual.

Definition 3.1 (Block-counting process) The *block-counting process* of the interacting seed-bank coalescent (defined in Definition 3.5) is the continuous-time Markov chain

$$Z^* = (Z^*(t))_{t\geq 0}, \qquad Z^*(t) = (n_t, m_t), \tag{3.5}$$

taking values in the state space $[N] \times [M]$ with transition rates

$$(n,m) \mapsto \begin{array}{ll} (n-1, m+1) & \text{at rate } \lambda n \left(1 - \frac{m}{M}\right), \\ (n+1, m-1) & \text{at rate } \lambda K m \left(1 - \frac{n}{N}\right), \\ (n-1, m) & \text{at rate } \frac{1}{N} \binom{n}{2} \mathbf{1}_{\{n \geq 2\}}, \end{array} \tag{3.6}$$

where $K = \frac{N}{M}$ is the *ratio* of the sizes of the active and the dormant population. □

The first two transitions in (3.6) correspond to exchange, the third transition to resampling. Later in this section we describe the associated *interacting seed-bank coalescent* process, which gives the genealogy of Z.

The following result gives the duality between Z and Z^*.

Theorem 3.2 (Duality) *The process Z is dual to the process Z^* via the duality relation*

$$\mathbb{E}_{(X,Y)} \left[\frac{\binom{X(t)}{n}}{\binom{N}{n}} \frac{\binom{Y(t)}{m}}{\binom{M}{m}} \mathbf{1}_{\{n \leq X(t), m \leq Y(t)\}} \right] = \mathbb{E}^{(n,m)} \left[\frac{\binom{X}{n(t)}}{\binom{N}{n(t)}} \frac{\binom{Y}{m(t)}}{\binom{M}{m(t)}} \mathbf{1}_{\{n(t) \leq X, m(t) \leq Y\}} \right], \quad t \geq 0, \tag{3.7}$$

where $\mathbb{E}$ stands for generic expectation. On the left the expectation is taken over Z with initial state $Z(0) = (X, Y) \in [N] \times [M]$; on the right the expectation is taken over Z^* with initial state $Z^*(0) = (n, m) \in [N] \times [M]$.

Note that the duality relation fixes the factorial moments and thereby the mixed moments of the random vector $(X(t), Y(t))$. This enables us to determine the equilibrium distribution of Z.

Although the above duality is new in the literature on seed-banks, the notion of factorial duality is not uncommon in mathematical models involving finite and fixed population sizes [8,12]. Similar types of dualities are often found for other models too (e.g. self-duality of independent random walks, exclusion and inclusion processes, etc. [11]). Remarkably, in the special case where $N = M = 2j$ for some $j \in \mathbb{N}/2$, Giardinà et al. (2009) [11, Section 3.2] identified the same duality relation as in (3.7) as a self-duality for the generalized $2j$-SEP on two-sites. This is not surprising given the fact that the exchange rates between active and dormant individuals defined in Table 1 are precisely the rates (up to rescaling) for the $2j$-SEP on two sites. We refer the reader to Sect. 4.1 to gain further insights into this.

Proposition 3.3 (Convergence of moments) *For any $(X, Y), (n, m) \in [N] \times [M]$ with $(n, m) \neq (0, 0)$,*

$$\lim_{t \to \infty} \mathbb{E}_{(X,Y)} \left[X(t)^n Y(t)^m \right] = N^n M^m \frac{X + Y}{N + M}. \tag{3.8}$$

Since the vector $(X(t), Y(t))$ takes values in $[N] \times [M]$, which has $(N+1)(M+1)$ points, the above proposition determines the limiting distribution of $(X(t), Y(t))$.

Corollary 3.4 (Equilibrium) *Suppose that Z starts from initial state $(X, Y) \in [N] \times [M]$. Then, $(X(t), Y(t))$ converges in law as $t \to \infty$ to a random vector (X_∞, Y_∞) whose distribution is given by*

$$\mathcal{L}_{(X,Y)}(X_\infty, Y_\infty) = \frac{X+Y}{N+M}\,\delta_{(N,M)} + \left(1 - \frac{X+Y}{N+M}\right)\delta_{(0,0)}. \tag{3.9}$$

Note that the equilibrium behaviour of Z is the same as for the classical Moran model without seed-bank. The fixation probability of type $\heartsuit$ is $\frac{X+Y}{N+M}$, which is nothing but the initial frequency of type-$\heartsuit$ individuals in the *entire population*. Even though the presence of the seed-bank delays the time of fixation, because its size is finite size it has no significant effect on the overall qualitative behaviour of the process. We will see in Sect. 3.2 that the situation is different in the multi-colony model.

3.1.3 Interacting Seed-Bank Coalescent

In our model, the *genealogy* of a sample taken from the finite population of $N + M$ individuals is governed by a partition-valued coalescent process similarly as for the genealogy of the classical Moran model. However, due to the presence of the seed-bank, blocks of a partition are marked as A (active) and D (dormant). Unlike in the genealogy of the classical Moran model, the blocks *interact* with each other. This interaction is present because of the restriction to *finite size* of the active and the dormant population. For this reason, we name the block process an *interacting seed-bank coalescent*. For convenience, we will use the word lineage to refer to a block in a partition.

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

$$\mathcal{M}_{j,k,l} = \Big\{ \boldsymbol{u} \in \{A, D\}^j : \text{ the numbers of } A \text{ and } D \text{ in } \boldsymbol{u} \text{ are at most } k \text{ and } l, \text{ respectively} \Big\}. \tag{3.10}$$

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

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

Definition 3.5 (Interacting seed-bank coalescent) The *interacting seed-bank coalescent* is the continuous-time Markov chain with state space $\mathcal{P}_{M,N}$ characterized by the following transition rates:

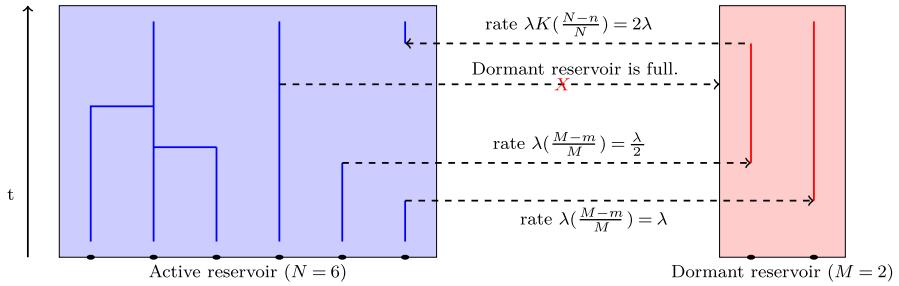

Fig. 1 Scheme of transitions for an interacting particle system with an active reservoir of size $N = 6$ and a dormant reservoir of size $M = 2$, so that $K = \frac{N}{M} = \frac{6}{2} = 3$. The effective rate for each of n active particles to become dormant is $\lambda \frac{M-m}{M}$ when the dormant reservoir has m particles. Similarly, the effective rate for each of m dormant particles to become active is $\lambda K \frac{N-n}{N}$ when the active reservoir has n particles

$$\pi \mapsto \pi' \text{ at rate}$$
$$\begin{array}{ll} \frac{1}{N} & \text{if } \pi \succ \pi', \\ \lambda\big(1 - \frac{|\pi|_D}{M}\big) & \text{if } \pi \bowtie \pi' \text{ by change of state of one lineage in } \pi \text{ from } A \text{ to } D, \\ \lambda K\big(1 - \frac{|\pi|_A}{N}\big) & \text{if } \pi \bowtie \pi' \text{ by change of state of one lineage in } \pi \text{ from } D \text{ to } A. \end{array} \tag{3.11}$$

□

The factor $1 - \frac{|\pi|_D}{M}$ in the transition rate of a single active lineage when π becomes dormant reflects the fact that as the seed-bank gets full, it becomes more difficult for an active lineage to enter the seed-bank. Similarly, as the number of active lineages decreases due to the coalescence, it becomes easier for a dormant lineage to leave the seed-bank and become active. This also tells us that there is a *repulsive interaction* between the lineages of the same state (A or D). Due to this interaction, it is tricky to study the coalescent. As N, M get large, the interaction becomes weak. As $N, M \to \infty$, after proper space-time scaling, the coalescent converges weakly to a limit coalescent where the interaction is no longer present. In fact, it can be shown that when both the time and the parameters are scaled properly, the coalescent converges weakly as $N, M \to \infty$ to the *seed-bank coalescent* described in [1].

We can also describe the coalescent in terms of an interacting particle system with the help of a graphical representation (see Fig. 1). The interacting particle system consists of two reservoirs, called *active* reservoir and *dormant* reservoir, having N and M labelled sites, respectively, each of which can be occupied by at most one particle. The particles in the active and dormant reservoir are called *active* and *dormant* particles, respectively. The active particles can coalesce with each other, in the sense that if an active particle occupies a labelled site where an active particle is present already, then the two particles are glued together to form a single particle at that site. Active particles can become dormant by moving to an empty site in the dormant reservoir, while dormant particles can become active by moving to an empty site in the active reservoir. The transition rates are as follows:

- An active particle tries to coalesce with another active particle at rate $\frac{1}{2}$ by choosing uniformly at random a labelled site in the active reservoir. If the chosen site is

empty, then it ignores the transition; otherwise, it coalesces with the active particle present at the new site.
- An active particle becomes dormant at rate λ by moving to a random labelled site in the dormant reservoir when the chosen site is empty; otherwise, it remains in the active reservoir.
- A dormant particle becomes active at rate λK by moving to a random labelled site in the active reservoir when the chosen site is empty; otherwise, it remains in the dormant reservoir.

Clearly, the particles *interact with each other* due to the finite capacity of the two reservoirs. If $N, M \to \infty$, then the probability to choose an empty site in a reservoir tends to 1, and so the system converges (after proper scaling) to an interacting particle system where the particles move independently between the two reservoirs.

Note that if we define n_t = number of active particles at time t and m_t = number of dormant particles at time t, then $Z^* = (n_t, m_t)_{t\geq 0}$ is the block-counting process defined in Definition 3.1. Also, if we remove the labels of the sites in the two reservoirs and represents the particle configuration by an element of $\mathcal{P}_{N,M}$, then we obtain the *interacting seed-bank coalescent* described in Definition 3.5. Even though it is natural to describe the genealogical process via a partition-valued stochastic process, we will stick with the interacting particle system description of the dual, since this will be more convenient for the multi-colony model.

3.2 Multi-colony Model

In this section, we consider multiple colonies, each with their own seed-bank. Each colony has an *active* population and a *dormant* population. We take $\mathbb{Z}^d$ as the underlying *geographic space* where the colonies are located (any countable Abelian group will do). With each colony $i \in \mathbb{Z}^d$ we associate a variable (X_i, Y_i), with X_i and Y_i the number of type-$\heartsuit$ active and dormant individuals, respectively, at colony i. Let (N_i, M_i) denote the size of the active and the dormant population at colony i. In each colony, active individuals are subject to resampling and migration, and to exchange with dormant individuals that are in the same colony. Dormant individuals are not subject to resampling and migration.

Since it is crucial for our duality to keep the population sizes constant, we consider migration of types without the individuals actually moving themselves. To be precise, by a migration from colony j to colony i we mean that an active individual from colony i randomly chooses an active individual from colony j and adopts its type. In this way, the *genetic information* moves from colony j to colony i, while the individuals themselves stay put.

3.2.1 Definition: Resampling, Exchange and Migration

We assume that each active individual at colony i resamples from colony j at rate $a(i, j)$, adopting the type of a uniformly chosen active individual at colony j. Here, the *migration kernel* $a(\cdot, \cdot)$ is assumed to satisfy Assumption 2.1. After a migration to colony i, the only variable that is affected is X_i, the number of type-$\heartsuit$ active

Table 2 Scheme of transitions in the multi-colony model

Initial state	Event	Final state	Transition rate
$(X_i, Y_i)_{i\in\mathbb{Z}^d}$	Migration from colony j to i	$(\cdots, (X_i - 1, Y_i), \cdots)$	$a(i, j)X_i(N_j - X_j)/N_j$
		$(\cdots, (X_i + 1, Y_i), \cdots)$	$a(i, j)(N_i - X_i)X_j/N_j$
	Exchange at colony i	$(\cdots, (X_i - 1, Y_i + 1), \cdots)$	$\lambda X_i(M_i - Y_i)/M_i$
		$(\cdots, (X_i + 1, Y_i - 1), \cdots)$	$\lambda(N_i - X_i)Y_i/M_i$

individuals at colony i. The final state can be either $X_i - 1$ or $X_i + 1$ depending on whether a type-♡ active individual from colony i chooses a type-♠ active individual from another colony or a type-♠ active individual from colony i chooses a type-♡ active individual from another colony. The rate at which X_i changes to $X_i - 1$ due to a migration from colony j is

$$a(i, j)X_i \tfrac{N_j - X_j}{N_j},$$

while the rate at which X_i changes to $X_i + 1$ due to a migration from colony j is

$$a(i, j)(N_i - X_i)\tfrac{X_j}{N_j}.$$

Note that for $i = j$ the migration rate is

$$a(i, i)X_i \tfrac{N_i - X_i}{N_i} = \tfrac{X_i(N_i - X_i)}{2N_i},$$

which is the same as the effective birth and death rate in the single-colony Moran model. Thus, the resampling within each colony is already taken care of via the migration.

It remains to define the associated exchange mechanism between the active and the dormant individuals in a colony. The exchange mechanism is the same as in the single-colony model, i.e. in each colony each active individual at rate λ performs an exchange with a dormant individual chosen uniformly from the seed-bank of that colony. For simplicity, we take the exchange rate λ to be the same in each colony.

The state space $\mathcal{X}$ of the process is

$$\mathcal{X} := \prod_{i\in\mathbb{Z}^d} \{0, 1, \ldots, N_i\} \times \{0, 1, \ldots, M_i\} = \prod_{i\in\mathbb{Z}^d} [N_i] \times [M_i]. \tag{3.12}$$

A configuration $\eta \in \mathcal{X}$ is denoted by $\eta = (X_i, Y_i)_{i\in\mathbb{Z}^d}$, with $X_i \in [N_i]$ and $Y_i \in [M_i]$. For each $i \in \mathbb{Z}^d$, let $\boldsymbol{\delta}_{i,A} = (X_n, Y_n)_{n\in\mathbb{Z}^d}$ and $\boldsymbol{\delta}_{i,D} = (\hat{X}_n, \hat{Y}_n)_{n\in\mathbb{Z}^d}$ be the configurations defined as

$$(X_n, Y_n) := (\mathbf{1}_{\{n=i\}}, 0), \quad (\hat{X}_n, \hat{Y}_n) := (0, \mathbf{1}_{\{n=i\}}) \qquad \forall\, n \in \mathbb{Z}^d. \tag{3.13}$$

For two configurations $\eta = (\bar{X}_i, \bar{Y}_i)_{i\in\mathbb{Z}^d}$ and $\xi = (\hat{X}_i, \hat{Y}_i)_{i\in\mathbb{Z}^d}$, we define $\eta \pm \xi := (X_i, Y_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}$ by setting, for each $i \in \mathbb{Z}^d$,

$$\begin{aligned} X_i &= (\bar{X}_i \pm \hat{X}_i)\mathbf{1}_{\{0\le \bar{X}_i \pm \hat{X}_i \le N_i\}} + N_i\, \mathbf{1}_{\{\bar{X}_i \pm \hat{X}_i > N_i\}}, \\ Y_i &= (\bar{Y}_i \pm \hat{Y}_i)\mathbf{1}_{\{0\le \bar{Y}_i \pm \hat{Y}_i \le M_i\}} + M_i\, \mathbf{1}_{\{\bar{Y}_i \pm \hat{Y}_i > M_i\}}. \end{aligned} \tag{3.14}$$

Throughout the remainder of this paper, we adopt the convention given in (3.14) for addition and subtraction of configurations in $\mathcal{X}$.

The generator L for the process, acting on functions in

$$\mathcal{D} = \big\{ f \in C(\mathcal{X}) \colon\ f \text{ depends on finitely many coordinates} \big\}, \tag{3.15}$$

is given by

$$L = L_{\mathrm{Mig}} + L_{\mathrm{Res}} + L_{\mathrm{Exc}}, \tag{3.16}$$

where

$$\begin{aligned} (L_{\mathrm{Mig}} f)(\eta) = \sum_{i\in\mathbb{Z}^d} \sum_{\substack{j\in\mathbb{Z}^d,\\ j\neq i}} \frac{a(i,j)}{N_j} \Big\{ & X_i(N_j - X_j)[f(\eta - \boldsymbol{\delta}_{i,A}) - f(\eta)] \\ & + X_j(N_i - X_i)[f(\eta + \boldsymbol{\delta}_{i,A}) - f(\eta)] \Big\} \end{aligned} \tag{3.17}$$

describes the resampling of active individuals in *different* colonies (= migration),

$$(L_{\mathrm{Res}} f)(\eta) = \sum_{i\in\mathbb{Z}^d} \frac{X_i(N_i - X_i)}{2N_i} [f(\eta - \boldsymbol{\delta}_{i,A}) + f(\eta + \boldsymbol{\delta}_{i,A}) - 2f(\eta)] \tag{3.18}$$

describes the resampling of active individuals in the *same* colony, and

$$\begin{aligned} (L_{\mathrm{Exc}} f)(\eta) = \sum_{i\in\mathbb{Z}^d} \frac{\lambda}{M_i} \Big\{ & X_i(M_i - Y_i)[f(\eta - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{i,D}) - f(\eta)] \\ & + Y_i(N_i - X_i)[f(\eta + \boldsymbol{\delta}_{i,A} - \boldsymbol{\delta}_{i,D}) - f(\eta)] \Big\} \end{aligned} \tag{3.19}$$

describes the exchange of active and dormant individuals in the *same* colony.

From now on, we denote the process associated with the generator L by

$$Z = (Z(t))_{t\ge 0}, \qquad Z(t) = (X_i(t), Y_i(t))_{i\in\mathbb{Z}^d}, \tag{3.20}$$

with $X_i(t)$ and $Y_i(t)$ representing the number of type-$\heartsuit$ active and dormant individuals at colony i at time t, respectively. Since Z is an interacting particle system, in order to show existence and uniqueness of the process, we can in principle follow the method described by Liggett in [25, Chapter I, Section 3]. However, for Liggett's method to

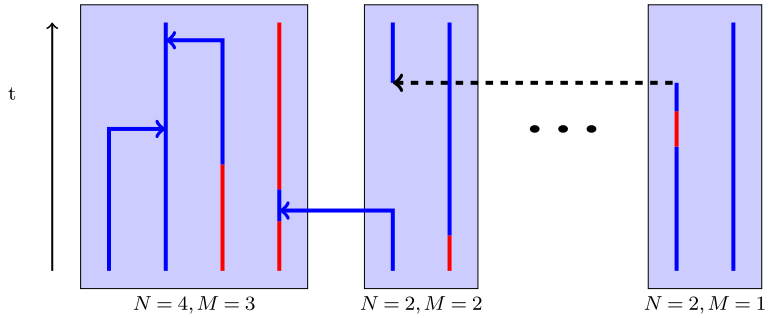

Fig. 2 Scheme of transitions in the interacting particle system. Each block depicts the reservoirs located at sites of $\mathbb{Z}^d$. The blue lines represent the evolution of active particles, and the red lines represent the evolution of dormant particles (Color figure online)

work, a uniform bound on the sizes $(N_i, M_i)_{i\in\mathbb{Z}^d}$ is needed that we want to avoid. Fortunately, if L is a Markov pregenerator (see [25, Definition 2.1]), then we can construct the process by providing a unique solution to the martingale problem for L. The following proposition tells us that L is indeed a Markov pregenerator and thus prepares the ground for proving the well-posedness of the martingale problem for L.

Proposition 3.6 (Pregenerator) *The generator L defined in* (3.16)*, acting on functions in $\mathcal{D}$ defined in* (3.15)*, is a Markov pregenerator.*

The existence of solutions to the martingale problem will be shown by using the techniques described in [25]. In order to establish uniqueness of the solution, we will need to exploit the dual process.

3.2.2 Duality

The dual process is a block-counting process associated with a spatial version of the interacting seed-bank coalescent described in Sect. 3.1.3. We briefly describe the spatial coalescent process in terms of an interacting particle system. At each site $i \in \mathbb{Z}^d$, there are two reservoirs, an *active* reservoir and a *dormant* reservoir, with $N_i \in \mathbb{N}$ and $M_i \in \mathbb{N}$ labelled locations, respectively. Each location in a reservoir can accommodate at most one particle. As before, we refer to the particles in an active and dormant reservoir as *active* particles and *dormant* particles, respectively. The dynamics of the interacting particle system is as follows (see Fig. 2).

- An active particle at site $i \in \mathbb{Z}^d$ becomes dormant at rate λ by moving to a random labelled location (out of M_i many) in the dormant reservoir at site i when the chosen labelled location is empty; otherwise, it remains in the active reservoir.
- A dormant particle at site $i \in \mathbb{Z}^d$ becomes active at rate λK_i with $K_i = \frac{N_i}{M_i}$ by moving to a random labelled location (out of N_i many) in the active reservoir at site i when the chosen labelled location is empty; otherwise, it remains in the dormant reservoir.
- An active particle at site i chooses a random labelled location (out of N_j many) from the active reservoir at site j at rate $a(i, j)$ and does the following:

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

Note that an active particle can migrate between different sites in $\mathbb{Z}^d$ and can coalesce with another active particle even when they are at different sites in $\mathbb{Z}^d$. For simplicity, we will impose the same assumptions on the migration kernel $a(\cdot,\cdot)$ as stated in Assumption 2.1. A configuration $(\eta_i)_{i\in\mathbb{Z}^d}$ of the particle system is an element of $\prod_{i\in\mathbb{Z}^d}\{0,1\}^{N_i}\times\{0,1\}^{M_i}$. For $i\in\mathbb{Z}^d$, η_i represents the state of the labelled locations in the active and the dormant reservoir at site i (1 means occupied by a particle, 0 means empty).

Below we give the definition of the block-counting process associated with the spatial coalescent process described above. Although it is an interesting problem to construct the block-counting process starting from a configuration with infinitely many particles, we will restrict ourselves to configurations with *finitely many particles* only because this makes the state space countable. Thus, the block-counting process is a continuous-time Markov chain on a countable state space, and hence, in the definition below, it suffices to specify the possible transitions and their respective rates only.

Definition 3.7 (Dual) The dual process

$$Z^* = (Z^*(t))_{t\geq 0}, \qquad Z^*(t) = (n_i(t), m_i(t))_{i\in\mathbb{Z}^d}, \tag{3.21}$$

is a continuous-time Markov chain with state space

$$\mathcal{X}^* := \Big\{(n_i,m_i)_{i\in\mathbb{Z}^d} \in \prod_{i\in\mathbb{Z}^d}[N_i]\times[M_i]\colon \sum_{i\in\mathbb{Z}^d}(n_i+m_i)<\infty\Big\} \tag{3.22}$$

and with transition rates

$$(n_k,m_k)_{k\in\mathbb{Z}^d} \to \begin{cases} (n_k,m_k)_{k\in\mathbb{Z}^d}-\boldsymbol{\delta}_{i,A} & \text{at rate } \frac{2a(i,i)}{N_i}\binom{n_i}{2}\mathbf{1}_{\{n_i\geq 2\}}+\sum_{\substack{j\in\mathbb{Z}^d,\\ j\neq i}}\frac{n_i a(i,j)n_j}{N_j} \text{ for } i\in\mathbb{Z}^d,\\ (n_k,m_k)_{k\in\mathbb{Z}^d}-\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{i,D} & \text{at rate } \frac{\lambda n_i(M_i-m_i)}{M_i} \quad \text{for } i\in\mathbb{Z}^d,\\ (n_k,m_k)_{k\in\mathbb{Z}^d}+\boldsymbol{\delta}_{i,A}-\boldsymbol{\delta}_{i,D} & \text{at rate } \frac{\lambda(N_i-n_i)m_i}{M_i} \quad \text{for } i\in\mathbb{Z}^d,\\ (n_k,m_k)_{k\in\mathbb{Z}^d}-\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,A} & \text{at rate } \frac{n_i a(i,j)(N_j-n_j)}{N_j} \quad \text{for } i\neq j\in\mathbb{Z}^d, \end{cases} \tag{3.23}$$

where the configurations $\boldsymbol{\delta}_{i,A}, \boldsymbol{\delta}_{i,D}\in\mathcal{X}^*\subset\mathcal{X}$ are as in (3.13), and additions and subtractions of configurations are performed in accordance with (3.14). □

Here, $n_i(t)$ and $m_i(t)$ are the number of active and dormant particles at site $i\in\mathbb{Z}^d$ at time t. The first transition describes the coalescence of an active particle at site i with other active particles elsewhere. The second and third transitions describe the movement of particles between the active and the dormant reservoir at site i. The fourth transition describes the migration of an active particle from site i to site j. The

following lemma tells us that the dual process Z^* is a well-defined and non-explosive (equivalent to uniqueness) Feller process on the countable state space $\mathcal{X}^*$.

Lemma 3.8 (Uniqueness of dual) *There exists a unique minimal Feller process* $(Z^*(t))_{t\geq 0}$ *on* $\mathcal{X}^*$ *with transition rates given in* (3.23).

Before we proceed we recall the definition of the martingale problem.

Definition 3.9 (Martingale problem) Suppose that $(L, \mathcal{D})$ is a Markov pregenerator, and let $\eta \in \mathcal{X}$. A probability measure $\mathbb{P}_\eta$ (or, equivalently, a process with law $\mathbb{P}_\eta$) on $D([0,\infty), \mathcal{X})$ is said to solve the martingale problem for L with initial point η if

- $\mathbb{P}_\eta[\xi_{(\cdot)} \in D([0,\infty), \mathcal{X}) : \xi_0 = \eta] = 1$.
- $(f(\eta_t) - \int_0^t (Lf)(\eta_s)\,\mathrm{d}s)_{s\geq 0}$ is a martingale relative to $(\mathbb{P}_\eta, (\mathcal{F}_t)_{t\geq 0})$ for all $f \in \mathcal{D}$, where $(\eta_t)_{t\geq 0}$ is the coordinate process on $D([0,\infty), \mathcal{X})$ and $(\mathcal{F}_t)_{t\geq 0}$ is the filtration given by $\mathcal{F}_t := \sigma(\eta_s \mid s \leq t)$ for $t \geq 0$. □

The following theorem gives the duality relation between the dual process Z^* and any solution to the martingale problem for $(L, \mathcal{D})$. This type of duality is sometimes referred to as martingale duality.

Theorem 3.10 (Duality relation) *Let the process* Z *with law* $\mathbb{P}_\eta$ *be a solution to the martingale problem for* $(L, \mathcal{D})$ *starting from initial state* $\eta = (X_i, Y_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}$. *Let* Z^* *be the dual process with law* $\mathbb{P}^\xi$ *starting from initial state* $\xi = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$. *For* $t \geq 0$, *let* $\Gamma(t)$ *be the random variable defined by*

$$\Gamma(t) := \max\Big\{\|i\| :\ i \in \mathbb{Z}^d,\ n_i(s) + m_i(s) > 0 \text{ for some } 0 \leq s \leq t\Big\}. \tag{3.24}$$

Suppose that the sizes $(N_i)_{i\in\mathbb{Z}^d}$ *of the active populations are such that for any* $T > 0$,

$$\sum_{i\in\mathbb{Z}^d} N_i\, \mathbb{P}^\xi\big(\Gamma(T) \geq \|i\|\big) < \infty. \tag{3.25}$$

Then, for any $t \geq 0$,

$$\begin{aligned}&\mathbb{E}_\eta\left[\prod_{i\in\mathbb{Z}^d} \frac{\binom{X_i(t)}{n_i}}{\binom{N_i}{n_i}} \frac{\binom{Y_i(t)}{m_i}}{\binom{M_i}{m_i}} \mathbf{1}_{\{n_i \leq X_i(t), m_i \leq Y_i(t)\}}\right] \\ &\quad = \mathbb{E}^\xi\left[\prod_{i\in\mathbb{Z}^d} \frac{\binom{X_i}{n_i(t)}}{\binom{N_i}{n_i(t)}} \frac{\binom{Y_i}{m_i(t)}}{\binom{M_i}{m_i(t)}} \mathbf{1}_{\{n_i(t) \leq X_i, m_i(t) \leq Y_i\}}\right],\end{aligned} \tag{3.26}$$

where the expectations are taken with respect to $\mathbb{P}_\eta$ *and* $\mathbb{P}^\xi$, *respectively.*

Note that the duality function is a product over all colonies of the duality function that appeared in the single-colony model. The infinite products are well defined: all but finitely many factors are 1, because of our assumption that there are only *finitely many*

particles in the dual process. Also note that there is no restriction on $(M_i)_{i\in\mathbb{Z}^d}$, the sizes of the dormant populations. This is because dormant individuals do not migrate and therefore do not feel the spatial extent of the system.

At first glance, it may seem that (3.25) places a severe restriction on $(N_i)_{i\in\mathbb{Z}^d}$, the sizes of the active populations. However, this is not the case. The following corollary provides us with a large class of active population sizes for which Theorem 3.10 is true under mild assumptions on the migration kernel $a(\cdot,\cdot)$.

Corollary 3.11 (Duality criterion) *Suppose that Assumption 2.1 is in force. Then, (3.25) and consequently the duality relation in (3.26) hold for every $(N_i)_{i\in\mathbb{Z}^d} \in \mathcal{N}$, where*

(a) either

$$\mathcal{N} := \left\{ (N_i)_{i\in\mathbb{Z}^d} \in \mathbb{N}^{\mathbb{Z}^d} : \lim_{\|i\|\to\infty} \frac{1}{\|i\|} \log N_i = 0 \right\} \tag{3.27}$$

when $\sum_{i\in\mathbb{Z}^d} \mathrm{e}^{\delta\|i\|} a(0,i) < \infty$ for some $\delta > 0$,

(b) or

$$\mathcal{N} := \left\{ (N_i)_{i\in\mathbb{Z}^d} \in \mathbb{N}^{\mathbb{Z}^d} : \sup_{i\in\mathbb{Z}^d\setminus\{0\}} \frac{N_i}{\|i\|^\delta} < \infty \right\} \tag{3.28}$$

when $\sum_{i\in\mathbb{Z}^d} \|i\|^\gamma a(0,i) < \infty$ for some $\delta > 0$ and $\gamma > d + \delta$.

Corollary 3.11 shows a *trade-off*: the more we restrict the tails of the migration kernel, the less we need to restrict the sizes of the active populations.

3.2.3 Well-Posedness

We use a martingale problem for the generator L defined in (3.16), in the sense of [9, p.173], to construct Z. The following proposition gives existence of solutions for any choice of the reservoir sizes. As for the uniqueness of solutions, we will see that a restriction on the sizes of the active populations is required.

Proposition 3.12 (Existence) *Let L be the generator defined in (3.16) acting on the set of local functions $\mathcal{D}$ defined in (3.15). Then, for all $\eta \in \mathcal{X}$ there exists a solution $\mathbb{P}_\eta$ (a probability measure on $D([0,\infty),\mathcal{X})$) to the martingale problem of $(L,\mathcal{D})$ with initial state η.*

The following theorem gives the well-posedness of the martingale problem for $(L,\mathcal{D})$ under a restricted class of sizes of the active populations and thus proves the existence of a unique Feller Markov process describing our multi-colony model.

Theorem 3.13 (Well-posedness) *Let $(N_i)_{i\in\mathbb{Z}^d} \in \mathcal{N}$ and $(M_i)_{i\in\mathbb{Z}^d} \in \mathbb{N}^{\mathbb{Z}^d}$, and let L be the generator defined in (3.16) acting on the set of local functions $\mathcal{D}$ defined in (3.15). Then, the following hold:*

- *For all* $\eta \in \prod_{i\in\mathbb{Z}^d}[N_i]\times[M_i]$*, there exists a unique solution* Z *in* $D([0,\infty),\mathcal{X})$ *of the martingale problem for* $(L,\mathcal{D})$ *with initial state* η.
- Z *is Feller and strong Markov, and its generator is an extension of* $(L,\mathcal{D})$.

In view of the above result, from here onwards, we implicitly assume that the restriction on $(N_i)_{i\in\mathbb{Z}^d}$ to $\mathcal{N}$ is always in force.

3.2.4 Equilibrium

Let us set $Z_i(t) := (X_i(t), Y_i(t))$ for $i \in \mathbb{Z}^d$ and denote by $\mu(t)$ the distribution of $Z(t)$. Further, for each $\theta \in [0,1]$ and $i \in \mathbb{Z}^d$, let ν_θ^i be the probability measure on $[N_i]\times[M_i]$ defined as

$$\nu_\theta^i := \text{Binomial}(N_i,\theta)\otimes\text{Binomial}(M_i,\theta). \tag{3.29}$$

For $\theta \in [0,1]$, let ν_θ be the distribution on $\mathcal{X}$ defined by $\nu_\theta := \bigotimes_{i\in\mathbb{Z}^d}\nu_\theta^i$ and set

$$\mathcal{J} := \{\nu_\theta \,|\, \theta\in[0,1]\}. \tag{3.30}$$

Let $D : \mathcal{X}\times\mathcal{X}^* \to [0,1]$ be the function defined by

$$D((X_k,Y_k)_{k\in\mathbb{Z}^d};(n_k,m_k)_{k\in\mathbb{Z}^d}) := \prod_{i\in\mathbb{Z}^d}\frac{\binom{X_i}{n_i}\binom{Y_i}{m_i}}{\binom{N_i}{n_i}\binom{M_i}{m_i}}\mathbf{1}_{\{n_i\le X_i, m_i\le Y_i\}}. \tag{3.31}$$

Theorem 3.14 (Convergence to equilibrium) *Suppose that* $\mu(0)=\nu_\theta\in\mathcal{J}$ *for some* $\theta\in[0,1]$. *Then, there exists a probability measure* ν *determined by the parameter* θ *such that*

- $\lim_{t\to\infty}\mu(t)=\nu$.
- ν *is an equilibrium for the process* Z.
- $\mathbb{E}_\nu[D(Z(0);\eta)] = \lim_{t\to\infty}\mathbb{E}^\eta[\theta^{|Z^*(t)|}]$, *where* $D(\cdot,\cdot)$ *is defined in* (3.31), *the right expectation is taken w.r.t. the dual process* Z^* *started at configuration* $\eta=(n_i,m_i)_{i\in\mathbb{Z}^d}\in\mathcal{X}^*$ *and* $|Z^*(t)| := \sum_{i\in\mathbb{Z}^d}[n_i(t)+m_i(t)]$ *is the total number of dual particles present at time* t.

Corollary 3.15 *Let* ν *be the equilibrium measure of* Z *in Theorem* 3.14 *corresponding to* $\theta\in[0,1]$. *Then,*

$$\mathbb{E}_\nu\left[\frac{X_i(0)}{N_i}\right] = \mathbb{E}_\nu\left[\frac{Y_i(0)}{M_i}\right] = \theta. \tag{3.32}$$

3.2.5 Clustering Criterion

We next analyse the long-time behaviour of the multi-colony Moran model with seed-banks. Our interest is to capture the nature of the equilibrium. To be precise, we

investigate whether coexistence of different types is possible in equilibrium. The measures $\bigotimes_{i\in\mathbb{Z}^d}\delta_{(0,0)}$ and $\bigotimes_{i\in\mathbb{Z}^d}\delta_{(N_i,M_i)}$ are the trivial equilibria where the system concentrates on only one of the two types. When the system converges to an equilibrium that is not a mixture of these two trivial equilibria, we say that *coexistence* happens. For $i\in\mathbb{Z}^d$, let us denote the frequency of type-♡ active and dormant individuals at colony i at time t by $x_i(t):=\frac{X_i(t)}{N_i}$ and $y_i(t):=\frac{Y_i(t)}{M_i}$ respectively.

Definition 3.16 (Clustering and Coexistence) The system is said to exhibit *clustering* if the following hold:

- $\lim_{t\to\infty}\mathbb{P}_\eta(x_i(t)\in\{0,1\})=1,\quad \lim_{t\to\infty}\mathbb{P}_\eta(y_i(t)\in\{0,1\})=1,$
- $\lim_{t\to\infty}\mathbb{P}_\eta(x_i(t)\neq x_j(t))=0,\quad \lim_{t\to\infty}\mathbb{P}_\eta(y_i(t)\neq y_j(t))=0,$
- $\lim_{t\to\infty}\mathbb{P}_\eta(x_i(t)\neq y_j(t))=0,$

for all $i,j\in\mathbb{Z}^d$ and any initial configuration $\eta\in\mathcal{X}$. Otherwise, the system is said to exhibit *coexistence*. □

The above conditions make sure that if an equilibrium exists, then it is a mixture of the two trivial equilibria.

The following criterion, which follows from Corollary 3.11, gives an equivalent condition for clustering.

Theorem 3.17 (Clustering criterion) *The system clusters if and only if in the dual process defined in Definition 3.7 two particles, starting from any locations in $\mathbb{Z}^d$ and any states (active or dormant), coalesce with probability* 1.

Note that the system clusters if and only if the genetic variability at time t between any two colonies converges to 0 as $t\to\infty$. From the duality relation in Theorem 3.10, it follows that this quantity is determined by the state of the dual process starting from two particles.

4 Proofs: Duality and Equilibrium for the Single-Colony Model

Section 4.1 contains the proof of Theorem 3.2, which follows the algebraic approach to duality described in [4,30]. Section 4.2 contains the proof of Proposition 3.3 and Corollary 3.4, which uses the duality in the single-colony model.

4.1 Duality and Change of Representation

Before we proceed with the proof of Theorem 3.2, and other results related to *stochastic duality*, it is worth stressing the importance of duality theory. Though originally introduced in the context of interacting particle systems, over the last decade duality theory has gained popularity in various fields, ranging from statistical physics and stochastic analysis to population genetics. One reason behind this wide interests is the simplification that duality provides: it often allows one to extract information about a complex stochastic process through a simpler process. To date, in the literature there

exist two systematic approaches towards duality, namely pathwise construction and Lie-algebraic framework. The former of the two approaches is more practical and widespread in the context of mathematical population genetics [7,16,20,21], while the latter has been developed more recently and reveals deeper mathematical structures behind duality, and often also provides a larger class of duality functions (see, for example, [4,10,17,30] for a general overview and further references). In what follows, we adopt the Lie algebraic framework suggested by Carinci et al. (2015) [4] and prepare the ground for this setting. The downside is that this approach does not capture the underlying genealogy of the original process. However, it does offer the opportunity to obtain a larger class of duality functions by applying symmetries from the Lie algebra to an already existing duality function [11]. In this paper, we refrain from exploring the latter aspect of the Lie-algebraic framework.

We start with briefly recalling that a (real) Lie algebra $\mathfrak{g}$ is a linear space over $\mathbb{R}$ endowed with a so-called *Lie bracket* $[\cdot,\cdot]\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ that is bilinear, skew-symmetric and satisfies the *Jacobi identity* [30]. The requirement of the bilinearity and skew-symmetry uniquely characterizes a Lie bracket by its action on a basis of $\mathfrak{g}$. An example of a (real) Lie algebra is the well-known $\mathfrak{su}(2)$-algebra, which is the three-dimensional vector space over $\mathbb{R}$ defined by the action of a Lie bracket on its basis elements $\{J^+, J^-, J^0\}$ as

$$[J^0, J^+] = J^+, \quad [J^0, J^-] = -J^-, \quad [J^-, J^+] = -2J^0. \tag{4.1}$$

For $\alpha\in\mathbb{N}$, let V_α be the linear space of all functions $f\colon [\alpha]\to\mathbb{R}$, and let $\mathfrak{gl}(V_\alpha)$ denote the space of all linear operators on V_α. Note that $\mathfrak{gl}(V_\alpha)$ is a $(1+\alpha)^2$-dimensional Lie algebra with the natural choice of Lie bracket given by $[A,B] := AB - BA$ for $A, B\in\mathfrak{gl}(V_\alpha)$. Let us define the operators $J^{\alpha,\pm}, J^{\alpha,0}, A^{\alpha,\pm}, A^{\alpha,0}\in\mathfrak{gl}(V_\alpha)$ acting on $f:[\alpha]\to\mathbb{R}$ as

$$\begin{aligned} &J^{\alpha,+}f(n) = (\alpha-n)f(n+1), \quad J^{\alpha,-}f(n) = nf(n-1), \quad J^{\alpha,0}f(n) = (n-\tfrac{\alpha}{2})f(n),\\ &A^{\alpha,+} = J^{\alpha,-} - J^{\alpha,+} - 2J^{\alpha,0}, \quad A^{\alpha,-} = J^{\alpha,+}, \quad A^{\alpha,0} = J^{\alpha,+} + J^{\alpha,0}. \end{aligned} \tag{4.2}$$

It is straightforward to see that

$$[A^{\alpha,0}, A^{\alpha,\pm}] = \pm A^{\alpha,\pm}, \quad [A^{\alpha,-}, A^{\alpha,+}] = -2A^{\alpha,0}, \tag{4.3}$$

which are the same commutation relations as in (4.1). Thus, for each $\alpha\in\mathbb{N}$, the Lie homomorphism $\phi_\alpha\colon \mathfrak{su}(2)\to\mathfrak{gl}(V_\alpha)$ defined by its action on the generators $\{J^+, J^-, J^0\}$ given by

$$J^+\mapsto A^{\alpha,+}, \quad J^-\mapsto A^{\alpha,-}, \quad J^0\mapsto A^{\alpha,0}, \tag{4.4}$$

is a finite-dimensional representation of $\mathfrak{su}(2)$. Similarly, we can verify that $\{J^{\alpha,+}, J^{\alpha,-}, J^{\alpha,0}\}$, $\alpha\in\mathbb{N}$, form a representation of the *dual* $\mathfrak{su}(2)$-algebra (defined by the commutation relations in (4.1), but with opposite signs).

Below we introduce the notion of duality between two operators and prove a lemma that will be crucial in the proof of duality of both the single-colony and the multi-colony model. The relevance to our context of the above discussion on $\mathfrak{su}(2)$ and its dual algebra will become clear as we go along.

Definition 4.1 (Operator duality) Let A and B be two operators acting on functions $f\colon\ \Omega \to \mathbb{R}$ and $g\colon\ \hat{\Omega} \to \mathbb{R}$, respectively. We say that A is dual to B with respect to the duality function $D\colon\ \Omega \times \hat{\Omega} \to \mathbb{R}$, denoted by $A \xrightarrow{D} B$, if $(AD(\cdot, y))(x) = (BD(x, \cdot))(y)$ for all $(x, y) \in \Omega \times \hat{\Omega}$. □

The following lemma intertwines the $\mathfrak{su}(2)$ and its dual algebra with a duality function.

Lemma 4.2 (Single-colony intertwiner) *For $\alpha \in \mathbb{N}$, let $d_\alpha\colon\ [\alpha] \times [\alpha] \to [0, 1]$ be the function defined by*

$$d_\alpha(x, n) = \frac{\binom{x}{n}}{\binom{\alpha}{n}} \mathbf{1}_{\{n \le x\}}. \tag{4.5}$$

Then, the following duality relations hold:

$$J^{\alpha,+} \xrightarrow{d_\alpha} A^{\alpha,+}, \quad J^{\alpha,-} \xrightarrow{d_\alpha} A^{\alpha,-}, \quad J^{\alpha,0} \xrightarrow{d_\alpha} A^{\alpha,0}. \tag{4.6}$$

Proof By straightforward calculations, it can be shown that $d_\alpha(x, n)$ satisfies the relations

$$\begin{aligned} (\alpha - x)\, d_\alpha(x+1, n) &= n\,[d_\alpha(x, n-1) - d_\alpha(x, n)] + (\alpha - n)\,[d_\alpha(x, n) - d_\alpha(x, n+1)], \\ x\, d_\alpha(x-1, n) &= (\alpha - n)\, d_\alpha(x, n), \\ x\, d_\alpha(x, n) &= (\alpha - n)\, d_\alpha(x, n+1) + n\, d_\alpha(x, n), \end{aligned} \tag{4.7}$$

from which the above dualities in (4.6) follow immediately. □

Remark 4.3 (Seed-bank and $\mathfrak{su}(2)$**-algebra)** The basic idea behind the algebraic approach to duality is to write the generator of a given process in terms of simple operators that form a representation of some known Lie algebra and to make an ansatz to obtain an intertwiner of the chosen representation. The intertwiner d_α in the above lemma was first identified in [12, Lemma 1] as a duality function in disguise for the classical duality between the Moran model and the block-counting process of Kingman's coalescent. Recently, in [4] this duality was put in the algebraic framework by deriving it from an intertwining via d_α of two representations of the *Heisenberg algebra* $\mathscr{H}(2)$. The connection of d_α to the $\mathfrak{su}(2)$-algebra was also made in [11, Section 3.2], where the authors obtained a self-duality function of $2j$-SEP factorized in terms of d_α by considering symmetries related to the $\mathfrak{su}(2)$-algebra. The relation of our seed-bank model to the $\mathfrak{su}(2)$-algebra becomes clear once we realize that the seed-bank component in our single-colony model is an *inhomogeneous* version of the $2j$-SEP on two sites. Thus, it is natural to expect that the classical duality of Moran model can be retrieved from representations of $\mathfrak{su}(2)$-algebra as well. The above lemma indeed

provides the ingredients to establish the duality of our single-colony model from representations of the $\mathfrak{su}(2)$-algebra. Although it is possible to guess the dual process of the single-colony model without going into the Lie-algebraic framework, the true usefulness of this approach lies in identifying the dual of the spatial model, where such speculation is no longer feasible. □

Proof of Theorem 3.2 Recall that both $Z = (X(t), Y(t))_{t\geq 0}$ and $Z^* = (n_t, m_t)_{t\geq 0}$ live on the state space $\Omega = [N] \times [M]$. Let $D\colon\ \Omega \times \Omega \to [0, 1]$ be the function defined by

$$D\big((X, Y); (n, m)\big) = \frac{\binom{X}{n}}{\binom{N}{n}} \frac{\binom{Y}{m}}{\binom{M}{m}} \mathbf{1}_{\{n\leq X, m\leq Y\}} = d_N(X, n) d_M(Y, m), \quad (X, Y), (n, m) \in \Omega. \tag{4.8}$$

Let $G = G_{\mathrm{Mor}} + G_{\mathrm{Exc}}$ be the generator of the process Z, where $G_{\mathrm{Mor}}, G_{\mathrm{Exc}}$ are as in (3.2)–(3.3). Also note from (3.7) that the generator $\widehat{G}$ of the dual process is given by $\widehat{G} = G_{\mathrm{King}} + G_{\mathrm{Exc}}$ where $G_{\mathrm{King}}\colon\ C(\Omega) \to C(\Omega)$ is defined as

$$(G_{\mathrm{King}} f)(n, m) = \frac{n(n-1)}{2N}[f(n-1, m) - f(n, m)], \quad (n, m) \in \Omega. \tag{4.9}$$

Since Ω is countable, it is enough to show the generator criterion for duality, i.e.

$$\big(GD(\cdot\,;\,(n, m))\big)(X, Y) = \big(\widehat{G} D((X, Y);\,\cdot)\big)(n, m), \quad (X, Y), (n, m) \in \Omega. \tag{4.10}$$

In our notation, (4.10) translates into $G \xrightarrow{D} \widehat{G}$. It is somewhat tedious to verify (4.10) by direct computation. Rather, we will write down a proof with the help of the elementary operators defined in (4.2). This approach will also reveal the underlying change of representation of the two operators $G, \widehat{G}$ that is embedded in the duality.

Note that

$$\begin{aligned}
G_{\mathrm{King}} &= \tfrac{1}{2N}\Big[(A_1^{N,+} - A_1^{N,-} + 2A_1^{N,0})A_1^{N,0} + \tfrac{N}{2}(A_1^{N,+} + A_1^{N,-} - N)\Big],\\
G_{\mathrm{Mor}} &= \tfrac{1}{2N}\Big[J_1^{N,0}(J_1^{N,+} - J_1^{N,-} + 2J_1^{N,0}) + \tfrac{N}{2}(J_1^{N,+} + J_1^{N,-} - N)\Big],\\
G_{\mathrm{Exc}} &= \tfrac{\lambda}{M}\Big[J_1^{N,+}J_2^{M,-} + J_1^{N,-}J_2^{M,+} + 2J_1^{N,0}J_2^{M,0} - \tfrac{NM}{2}\Big]\\
&= \tfrac{\lambda}{M}\Big[A_1^{N,+}A_2^{M,-} + A_1^{N,-}A_2^{M,+} + 2A_1^{N,0}A_2^{M,0} - \tfrac{NM}{2}\Big],
\end{aligned} \tag{4.11}$$

where the subscripts indicate which variable of the associated function the operators act on. For example, $J_1^{N,+}$ and $J_2^{M,+}$ act on the first and second variable, respectively. So, for a function $f\colon\ [N] \times [M] \to \mathbb{R}$, we have $(J_1^{N,+} f)(n, m) = (J^{N,+} f(\cdot\,;\,m))(n)$ and $(J_2^{M,+} f)(n, m) = (J^{M,+} f(n\,;\,\cdot))(m)$. The equivalent version of Lemma 4.2 holds for these operators with subscript as well, except that the duality function is

D. In other words, $J_1^{N,+} \xrightarrow{D} A_1^{N,+}$, $J_2^{M,+} \xrightarrow{D} A_2^{M,+}$, and so on. Using these duality relations and the representations in (4.11), we have $G_{\mathrm{Mor}} \xrightarrow{D} G_{\mathrm{King}}$ and $G_{\mathrm{Exc}} \xrightarrow{D} G_{\mathrm{Exc}}$, where we use:

- Two operators acting on different sites commute with each other.
- For some duality function d and operators $A, B, \hat{A}, \hat{B}$, if $A \xrightarrow{d} \hat{A}$, $B \xrightarrow{d} \hat{B}$, then, for any constants c_1, c_2, $AB \xrightarrow{d} \hat{B}\hat{A}$ and $c_1 A + c_2 B \xrightarrow{d} c_1 \hat{A} + c_2 \hat{B}$.

Since $G = G_{\mathrm{Mor}} + G_{\mathrm{Exc}}$ and $\widehat{G} = G_{\mathrm{King}} + G_{\mathrm{Exc}}$, we have $G \xrightarrow{D} \widehat{G}$, which proves the claim. □

4.2 Equilibrium

Proof *(Proof of Proposition 3.3.)* For $x \in \mathbb{R}$ and $r \in \mathbb{N}$, let $(x)_r$ be the *falling factorial* defined as

$$(x)_r = x(x-1) \times \cdots \times (x-r+1), \tag{4.12}$$

where we put $(x)_r = 1$ when $r = 0$. For any $n \in \mathbb{N}_0$, we can write x^n as

$$x^n = \sum_{j=0}^{n} c_{n,j} \, (x)_j, \tag{4.13}$$

where the constants $c_{n,j}$ (known as the Stirling numbers of the second kind) are unique and depend only on n and $j \in [n]$. Let $(n, m) \in \Omega = [N] \times [M]$ be such that $(n, m) \neq (0, 0)$, and let $(n_t, m_t)_{t \geq 0}$ be the dual process in Definition 3.7. It follows from (4.13) and Theorem 3.2 that

$$\begin{aligned}
&\lim_{t \to \infty} \mathbb{E}_{(X,Y)}[X(t)^n Y(t)^m] \\
&\quad = \sum_{i=0}^{n} \sum_{j=0}^{m} c_{n,i} c_{m,j} \lim_{t \to \infty} \mathbb{E}_{(X,Y)}[(X(t))_i (Y(t))_j] \\
&\quad = \sum_{i=0}^{n} \sum_{j=0}^{m} c_{n,i} c_{m,j} (N)_i (M)_j \lim_{t \to \infty} \mathbb{E}_{(X,Y)}[D((X(t), Y(t)); (i, j))] \\
&\quad = \sum_{i=0}^{n} \sum_{j=0}^{m} c_{n,i} c_{m,j} (N)_i (M)_j \lim_{t \to \infty} \mathbb{E}^{(i,j)}[D((X, Y); (n_t, m_t))],
\end{aligned} \tag{4.14}$$

where $D\colon \Omega \times \Omega \to [0, 1]$ is the duality function in Theorem 3.2, defined by

$$D((X, Y); (n, m)) = \frac{\binom{X}{n}}{\binom{N}{n}} \frac{\binom{Y}{m}}{\binom{M}{m}} \mathbf{1}_{\{n \leq X, m \leq Y\}} \equiv \frac{(X)_n (Y)_m}{(N)_n (M)_m}, \tag{4.15}$$

and the expectation in the last line of (4.14) is with respect to the dual process. Let T be the first time at which there is only one particle left in the dual, i.e. $T = \inf\{t > 0\colon\ n_t + m_t = 1\}$. Note that, for any initial state $(i, j) \in \Omega\backslash\{(0,0)\}$, $T < \infty$ with probability 1, and the distribution of (n_t, m_t) converges as $t \to \infty$ to the invariant distribution $\frac{N}{N+M}\delta_{(1,0)} + \frac{M}{N+M}\delta_{(0,1)}$. So, for any $(i, j) \in \Omega\backslash\{(0,0)\}$,

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}^{(i,j)}[D((X,Y);(n_t,m_t))] \\
&\quad = \lim_{t\to\infty} \mathbb{E}^{(i,j)}[D((X,Y);(n_t,m_t)) \mid T \le t]\, \mathbb{P}^{(i,j)}(T \le t) \\
&\qquad + \lim_{t\to\infty} \underbrace{\mathbb{E}^{(i,j)}[D((X,Y);(n_t,m_t)) \mid T > t]}_{\le 1}\, \mathbb{P}^{(i,j)}(T > t) \\
&\quad = \lim_{t\to\infty} \left[\tfrac{X}{N}\, \mathbb{P}^{(i,j)}(n_t = 1, m_t = 0) + \tfrac{Y}{M}\, \mathbb{P}^{(i,j)}(n_t = 0, m_t = 1)\right] \\
&\quad = \frac{X}{N}\frac{N}{N+M} + \frac{Y}{M}\frac{M}{N+M} = \frac{X+Y}{N+M},
\end{aligned} \tag{4.16}$$

where we use that the second term after the first equality converges to 0 because $T < \infty$ with probability 1. Combining (4.16) with (4.14), we get

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}_{(X,Y)}[X(t)^n Y(t)^m] \\
&\quad = \sum_{i=0}^{n}\sum_{j=0}^{m} c_{n,i}c_{m,j}(N)_i(M)_j \lim_{t\to\infty} \mathbb{E}^{(i,j)}[D((X,Y);(n_t,m_t))] \\
&\quad = \frac{X+Y}{N+M}\left(\sum_{i=0}^{n} c_{n,i}(N)_i\right)\left(\sum_{j=0}^{m} c_{m,j}(M)_j\right) + \left(1 - \tfrac{X+Y}{N+M}\right) c_{n,0}c_{m,0} \\
&\quad = N^n M^m \frac{X+Y}{N+M},
\end{aligned} \tag{4.17}$$

where the last equality follows from (4.13) and the fact that $c_{n,0}c_{m,0} = 0$ when $(n, m) \neq (0, 0)$. □

Proof *(Proof of Corollary 3.4.)* Note that the distribution of a two-dimensional random vector (Z_1, Z_2) taking values in $[N] \times [M]$ is determined by the mixed moments $\mathbb{E}[Z_1^i Z_2^j]$, $i, j \in [N] \times [M]$. For $i \in I = [NM]$, let $p_i = \mathbb{P}((Z_1, Z_2) = f^{-1}(i))$, where $f\colon\ [N] \times [M] \to I$ is a bijection. For $i \in I$, let $c_i = \mathbb{E}[Z_1^x Z_2^y]$, where $(x, y) = f^{-1}(i)$. We can write $\boldsymbol{c} = A\boldsymbol{p}$, where $\boldsymbol{p} = (p_i)_{i\in I}$, $\boldsymbol{c} = (c_i)_{i\in I}$ and A is an invertible $(N+1)(M+1) \times (N+1)(M+1)$ matrix. Hence, $\boldsymbol{p} = A^{-1}\boldsymbol{c}$ is uniquely determined by the mixed moments, and convergence of the mixed moments of $(X(t), Y(t))$ as shown in Proposition 3.3 is enough to conclude that $(X(t), Y(t))$ converges in distribution as $t \to \infty$ to a random vector (X_∞, Y_∞) taking values in $[N] \times [M]$. The distribution of (X_∞, Y_∞) is also uniquely determined and is given by $\frac{X+Y}{N+M}\delta_{(N,M)} + (1 - \frac{X+Y}{N+M})\delta_{(0,0)}$. □

5 Proofs: Duality and Well-Posedness for the Multi-colony Model

In Sect. 5.1, we give the proof of Lemma 3.8. In Sect. 5.2, we introduce equivalent versions for the multi-colony setting of the operators defined in (4.2) for the single-colony setting, and use these to prove Theorem 3.10 and Corollary 3.11. In Sect. 5.3, we prove Propositions 3.6, 3.12 and Theorem 3.13.

5.1 Proof of Lemma 3.8

Proof Note that the rate-matrix is nothing but the dual generator L_{dual} obtained from the rates specified in (3.23). The action of L_{dual} on a function $f\colon\ \mathcal{X}^* \to \mathbb{R}$ is given by

$$\begin{aligned}(L_{\text{dual}} f)(\xi) = &\sum_{i\in\mathbb{Z}^d} \left[\tfrac{n_i(n_i-1)}{2N_i} + n_i \sum_{\substack{j\in\mathbb{Z}^d,\\ j\neq i}} a(i,j)\tfrac{n_j}{N_j}\right]\big[f(\xi-\boldsymbol{\delta}_{i,A})-f(\xi)\big]\\ &+\sum_{i\in\mathbb{Z}^d}\lambda\, n_i\tfrac{(M_i-m_i)}{M_i}\big[f(\xi-\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{i,D})-f(\xi)\big]\\ &+\sum_{i\in\mathbb{Z}^d}\lambda(N_i-n_i)\tfrac{m_i}{M_i}\big[f(\xi+\boldsymbol{\delta}_{i,A}-\boldsymbol{\delta}_{i,D})-f(\xi)\big]\\ &+\sum_{i\in\mathbb{Z}^d}\sum_{\substack{j\in\mathbb{Z}^d\\ j\neq i}} a(i,j)n_i\tfrac{N_j-n_j}{N_j}\big[f(\xi-\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,A})-f(\xi)\big],\end{aligned} \tag{5.1}$$

where $\xi = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$ and the configurations $\boldsymbol{\delta}_{i,A}, \boldsymbol{\delta}_{i,D} \in \mathcal{X}^* \subset \mathcal{X}$ are as in (3.13). It is enough to show that L_{dual} satisfies the well-known Foster–Lyapunov criterion for stability (see, for example, [27, Theorem 2.1] or [6, Theorem (1.11)] for Markov processes on countable state spaces), i.e.

$$(L_{\text{dual}} V)(\xi) \le pV(\xi) \quad \forall \xi \in \mathcal{X}^*, \tag{5.2}$$

for some $p > 0$ with $V\colon\ \mathcal{X}^* \to (0,\infty)$ a function such that there exist $(E_k)_{k\in\mathbb{N}}$ with $E_k \uparrow \mathcal{X}^*$ and $\inf_{x\notin E_k} V(x) \to \infty$ as $k \to \infty$.

Let us define the function $V\ :\ \mathcal{X}^* \to (0,\infty)$ as

$$V((n_i,m_i)_{i\in\mathbb{Z}^d}) := 1 + \sum_{i\in\mathbb{Z}^d}(n_i+m_i), \quad (n_i,m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*, \tag{5.3}$$

and, for $k \in \mathbb{N}$, set

$$E_k := \Big\{(n_i,m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*\ :\ \sum_{i\in\mathbb{Z}^d} n_i + m_i \le k\Big\}. \tag{5.4}$$

Since $\mathcal{X}^*$ contains configurations with finitely many particles, V is well defined. It is straightforward to see that

$$E_k \uparrow \mathcal{X}^*, \quad \lim_{k\to\infty} \inf_{x\notin E_k} V(x) = \infty. \tag{5.5}$$

Let $\xi = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$ be arbitrary. Note that, for any $i, j \in \mathbb{Z}^d$ with $i \neq j$,

$$\begin{aligned} &[V(\xi - \boldsymbol{\delta}_{i,A}) - V(\xi)] = -\mathbf{1}_{\{n_i\geq 1\}}, && [V(\xi - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{i,D}) - V(\xi)]\, n_i(M_i - m_i) = 0,\\ &[V(\xi + \boldsymbol{\delta}_{i,A} - \boldsymbol{\delta}_{i,D}) - V(\xi)](N_i - n_i)m_i = 0, && [V(\xi - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{j,A}) - V(\xi)]\, n_i(N_j - n_j) = 0 \end{aligned} \tag{5.6}$$

and so by using (5.1) we obtain

$$\begin{aligned} |(L_{\text{dual}}V)(\xi)| &\leq \sum_{i\in\mathbb{Z}^d} \left[\tfrac{n_i(n_i-1)}{2N_i} + n_i \sum_{\substack{j\in\mathbb{Z}^d,\\ j\neq i}} a(i,j)\tfrac{n_j}{N_j} \right] |V(\xi - \boldsymbol{\delta}_{i,A}) - V(\xi)| \\ &\leq \sum_{i\in\mathbb{Z}^d} \left[\tfrac{n_i}{2} + n_i \sum_{j\in\mathbb{Z}^d} a(i,j) \right] \leq \max\{1,c\} \sum_{i\in\mathbb{Z}^d} n_i \leq \max\{1,c\} V(\xi), \end{aligned} \tag{5.7}$$

where $c = \sum_{i\in\mathbb{Z}^d} a(0,i) < \infty$. Hence, setting $p := \max\{1, c\} > 0$, we have that

$$(L_{\text{dual}}V)(\xi) \leq |(L_{\text{dual}}V)(\xi)| \leq p\, V(\xi), \tag{5.8}$$

which proves our the claim. □

5.2 Duality

5.2.1 Generators and Intertwiners

Let $f \in C(\mathcal{X})$ and $\eta = (X_i, Y_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}$, and let $\boldsymbol{\delta}_{i,A}, \boldsymbol{\delta}_{i,D}$ be as in (3.13). Define the action of the multi-colony operators as in Table 3.
The same duality relations as in Lemma 4.2 hold for these operators as well. The only difference is that the duality function becomes the site-wise product of the duality functions appearing in the single-colony model.

Lemma 5.1 (Multi-colony intertwiner) *Let* $D\colon \mathcal{X}\times\mathcal{X}^* \to [0,1]$ *be the function defined by*

$$D((X_k,Y_k)_{k\in\mathbb{Z}^d}; (n_k,m_k)_{k\in\mathbb{Z}^d}) = \prod_{i\in\mathbb{Z}^d} \frac{\binom{X_i}{n_i}}{\binom{N_i}{n_i}} \frac{\binom{Y_i}{m_i}}{\binom{M_i}{m_i}} \mathbf{1}_{\{n_i\leq X_i, m_i\leq Y_i\}}, \tag{5.9}$$

Table 3 Action of operators on $f \in C(\mathcal{X})$

Operators acting on variable $X_i, i \in \mathbb{Z}^d$	Operators acting on variable $Y_i, i \in \mathbb{Z}^d$
$J_{i,A}^{N_i,+} f(\eta) = (N_i - X_i) f(\eta + \boldsymbol{\delta}_{i,A})$	$J_{i,D}^{M_i,+} f(\eta) = (M_i - Y_i) f(\eta + \boldsymbol{\delta}_{i,D})$
$J_{i,A}^{N_i,-} f(\eta) = X_i f(\eta - \boldsymbol{\delta}_{i,A})$	$J_{i,D}^{M_i,-} f(\eta) = Y_i f(\eta - \boldsymbol{\delta}_{i,D})$
$J_{i,A}^{N_i,0} f(\eta) = (X_i - \frac{N_i}{2}) f(\eta)$	$J_{i,D}^{M_i,0} f(\eta) = (Y_i - \frac{M_i}{2}) f(\eta)$
$A_{i,A}^{N_i,+} = J_{i,A}^{N_i,-} - J_{i,A}^{N_i,+} - 2J_{i,A}^{N_i,0}$	$A_{i,D}^{M_i,+} = J_{i,D}^{M_i,-} - J_{i,D}^{M_i,+} - 2J_{i,D}^{M_i,0}$
$A_{i,A}^{N_i,-} = J_{i,A}^{N_i,+}$	$A_{i,D}^{M_i,-} = J_{i,D}^{M_i,+}$
$A_{i,A}^{N_i,0} = J_{i,A}^{N_i,+} + J_{i,A}^{N_i,0}$	$A_{i,D}^{M_i,0} = J_{i,D}^{M_i,+} + J_{i,D}^{M_i,0}$

where $(X_k, Y_k)_{k\in\mathbb{Z}^d} \in \mathcal{X}$ *and* $(n_k, m_k)_{k\in\mathbb{Z}^d} \in \mathcal{X}^*$. *Then, for every* $i \in \mathbb{Z}^d$ *and* $s \in \{0, +, -\}$,

$$J_{i,A}^{N_i,s} \xrightarrow{D} A_{i,A}^{N_i,s}, \quad J_{i,D}^{M_i,s} \xrightarrow{D} A_{i,D}^{M_i,s}. \tag{5.10}$$

Proof The proof is exactly the same as the proof of Lemma 4.2. □

Proposition 5.2 (Generator criterion) *Let* L *be the generator defined in* (3.16)*, and* $\hat{L}$ *the generator of the dual process defined in Definition* 3.7. *Furthermore, let* $D\colon \mathcal{X} \times \mathcal{X}^* \to [0, 1]$ *be the function defined in Lemma* 5.1. *Then,* $L \xrightarrow{D} \hat{L}$.

Proof Recall that $L = L_{\text{Mig}} + L_{\text{Res}} + L_{\text{Exc}}$, where $L_{\text{Mig}}, L_{\text{Res}}, L_{\text{Ex}}$ are defined in (3.17)–(3.19). In terms of the operators defined earlier, these have the following representations:

$$\begin{aligned}
L_{\text{Mig}} &= \sum_{i\in\mathbb{Z}^d} \sum_{\substack{j\in\mathbb{Z}^d \\ j\neq i}} \frac{a(i,j)}{N_j} \Big[\Big(J_{i,A}^{N_i,+} - J_{i,A}^{N_i,-} + 2J_{i,A}^{N_i,0} \Big) J_{j,A}^{N_j,0} \\
&\quad + \tfrac{N_j}{2} \Big(J_{i,A}^{N_i,+} + J_{i,A}^{N_i,-} - N_i \Big) \Big], \\
L_{\text{Res}} &= \sum_{i\in\mathbb{Z}^d} \frac{1}{2N_i} \Big[J_{i,A}^{N_i,0} \Big(J_{i,A}^{N_i,+} - J_{i,A}^{N_i,-} + 2J_{i,A}^{N_i,0} \Big) \\
&\quad + \tfrac{N_i}{2} \Big(J_{i,A}^{N_i,+} + J_{i,A}^{N_i,-} - N_i \Big) \Big], \\
L_{\text{Exc}} &= \sum_{i\in\mathbb{Z}^d} \frac{\lambda}{M_i} \Big[J_{i,A}^{N_i,+} J_{i,D}^{M_i,-} + J_{i,A}^{N_i,-} J_{i,D}^{M_i,+} + 2J_{i,A}^{N_i,0} J_{i,D}^{M_i,0} - \tfrac{N_i M_i}{2} \Big] \\
&= \sum_{i\in\mathbb{Z}^d} \frac{\lambda}{M_i} \Big[A_{i,A}^{N_i,+} A_{i,D}^{M_i,-} + A_{i,A}^{N_i,-} A_{i,D}^{M_i,+} + 2A_{i,A}^{N_i,0} A_{i,D}^{M_i,0} - \tfrac{N_i M_i}{2} \Big].
\end{aligned} \tag{5.11}$$

Similarly, the generator $\hat{L}$ of the dual process defined in Definition 3.7 acting on $f \in C(\mathcal{X}^*)$ is given by $\hat{L} = \hat{L}_{\text{Mig}} + L_{\text{Exc}} + L_{\text{King}}$, where

$$\begin{aligned}\hat{L}_{\mathrm{Mig}} f(\xi) &= \sum_{i\in\mathbb{Z}^d}\sum_{\substack{j\in\mathbb{Z}^d\\ j\neq i}} \frac{a(i,j)}{N_j}\Big\{n_i(N_j-n_j)[f(\xi-\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,A})-f(\xi)]\\ &\qquad + n_i n_j[f(\xi-\boldsymbol{\delta}_{i,A})-f(\xi)]\Big\},\\ L_{\mathrm{King}} f(\xi) &= \sum_{i\in\mathbb{Z}^d} \frac{n_i(n_i-1)}{2N_i}[f(\xi-\boldsymbol{\delta}_{i,A})+f(\xi+\boldsymbol{\delta}_{i,A})-2f(\xi)],\end{aligned} \tag{5.12}$$

for $\xi = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$. The representations of these operators are

$$\begin{aligned}\hat{L}_{\mathrm{Mig}} &= \sum_{i\in\mathbb{Z}^d}\sum_{\substack{j\in\mathbb{Z}^d\\ j\neq i}} \frac{a(i,j)}{N_j}\Big[A_{j,A}^{N_j,0}\left(A_{i,A}^{N_i,+}-A_{i,A}^{N_i,-}+2A_{i,A}^{N_i,0}\right)\\ &\qquad +\tfrac{N_j}{2}\left(A_{i,A}^{N_i,+}+A_{i,A}^{N_i,-}-N_i\right)\Big],\\ L_{\mathrm{King}} &= \sum_{i\in\mathbb{Z}^d} \frac{1}{2N_i}\Big[\left(A_{i,A}^{N_i,+}A_{i,A}^{N_i,-}+2A_{i,A}^{N_i,0}\right)A_{i,A}^{N_i,0}\\ &\qquad +\tfrac{N_i}{2}\left(A_{i,A}^{N_i,+}+A_{i,A}^{N_i,-}-N_i\right)\Big].\end{aligned} \tag{5.13}$$

From Lemma 5.1 and the representations in (5.11)–(5.13), we see that $L_{\mathrm{Mig}} \xrightarrow{D} \hat{L}_{\mathrm{Mig}}$, $L_{\mathrm{Res}} \xrightarrow{D} L_{\mathrm{King}}$ and $L_{\mathrm{Ex}} \xrightarrow{D} L_{\mathrm{Ex}}$, which yields $L \xrightarrow{D} \hat{L}$. □

As shown in [22, Proposition 1.2], the generator criterion is enough to get the required duality relation of Theorem 3.10 when both L and $\hat{L}$ are Markov generators of Feller processes. Since it is not a priori clear whether L (or its extension) is a Markov generator, we need to use [9, Theorem 4.11, Corollary 4.13].

5.2.2 Proof of Duality Relation

Proof of Theorem 3.10 We combine [9, Theorem 4.11 and Corollary 4.13] and reinterpret these in our context:

- Let $(\eta_t)_{t\geq 0}$ and $(\xi_t)_{t\geq 0}$ be two independent processes on E_1 and E_2 that are solutions to the martingale problem for $(L_1, \mathcal{D}_1)$ and $(L_2, \mathcal{D}_2)$ with initial states $x \in E_1$ and $y \in E_2$. Assume that $D\colon E_1 \times E_2 \to \mathbb{R}$ is such that $D(\cdot\,;\xi) \in \mathcal{D}_1$ for any $\xi \in E_2$ and $D(\eta\,;\cdot) \in \mathcal{D}_2$ for any $\eta \in E_1$. Also assume that for each $T > 0$, there exists an integrable random variable U_T such that

$$\begin{aligned}&\sup_{0\leq s,t\leq T} |D(\eta_t;\xi_s)| \leq U_T, \quad \sup_{0\leq s,t\leq T} |(L_1 D(\cdot\,;\xi_s))(\eta_t)| \leq U_T,\\ &\sup_{0\leq s,t\leq T} |(L_2 D(\eta_t;\,\cdot))(\xi_s)| \leq U_T.\end{aligned} \tag{5.14}$$

  If $(L_1 D(\cdot\,;y))(x) = (L_2 D(x\,;\cdot))(y)$, then $\mathbb{E}_x[D(\eta_t;y)] = \mathbb{E}^y[D(x,\xi_t)]$ for all $t \geq 0$.

To apply the above, pick $E_1 = \mathcal{X}$, $E_2 = \mathcal{X}^*$, $L_1 = L$, $L_2 = L_{\text{dual}}$, $\mathcal{D}_1 = \mathcal{D}$, $\mathcal{D}_2 = C(\mathcal{X}^*)$, where L_{dual} is the generator of the dual process Z^* and set D to be the function defined in Lemma 5.1. Note that since $\mathcal{D}$ contains local functions only, $D(\cdot\,;\xi) \in \mathcal{D}$ for any $\xi \in \mathcal{X}^*$ and, since $\mathcal{X}^*$ is countable, $D(\eta\,;\cdot) \in C(\mathcal{X}^*)$ for any $\eta \in \mathcal{X}$. Fix $x = (X_i, Y_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}$ and $y = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$. Note that by Proposition 5.2, $(L_1 D(\cdot\,; y))(x) = (L_2 D(x\,;\cdot))(y)$. Pick $(\xi_t)_{t\geq 0}$ to be the process Z^* with initial state y. Note that $(\xi_t)_{t\geq 0}$ is the unique solution to the martingale problem for $(L_{\text{dual}}, C(\mathcal{X}^*))$ with initial state y. Let $(\eta_t)_{t\geq 0}$ denote any solution Z to the martingale problem for $(L, \mathcal{D})$ with initial state x. Fix $T > 0$ and note that for $0 \leq s, t < T$,

$$
\begin{aligned}
(L_1 D(\cdot\,;\xi_s))(\eta_t) = {} & \sum_{i\in\mathbb{Z}^d} X_i(t) \left[\sum_{j\in\mathbb{Z}^d} a(i,j) \tfrac{N_j - X_j(t)}{N_j} \right] \big[D(\eta_t - \boldsymbol{\delta}_{i,A}; \xi_s) - D(\eta_t; \xi_s) \big] \\
& + \sum_{i\in\mathbb{Z}^d} (N_i - X_i(t)) \left[\sum_{j\in\mathbb{Z}^d} a(i,j) \tfrac{X_j(t)}{N_j} \right] \\
& \times \big[D(\eta_t + \boldsymbol{\delta}_{i,A}; \xi_s) - D(\eta_t; \xi_s) \big] \\
& + \sum_{i\in\mathbb{Z}^d} \lambda X_i(t) \tfrac{M_i - Y_i(t)}{M_i} \big[D(\eta_t - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{i,D}; \xi_s) - D(\eta_t; \xi_s) \big] \\
& + \sum_{i\in\mathbb{Z}^d} \lambda (N_i - X_i(t)) \tfrac{Y_i(t)}{M_i} \big[D(\eta_t + \boldsymbol{\delta}_{i,A} - \boldsymbol{\delta}_{i,D}; \xi_s) - D(\eta_t; \xi_s) \big]
\end{aligned}
\tag{5.15}
$$

and

$$
\begin{aligned}
(L_2 D(\eta_t\,;\cdot))(\xi_s) = {} & \sum_{i\in\mathbb{Z}^d} \left[\tfrac{n_i(s)(n_i(s)-1)}{2N_i} + n_i(s) \sum_{\substack{j\in\mathbb{Z}^d, \\ j\neq i}} a(i,j) \tfrac{n_j(s)}{N_j} \right] \\
& \times \big[D(\eta_t; \xi_s - \boldsymbol{\delta}_{i,A}) - D(\eta_t; \xi_s) \big] \\
& + \sum_{i\in\mathbb{Z}^d} \lambda\, n_i(s) \tfrac{M_i - m_i(s)}{M_i} \big[D(\eta_t; \xi_s - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{i,D}) - D(\eta_t; \xi_s) \big] \\
& + \sum_{i\in\mathbb{Z}^d} \lambda (N_i - n_i(s)) \tfrac{m_i(s)}{M_i} \big[D(\eta_t; \xi_s + \boldsymbol{\delta}_{i,A} - \boldsymbol{\delta}_{i,D}) - D(\eta_t; \xi_s) \big] \\
& + \sum_{i\in\mathbb{Z}^d} \sum_{\substack{j\in\mathbb{Z}^d \\ j\neq i}} a(i,j) n_i(s) \tfrac{N_j - n_j(s)}{N_j} \\
& \times \big[D(\eta_t; \xi_s - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{j,A}) - D(\eta_t; \xi_s) \big].
\end{aligned}
\tag{5.16}
$$

The random variable $\Gamma(t)$ defined in Theorem 3.10 is stochastically increasing in time t, and if we change the configuration η_t outside the box $[0, \Gamma(s)]^d \cap \mathbb{Z}^d$, then the value of $D(\eta_t; \xi_s)$ does not change. Consequently, all the summands in (5.15) for

$\|i\| > \Gamma(s), i \in \mathbb{Z}^d$, are 0, and since $\Gamma(s) \le \Gamma(T)$, we have the estimate

$$|(L_1 D(\cdot\,; \xi_s))(\eta_t)| \le 2(c+\lambda) \sum_{\substack{i\in\mathbb{Z}^d \\ \|i\|\le\Gamma(s)}} N_i \;\le\; 2(c+\lambda) \sum_{\substack{i\in\mathbb{Z}^d \\ \|i\|\le\Gamma(T)}} N_i, \tag{5.17}$$

where $c = \sum_{i\in\mathbb{Z}^d} a(0,i)$. Now, by Definition 3.7, the process $(\xi_t)_{t\ge 0}$ is the interacting particle system with coalescence in which the total number of particles can only decrease in time, and so $\sum_{i\in\mathbb{Z}^d}(n_i(s)+m_i(s)) \le N$, where $N = \sum_{i\in\mathbb{Z}^d}(n_i+m_i)$. Also, since $s \le T$, for $i \in \mathbb{Z}^d$ with $\|i\| > \Gamma(T)$ we have $n_i(s) = m_i(s) = 0$. Hence, from (5.16) we get

$$|(L_2 D(\eta_t\,; \cdot))(\xi_s)| \le 2(c+\lambda)N + 2\lambda \sum_{\substack{i\in\mathbb{Z}^d \\ \|i\|\le\Gamma(T)}} N_i. \tag{5.18}$$

Define the random variable U_T by

$$U_T = 1 + 2(c+\lambda)N + 2(c+\lambda) \sum_{\substack{i\in\mathbb{Z}^d \\ \|i\|\le\Gamma(T)}} N_i. \tag{5.19}$$

Then, combining (5.17)–(5.18) with the fact that the function D takes values in $[0, 1]$, we see that U_T satisfies all the conditions in (5.14), while assumption (3.25) in Theorem 3.10 ensures the integrability of U_T. □

5.2.3 Proof of Duality Criterion

Proof of Corollary 3.11 Let $\xi = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$ and $T > 0$ be fixed. By Theorem 3.10, it suffices to show that for any $(N_i)_{i\in\mathbb{Z}^d} \in \mathcal{N}$,

$$\sum_{i\in\mathbb{Z}^d} N_i\, \mathbb{P}^{\xi}(\Gamma(T) \ge \|i\|) < \infty, \tag{5.20}$$

where $\mathbb{P}^{\xi}$ is the law of the dual process Z^* started from initial state ξ. Let $n = \sum_{i\in\mathbb{Z}^d}(n_i+m_i)$ be the initial number of particles, and let $N(t)$ be the total number of migration events within the time interval $[0, t]$. We will construct a Poisson process N^* via coupling such that $N(t) \le N^*(t)$ for all $t \ge 0$ with probability 1. For this purpose, let us consider n independent particles performing a random walk on $\mathbb{Z}^d$ according to the migration kernel $a(\cdot,\cdot)$. For each $k = 1, \ldots, n$, let $\xi_k(t)$ and $\xi_k^*(t)$ denote the position of the k-th dependent and independent particle at time t, respectively. We take $\xi_k(0) = \xi_k^*(0)$ and couple each k-th interacting particle with the k-th independent particle as below:

- If the independent particle makes a jump from site $\xi_k^*(t)$ to $j^* \in \mathbb{Z}^d$, then the dependent particle jumps from $\xi_k(t)$ to $j = \xi_k(t) + (j^* - \xi_k^*(t))$ with probability $p_k(t)$ given by

$$p_k(t) = \begin{cases} 1 - \frac{n_j(t)}{N_j} & \text{if the dependent particle is in an } \textit{active} \text{ and } \textit{non-coalesced} \text{ state,} \\ 0 & \text{otherwise,} \end{cases} \tag{5.21}$$

  where $n_j(t)$ is the number of active particles at site j.
- The dependent particle does the other transitions (waking up, becoming dormant and coalescence) independently of the previous migration events, with the prescribed rates defined in Definition 3.7.

Note that since the migration kernel is translation invariant, under the above coupling the effective rate at which a dependent particle migrates from site i to j is $n_i a(i, j)(1 - \frac{n_j}{N_j})$ when there are n_i and n_j active particles at site i and j, respectively. Also, if $N_k(t)$ and $N_k^*(t)$ are the number of migration steps made within the time interval $[0, t]$ by the k-th dependent and independent particle, respectively, then under this coupling $N_k(t) \le N_k^*(t)$ with probability 1. Set $N^*(\cdot) = \sum_{k=1}^n N_k^*(\cdot)$. Then, clearly,

$$N(\cdot) = \sum_{k=1}^{n} N_k(\cdot) \le N^*(\cdot) \text{ with probability 1.} \tag{5.22}$$

Also, N^* is a Poisson process with intensity cn, since each independent particle migrates at a total rate c.

Let $Y_l, X_l \in \mathbb{Z}^d$ denote the step at the l-th migration event in the dependent and independent particle systems, respectively. Note that $(X_l)_{l \in \mathbb{N}}$ are i.i.d. with distribution $(a(0, i))_{i \in \mathbb{Z}^d}$. Since, under the above coupling, a dependent particle copies the step of an independent particle with a certain probability (possibly 0), and $\Gamma(0)$ is the minimum length of the box within which all n dependent particles at time 0 are located, we have, for any $t \ge 0$,

$$\Gamma(t) \le \Gamma(0) + \sum_{l=1}^{N(t)} |Y_l| \le \Gamma(0) + \sum_{l=1}^{N^*(t)} |X_l|. \tag{5.23}$$

Therefore,

$$\mathbb{P}^{\xi}(\Gamma(T) \ge k) \le \mathbb{P}\big(S_{N^*(T)} \ge k - \Gamma(0)\big) \qquad \forall\, k \ge 0, \tag{5.24}$$

where $S_{N^*(T)} = \sum_{l=1}^{N*(T)} |X_l|$.

To prove part (a), note that $\mathbb{E}[\mathrm{e}^{\delta S_{N^*(T)}}] < \infty$ and so, by Chebyshev's inequality,

$$\mathbb{P}(S_{N^*(T)} \ge x) = \mathbb{P}(\mathrm{e}^{\delta S_{N^*(T)}} \ge \mathrm{e}^{\delta x}) \le \mathbb{E}[\mathrm{e}^{\delta S_{N^*(T)}}]\, \mathrm{e}^{-\delta x}. \tag{5.25}$$

Thus, the inequality in (5.24) reduces to

$$\mathbb{P}^{\xi}(\Gamma(T) \geq k) \leq V \mathrm{e}^{-\delta k} \qquad \forall\, k \geq 0, \tag{5.26}$$

where

$$V = \mathbb{E}\left[\exp\{\delta\Gamma(0) + \delta S_{N^*(T)}\}\right] < \infty. \tag{5.27}$$

For $k \in \mathbb{N}$, let $\alpha_k = \#\{i \in \mathbb{Z}^d \colon \|i\|_\infty = k\}$. Then, $\alpha_k = (2k+1)^d - (2k-1)^d \leq 4^d k^{d-1}$. Hence,

$$\begin{aligned}\sum_{i \in \mathbb{Z}^d \setminus \{0\}} N_i\, \mathbb{P}^{\xi}(\Gamma(T) \geq \|i\|) &\leq \sum_{k \in \mathbb{N}} c_k \alpha_k\, \mathbb{P}^{\xi}(\Gamma(T) \geq k) \\ &\leq \sum_{k \in \mathbb{N}} c_k 4^d k^{d-1}\, \mathbb{P}^{\xi}(\Gamma(T) \geq k),\end{aligned} \tag{5.28}$$

where $c_k = \sup\{N_i \colon \|i\|_\infty = k, i \in \mathbb{Z}^d\}$. Since under the assumption of part (a), $\lim_{k\to\infty} \frac{1}{k} \log c_k = 0$, there exists a $K \in \mathbb{N}$ such that $c_k \leq \mathrm{e}^{\delta k/2}$ for all $k \geq K$. Hence, using (5.26), we find that

$$\sum_{i \in \mathbb{Z}^d} N_i\, \mathbb{P}^{\xi}(\Gamma(T) \geq \|i\|) \leq N_0 + \sum_{k=1}^{K-1} c_k \alpha_k + 4^d V \sum_{k=K}^{\infty} k^{d-1}\, \mathrm{e}^{-\delta k/2} < \infty, \tag{5.29}$$

which settles part (a).

To prove part (b), note that under the assumption $\sum_{i \in \mathbb{Z}^d} \|i\|^\gamma a(0,i) < \infty$ for some $\gamma > d + \delta$, we have $\mathbb{E}[S^\gamma_{N^*(T)}] < \infty$, and since $S_{N^*(T)}$ is a positive random variable, we get

$$\mathbb{P}(S_{N^*(T)} \geq x) \leq \mathbb{E}[S^\gamma_{N^*(T)}]\, x^{-\gamma}. \tag{5.30}$$

From (5.24) we get

$$\mathbb{P}^{\xi}(\Gamma(T) \geq k) \leq \frac{V}{(k - \Gamma(0))^\gamma} \qquad \forall\, k > \Gamma(0), \tag{5.31}$$

where $V = \mathbb{E}[S^\gamma_{N^*(T)}]$. By the assumption of part (b), there exists a $C > 0$ such that

$$c_k = \sup\{N_i \colon \|i\|_\infty = k, i \in \mathbb{Z}^d\} \leq C k^\delta \tag{5.32}$$

and so using (5.28), we obtain

$$\sum_{i\in\mathbb{Z}^d} N_i\,\mathbb{P}^{\xi}(\Gamma(T)\geq\|i\|)\leq N_0+\sum_{k\leq\Gamma(0)} c_k\alpha_k$$
$$+4^d CV\sum_{k>\Gamma(0)}\frac{k^{d+\delta-1}}{(k-\Gamma(0))^{\gamma}}<\infty, \qquad (5.33)$$

which settles part (b). □

5.3 Well-Posedness

In this section, we prove Propositions 3.6, 3.12 and Theorem 3.13.

5.3.1 Existence

Since the state space $\mathcal{X}$ is compact, the theory described in [25, Chapter I, Section 3] is applicable in our setting without any significant changes. The interacting particle systems in [25] have state space W^S, where W is a compact phase space and S is a countable site space. In our setting, the site space is $S=\mathbb{Z}^d$, but the phase space differs at each site, i.e. $[N_i]\times[M_i]$ at site $i\in\mathbb{Z}^d$. The general form of the generator of an interacting particle system in [25] is

$$(\Omega f)(\eta)=\sum_T\int_{W_T} c_T(\eta,\mathrm{d}\xi)[f(\eta^{\xi})-f(\eta)],\qquad \eta\in\mathcal{X}, \qquad (5.34)$$

where the sum is taken over all finite subsets T of S, and η^{ξ} is the configuration

$$\eta_i^{\xi}=\begin{cases}\xi_i & \text{if } i\in T,\\ \eta_i & \text{else.}\end{cases} \qquad (5.35)$$

For finite $T\Subset\mathcal{X}$, $c_T(\eta,d\xi)$ is a finite positive measure on $W_T=W^T$. To make the latter compatible with our setting, we define $W_T=\prod_{i\in T}[N_i]\times[M_i]$. The interpretation is that η is the current configuration of the system, $c_T(\eta,W_T)$ is the total rate at which a transition occurs involving *all* the coordinates in T, and $c_T(\eta,d\xi)/c_T(\eta,W_T)$ is the distribution of the restriction to T of the new configuration after that transition has taken place. Fix $\eta=(X_i,Y_i)_{i\in\mathbb{Z}^d}\in\mathcal{X}$. Comparing (5.34) with the formal generator L defined in (3.16), we see that the form of $c_T(\cdot,\cdot)$ is as follows:

- $c_T(\eta,d\xi)=0$ if $|T|\geq 2$.
- For $|T|=1$, let $T=\{i\}$ for some $i\in\mathbb{Z}^d$. Then, $c_T(\eta,\cdot)$ is the measure on $[N_i]\times[M_i]$ given by

$$c_T(\eta,\cdot)=X_i\left[\sum_{j\in\mathbb{Z}^d} a(i,j)\tfrac{N_j-X_j}{N_j}\right]\delta_{(X_i-1,Y_i)}(\cdot)+(N_i-X_i)$$

$$\times \left[\sum_{j \in \mathbb{Z}^d} a(i,j) \frac{X_j}{N_j} \right] \delta_{(X_i+1, Y_i)}(\cdot)$$
$$+\lambda X_i \frac{M_i - Y_i}{M_i} \delta_{(X_i-1, Y_i+1)}(\cdot) + \lambda (N_i - X_i) \frac{Y_i}{M_i} \delta_{(X_i+1, Y_i-1)}(\cdot). \tag{5.36}$$

Note that the total mass is

$$c_T(\eta, W_T) = X_i \left[\sum_{j \in \mathbb{Z}^d} a(i,j) \frac{N_j - X_j}{N_j} \right] + (N_i - X_i) \left[\sum_{j \in \mathbb{Z}^d} a(i,j) \frac{X_j}{N_j} \right]$$
$$+ \lambda X_i \frac{M_i - Y_i}{M_i} + \lambda (N_i - X_i) \frac{Y_i}{M_i}. \tag{5.37}$$

Lemma 5.3 (Bound on rates) *Let* $c = \sum_{i \in \mathbb{Z}^d} a(0,i) < \infty$. *For a finite set* $T \Subset \mathbb{Z}^d$, *let* $c_T = \sup_{\eta \in \mathcal{X}} c_T(\eta, W_T)$. *Then,* $c_T \le (c+\lambda) \mathbf{1}_{\{|T|=1\}} \sup_{i \in T} N_i$ *with* $c = \sum_{i \in \mathbb{Z}^d} a(0,i)$.

Proof Clearly, $c_T = 0$ if $|T| \ge 2$. So let $T = \{i\}$ for some $i \in \mathbb{Z}^d$. We see that for $\eta = (X_k, Y_k)_{k \in \mathbb{Z}^d}$, $c_T(\eta, W_T) \le cX_i + c(N_i - X_i) + \lambda X_i + \lambda(N_i - X_i) = (c+\lambda) N_i = (c+\lambda) \sup_{i \in T} N_i$. □

Proof of Proposition 3.6 By [25, Proposition 6.1 of Chapter I], it suffices to show that

$$\sum_{T \ni i} c_T < \infty \qquad \forall\, i \in S, \tag{5.38}$$

where the sum is taken over all finite subsets $T \Subset S$ containing $i \in S$. Since in our case $S = \mathbb{Z}^d$, we let $i \in \mathbb{Z}^d$ be fixed. By Lemma 5.3, the sum reduces to $c_{\{i\}}$, and clearly $c_{\{i\}} \le (c+\lambda) N_i < \infty$. □

Proof of Proposition 3.12 By [25, Proposition 6.1 and Theorem 6.7 of Chapter I], to show existence of solutions to the martingale problem for $(L, \mathcal{D})$, it is enough to prove that (5.38) is satisfied. But we already showed this in the proof of Proposition 3.6. □

5.3.2 Uniqueness

Before we turn to the proof of Theorem 3.13, we state and prove the following proposition, which, along with the duality established in Corollary 3.11, will play a key role in the proof of the uniqueness of solutions to the martingale problem.

Proposition 5.4 (Separation) *Let* $D\colon \mathcal{X} \times \mathcal{X}^* \to [0,1]$ *be the duality function defined in Lemma 5.1. Define the set of functions* $\mathcal{M} = \{D(\cdot\,; \xi)\colon \xi \in \mathcal{X}^*\}$. *Then,* $\mathcal{M}$ *is separating on the set of probability measures on* $\mathcal{X}$.

Proof Let $\mathbb{P}$ be a probability measure on $\mathcal{X} = \prod_{i\in\mathbb{Z}^d}[N_i]\times[M_i]$. It suffices to show that the finite-dimensional distributions of $\mathbb{P}$ are determined by $\{\int_{\mathcal{X}} f\,d\,\mathbb{P}\colon\ f\in\mathcal{M}\}$. Note that it is enough to show the following:

- Let $X=(X_1,X_2,\ldots,X_n)\in\prod_{i=1}^n[N_i]$ be an n-dimensional random vector with some distribution $\mathbb{P}_X$ on $\prod_{i=1}^n[N_i]$. Then, $\mathbb{P}_X$ is determined by the family

$$\mathcal{F}=\left\{\mathbb{E}\left[\prod_{i=1}^n\frac{\binom{X_i}{\alpha_i}}{\binom{N_i}{\alpha_i}}\right]\colon\ (\alpha_i)_{1\le i\le n}\in\prod_{i=1}^n[N_i]\right\}. \tag{5.39}$$

By (4.13), the family $\mathcal{F}$ is equivalent to the family

$$\mathcal{F}^*=\left\{\mathbb{E}\left[\prod_{i=1}^n X_i^{\alpha_i}\right]\colon\ (\alpha_i)_{1\le i\le n}\in\prod_{i=1}^n[N_i]\right\} \tag{5.40}$$

containing the mixed moments of $(X_1,\ldots,X_n)$. Since X takes a total of $N=\prod_{i=1}^n(N_i+1)$ many values, we can write the distribution $\mathbb{P}_X$ as the N-dimensional vector $\boldsymbol{p}=(p_1,\ldots,p_N)$, where $p_i=\mathbb{P}_X(X=f^{-1}(i))$ and $f\colon\ \prod_{i=1}^n[N_i]\to\{1,\ldots,N\}$ is the bijection defined by

$$f(x_1,x_2,\ldots,x_n)=\sum_{i=1}^{n-1}\left(\prod_{j=i+1}^n(N_j+1)\right)x_i+x_n+1,\quad (x_1,\ldots,x_n)\in\prod_{i=1}^n[N_i]. \tag{5.41}$$

Note that $\mathcal{F}^*$ also contains N elements, and so we can write $\mathcal{F}^*$ as the N-dimensional vector $\boldsymbol{e}=(e_1,\ldots,e_N)$, where $e_i=\mathbb{E}[\prod_{k=1}^n X_k^{\alpha_k}]$, $(\alpha_1,\ldots,\alpha_n)=f^{-1}(i)$. We show that there exists an invertible linear operator that maps $\boldsymbol{p}$ to $\boldsymbol{e}$. Indeed, for $i=1,\ldots,n$, define the $(N_i+1)\times(N_i+1)$ Vandermonde matrix A_i,

$$A_i=\begin{bmatrix}1 & 1 & 1 & \ldots & 1\\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_{N_i+1}\\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \ldots & \alpha_{N_i+1}^2\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \alpha_1^{N_i} & \alpha_2^{N_i} & \alpha_3^{N_i} & \ldots & \alpha_{N_i+1}^{N_i}\end{bmatrix},\qquad (\alpha_1,\alpha_2\ldots,\alpha_{N_i+1})=(0,1,\ldots,N_i). \tag{5.42}$$

Being Vandermonde matrices, all A_i are invertible. Finally, define the $N\times N$ matrix A by $A=A_1\otimes A_2\otimes\cdots\otimes A_n$, where $\otimes$ denotes the Kronecker product for matrices. Then, A is invertible because all A_i are. Also, we can check that $A\boldsymbol{p}=\boldsymbol{e}$, and hence, the distribution of X given by $\boldsymbol{p}=A^{-1}\boldsymbol{e}$ is uniquely determined by $\boldsymbol{e}$, i.e. the family $\mathcal{F}^*$. □

Proof of Theorem 3.13 We use [9, Proposition 4.7], which states the following (reinterpreted in our setting):

- Let $\mathcal{S}_1$ be compact and $\mathcal{S}_2$ be separable. Let $x \in \mathcal{S}_1, y \in \mathcal{S}_2$ be arbitrary and $D\colon\ \mathcal{S}_1 \times \mathcal{S}_2 \to \mathbb{R}$ be such that the set $\{D(\cdot\,; z)\colon\ z \in \mathcal{S}_2\}$ is separating on the set of probability measures on $\mathcal{S}_1$. Assume that for any two solutions $(\eta_t)_{t\geq 0}$ and $(\xi_t)_{t\geq 0}$ of the martingale problem for $(L_1, \mathcal{D}_1)$ and $(L_2, \mathcal{D}_2)$ with initial states x and y, the duality relation holds: $\mathbb{E}_x[D(\eta_t, y)] = \mathbb{E}^y[D(x, \xi_t)]$ for all $t \geq 0$. If for every $z \in \mathcal{S}_2$ there exists a solution to the martingale problem for $(L_2, \mathcal{D}_2)$ with initial state z, then for every $\eta \in \mathcal{S}_1$ uniqueness holds for the martingale problem for $(L_1, \mathcal{D}_1)$ with initial state η.

Pick $\mathcal{S}_1 = \mathcal{X}, \mathcal{S}_2 = \mathcal{X}^*$, $(L_1, \mathcal{D}_1) = (L, \mathcal{D})$ and $(L_2, \mathcal{D}_2) = (L_{\text{dual}}, C(\mathcal{X}^*))$, where L_{dual} is the generator of the dual process Z^*. Note that in our setting the martingale problem for $(L_{\text{dual}}, C(\mathcal{X}^*))$ is already well-posed (the unique solution is the dual process Z^* in Lemma 3.8). Hence, combining the above observations with Proposition 5.4 and Corollary 3.11, we get uniqueness of the solutions to the martingale problem for $(L, \mathcal{D})$ for every initial state $\eta \in \mathcal{X}$.

The second claim follows from [25, Theorem 6.8 of Chapter I]. □

6 Proofs: Equilibrium and Clustering Criterion

In Sect. 6.1, we prove Theorem 3.14 and Corollary 3.15. In Sect. 6.2, we derive expressions for the single-site genetic variability in terms of the dual process. In Sect. 6.3, we use one dual particle to write down expressions for first moments. In Sect. 6.4, we use two dual particles to write down expressions for second moments. In Sect. 6.5, we use these expressions to prove Theorem 3.17.

6.1 Convergence to Equilibrium

Proof of Theorem 3.14 Since the state space $\mathcal{X}$ is compact, the set of all probability measures on $\mathcal{X}$ is compact as well, by Prokhorov's theorem. It therefore suffices to prove convergence of the finite-dimensional distributions of $Z(t) = (X_i(t), Y_i(t))_{i\in\mathbb{Z}^d}$. Now recall from the proof of Proposition 5.4 that the distribution of an n-dimensional random vector $X(t) := (X_1(t), \ldots, X_n(t))$ taking values in $\prod_{l=1}^n [N_l]$ is determined by

$$\mathcal{F}_t = \left\{ \mathbb{E}\left[\prod_{l=1}^n \frac{\binom{X_l(t)}{\alpha_l}}{\binom{N_l}{\alpha_l}} \right] :\ (\alpha_l)_{1\leq l\leq n} \in \prod_{l=1}^n [N_l] \right\}. \tag{6.1}$$

In fact, the distribution of $X(t)$ converges if and only if $\mathbb{E}\left[\prod_{l=1}^n \binom{X_l(t)}{/} \alpha_l \binom{N_l}{\alpha_l} \right]$ converges for all $(\alpha_l)_{1\leq l\leq n} \in \prod_{l=1}^n [N_l]$ as $t \to \infty$. Since our duality function is given by

$$D((X_k, Y_k)_{k\in\mathbb{Z}^d}; (n_k, m_k)_{k\in\mathbb{Z}^d}) = \prod_{i\in\mathbb{Z}^d} \frac{\binom{X_i}{n_i}}{\binom{N_i}{n_i}} \frac{\binom{Y_i}{m_i}}{\binom{M_i}{m_i}} \mathbf{1}_{\{n_i\leq X_i, m_i\leq Y_i\}}, \tag{6.2}$$

it suffices to show that $\lim_{t\to\infty} \mathbb{E}_{\nu_\theta}[D(Z(t);\eta)]$ exists for all $\eta \in \mathcal{X}^*$. Let $\eta \in \mathcal{X}^*$ be fixed. By duality, we have

$$\begin{aligned}\mathbb{E}_{\nu_\theta}[D(Z(t);\eta)] &= \int_{\mathcal{X}} \mathbb{E}_\xi[D(Z(t);\eta)]\,\mathrm{d}\nu_\theta(\xi) = \int_{\mathcal{X}} \mathbb{E}^\eta[D(\xi;Z^*(t))]\,\mathrm{d}\nu_\theta(\xi)\\ &= \mathbb{E}^\eta\left[\int_{\mathcal{X}} D(\xi;Z^*(t))\,\mathrm{d}\nu_\theta(\xi)\right],\end{aligned} \tag{6.3}$$

where $\mathbb{E}_\xi$ denotes expectation w.r.t the law of $Z(t)$ started at configuration $\xi \in \mathcal{X}$, $Z^*(t) = (n_i(t), m_i(t))_{i\in\mathbb{Z}^d}$ is the dual process started at configuration η, and $\mathbb{E}^\eta$ denotes expectation w.r.t the law of the dual process. A simple calculation shows that if V is a random variable with distribution Binomial(N, p), then $\mathbb{E}\left[\binom{V}{n}/\binom{N}{n}\right] = p^n$ for $0 \le n \le N$. Since $(X_i(0), Y_i(0))_{i\in\mathbb{Z}^d}$ are all independent under ν_θ with Binomials as marginal distributions, we have

$$\mathbb{E}_{\nu_\theta}[D(Z(t);\eta)] = \mathbb{E}^\eta\left[\prod_{i\in\mathbb{Z}^d} \theta^{n_i(t)}\,\theta^{m_i(t)}\right] = \mathbb{E}^\eta[\theta^{|Z^*(t)|}], \tag{6.4}$$

where $|Z^*(t)| := \sum_{i\in\mathbb{Z}^d} n_i(t) + m_i(t)$ is total number of particles in the dual process at time t. Now, since the dual process is coalescing, $|Z^*(t)|$ is decreasing in t. Since $\theta \in [0,1]$, we see that $\mathbb{E}_{\nu_\theta}[D(Z(t);\eta)]$ is increasing in t. Thus, $\lim_{t\to\infty} \mathbb{E}_{\nu_\theta}[D(Z(t);\eta)]$ exists, which proves the existence of an equilibrium measure ν such that the distribution of $Z(t)$ weakly converges to ν. Also, by definition, $\mathbb{E}_\nu[D(Z(0);\eta)] = \lim_{t\to\infty} \mathbb{E}_{\nu_\theta}[D(Z(t);\eta)] = \lim_{t\to\infty} \mathbb{E}^\eta[\theta^{|Z^*(t)|}]$. □

Proof of Corollary 3.15 This follows by choosing $\eta = \delta_{i,A}$ and $\eta = \delta_{i,D}$ in the last part of Theorem 3.14 and noting that $\mathbb{E}^\eta[\theta^{|Z^*(t)|}] = \theta$ when $|\eta| = 1$. □

6.2 Genetic Variability

For $i, j \in \mathbb{Z}^d$ and $t \ge 0$, define

$$\Delta_{i,j}(t) = \Delta_{(i,A),(j,A)}(t) + \Delta_{(i,A),(j,D)}(t), \tag{6.5}$$

where

$$\Delta_{(i,A),(j,A)}(t) = \begin{cases} \frac{X_i(t)(N_j - X_j(t))}{N_i N_j} + \frac{X_j(t)(N_i - X_i(t))}{N_j N_i} & \text{if } i \ne j,\\ \frac{2X_i(t)(N_i - X_i(t))}{N_i(N_i-1)} & \text{if } i = j \text{ and } N_i \ne 1,\\ 0 & \text{otherwise,}\end{cases} \tag{6.6}$$

is the genetic variability (also frequently referred to as 'sample heterozygosity') at time t between the active populations of colony i and j, i.e. the probability that two

individuals drawn randomly from the two populations at time t are of different type, and

$$\Delta_{(i,A),(j,D)}(t) = \frac{X_i(t)(M_j - Y_j(t))}{N_i M_j} + \frac{(N_i - X_i(t))Y_j(t)}{N_i M_j} \tag{6.7}$$

is the genetic variability at time t between the active population of colony i and the dormant population of colony j. Note that the conditions in Definition 3.16 are equivalent to

$$\lim_{t\to\infty} \mathbb{E}[\Delta_{i,j}(t)] = 0 \qquad \forall\, i, j \in \mathbb{Z}^d, \tag{6.8}$$

where the expectation is taken conditional on an arbitrary initial condition $(X_i(0), Y_i(0))_{i\in\mathbb{Z}^d}$, which we suppress from the notation.

We use the dual process to compute $\mathbb{E}(\Delta_{(i,A),(j,A)}(t))$ and $\mathbb{E}(\Delta_{(i,A),(j,D)}(t))$, namely

$$\mathbb{E}(\Delta_{(i,A),(j,A)}(t)) = \begin{cases} \mathbb{E}\left[\frac{X_i(t)}{N_i}\right) + \mathbb{E}\left(\frac{X_j(t)}{N_j}\right] - 2\,\mathbb{E}\left[\frac{X_i(t)X_j(t)}{N_i N_j}\right] & \text{if } i \neq j, \\ 2\left(\mathbb{E}\left[\frac{X_i(t)}{N_i}\right] - \mathbb{E}\left[\frac{X_i(t)(X_i(t)-1)}{N_i(N_i-1)}\right]\right) & \text{otherwise,} \end{cases} \tag{6.9}$$

and

$$\mathbb{E}[\Delta_{(i,A),(j,D)}(t)] = \mathbb{E}\left[\frac{X_i(t)}{N_i}\right] + \mathbb{E}\left[\frac{Y_j(t)}{M_j}\right] - 2\,\mathbb{E}\left(\frac{X_i(t)Y_j(t)}{N_i M_j}\right). \tag{6.10}$$

Thus, in terms of the duality function D defined in Lemma 5.1,

$$\begin{aligned} \mathbb{E}[\Delta_{(i,A),(j,A)}(t)] = \mathbb{E}\Big[D(Z(t); \boldsymbol{\delta}_{i,A})\Big] + \mathbb{E}\Big[D(Z(t); \boldsymbol{\delta}_{j,A})\Big] \\ -2\,\mathbb{E}\Big[D(Z(t); \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{j,A})\Big], \end{aligned} \tag{6.11}$$

where $\boldsymbol{\delta}_{i,A}, \boldsymbol{\delta}_{j,A}$ are defined in (3.13). Similarly,

$$\begin{aligned} \mathbb{E}[\Delta_{(i,A),(j,D)}(t))] = \mathbb{E}\Big[D(Z(t); \boldsymbol{\delta}_{i,A})\Big] + \mathbb{E}\Big[D(Z(t); \boldsymbol{\delta}_{j,D})\Big] \\ -2\,\mathbb{E}\Big[D(Z(t); \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{j,D})\Big]. \end{aligned} \tag{6.12}$$

Since, by the duality relation in (3.26),

$$\mathbb{E}\Big[D(Z(t); Z^*(0))\Big] = \mathbb{E}\Big[D(Z(0); Z^*(t))\Big], \tag{6.13}$$

we have

$$\begin{aligned}
&\mathbb{E}^{\boldsymbol{\delta}_{i,A}}\Big[D(\eta_0;\xi_t)\Big] = \mathbb{E}\Big[\tfrac{X_i(t)}{N_i}\Big], \quad \mathbb{E}^{\boldsymbol{\delta}_{i,D}}\Big[D(\eta_0;\xi_t)\Big] = \mathbb{E}\Big[\tfrac{Y_i(t)}{M_i}\Big],\\
&\mathbb{E}^{\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,A}}\Big[D(\eta_0;\xi_t)\Big] = \begin{cases} \mathbb{E}\Big[\frac{X_i(t)(X_i(t)-1)}{N_i(N_i-1)}\Big] & \text{if } i=j,\\ \mathbb{E}\Big[\frac{X_i(t)X_j(t)}{N_iN_j}\Big] & \text{otherwise,}\end{cases}\\
&\mathbb{E}^{\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,D}}\Big(D(\eta_0;\xi_t)\Big) = \mathbb{E}\left(\tfrac{X_i(t)Y_j(t)}{N_iM_j}\right),
\end{aligned} \tag{6.14}$$

where $\eta_0 = Z^*(0)$ and the expectation in the left-hand side is taken with respect to the dual process $(\xi_t)_{t\geq 0} = Z^*$ defined in Definition 3.7. Combining the above with (6.11)–(6.12), we get

$$\begin{aligned}
\mathbb{E}[\Delta_{(i,A),(j,A)}(t)] = &\Big(\mathbb{E}^{\boldsymbol{\delta}_{i,A}}\Big[D(\eta_0;\xi_t)\Big] - \mathbb{E}^{\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,A}}\Big[D(\eta_0;\xi_t)\Big]\Big)\\
&+\Big(\mathbb{E}^{\boldsymbol{\delta}_{j,A}}\Big[D(\eta_0;\xi_t)\Big] - \mathbb{E}^{\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,A}}\Big[D(\eta_0;\xi_t)\Big]\Big)
\end{aligned} \tag{6.15}$$

and

$$\begin{aligned}
\mathbb{E}[\Delta_{(i,A),(j,D)}(t)] = &\Big(\mathbb{E}^{\boldsymbol{\delta}_{i,A}}\Big[D(\eta_0;\xi_t)\Big] - \mathbb{E}^{\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,D}}\Big[D(\eta_0;\xi_t)\Big]\Big)\\
&+\Big(\mathbb{E}^{\boldsymbol{\delta}_{j,D}}\Big[D(\eta_0;\xi_t)\Big] - \mathbb{E}^{\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{j,D}}\Big[D(\eta_0;\xi_t)\Big]\Big).
\end{aligned} \tag{6.16}$$

In Sects. 6.3–6.4, we derive expressions for the terms appearing in (6.15)–(6.16).

6.3 Dual: Single Particle

We saw earlier that in order to compute the first moment of $X_i(t)$ and $Y_i(t)$, we need to put a single particle at site i in the active and the dormant state as initial configurations, respectively. This motivates us to analyse the dual process when it starts with a single particle. The generator L_{dual} of the dual process can be written as

$$L_{\text{dual}} = L_{\text{Coal}} + L_{AD} + L_{DA} + L_{\text{Mig}}, \tag{6.17}$$

where

$$\begin{aligned}
(L_{\text{Coal}}f)(\xi) = &\sum_{i\in\mathbb{Z}^d} \frac{n_i(n_i-1)}{2N_i}[f(\xi-\boldsymbol{\delta}_{i,A}) - f(\xi)]\\
&+\sum_{i\in\mathbb{Z}^d}\sum_{\substack{j\in\mathbb{Z}^d\\ j\neq i}} \frac{a(i,j)}{N_j} n_i n_j [f(\xi-\boldsymbol{\delta}_{i,A}) - f(\xi)],
\end{aligned} \tag{6.18}$$

$$(L_{AD}f)(\xi) = \sum_{i\in\mathbb{Z}^d} \frac{\lambda\, n_i(M_i-m_i)}{M_i}[f(\xi-\boldsymbol{\delta}_{i,A}+\boldsymbol{\delta}_{i,D}) - f(\xi)], \tag{6.19}$$

$$(L_{DA}f)(\xi) = \sum_{i\in\mathbb{Z}^d} \frac{\lambda\, m_i(N_i - n_i)}{M_i}[f(\xi + \boldsymbol{\delta}_{i,A} - \boldsymbol{\delta}_{i,D}) - f(\xi)], \tag{6.20}$$

$$(L_{\mathrm{Mig}}f)(\xi) = \sum_{i\in\mathbb{Z}^d} \sum_{\substack{j\in\mathbb{Z}^d \\ j\neq i}} \frac{a(i,j)}{N_j} n_i(N_j - n_j)[f(\xi - \boldsymbol{\delta}_{i,A} + \boldsymbol{\delta}_{j,A}) - f(\xi)], \tag{6.21}$$

for $f \in C(\mathcal{X}^*)$ and $\xi = (n_i, m_i)_{i\in\mathbb{Z}^d} \in \mathcal{X}^*$.

When there is a single particle in the system at time 0, and consequently at any later time, the only parts of the generator that are nonzero are L_{AD}, L_{DA} and L_{Mig}. Here, L_{AD} turns an active particle at site i into a dormant particle at site i at rate λ, L_{DA} turns a dormant particle at site i into an active particle at site i at rate λK_i, with $K_i = \frac{N_i}{M_i}$, while L_{Mig} moves an active particle at site i to site $j \neq i$ at rate $a(i,j)$. Let us denote the state of the particle at time t by $\xi(t) \in \mathbb{Z}^d \times \{A, D\}$, where the first coordinate of $\xi(t)$ is the location of the particle and the second coordinate indicates whether the particle is active (A) or dormant (D). Let $\mathbb{P}^\xi$ be the law of the process $(\xi(t))_{t\geq 0}$ with initial state ξ.

Lemma 6.1 (First moments)

$$\begin{aligned}
\mathbb{E}\left[\frac{X_i(t)}{N_i}\right] &= \sum_{k\in\mathbb{Z}^d} \frac{X_k(0)}{N_k}\,\mathbb{P}^{(i,A)}(\xi(t) = (k,A)) + \frac{Y_k(0)}{M_k}\mathbb{P}^{(i,A)}(\xi(t) = (k,D)), \\
\mathbb{E}\left[\frac{Y_i(t)}{M_i}\right] &= \sum_{k\in\mathbb{Z}^d} \frac{X_k(0)}{N_k}\,\mathbb{P}^{(i,D)}(\xi(t) = (k,A)) + \frac{Y_k(0)}{M_k}\mathbb{P}^{(i,D)}(\xi(t) = (k,D)).
\end{aligned} \tag{6.22}$$

Proof *(Proof.)* Recall that, via the duality relation,

$$\mathbb{E}\left[\frac{X_i(t)}{N_i}\right] = \mathbb{E}^{\boldsymbol{\delta}_{i,A}}\left[\prod_{k\in\mathbb{Z}^d} \frac{\binom{X_k(0)}{n_k(t)}}{\binom{N_k}{n_k(t)}} \frac{\binom{Y_k(0)}{m_k(t)}}{\binom{M_k}{m_k(t)}} \mathbf{1}_{\{n_k(t)\leq X_k(0), m_k(t)\leq Y_k(0)\}}\right], \tag{6.23}$$

where the expectation in the right-hand side is taken with respect to the dual process with initial state $\boldsymbol{\delta}_{i,A}$ (a single active particle at site i), which has law $\mathbb{P}_{(i,A)}$. Since the term inside the expectation is equal to $\frac{X_k(0)}{N_k}$ or $\frac{Y_k(0)}{M_k}$, depending on whether $\xi(t) = (k,A)$ or $\xi(t) = (k,D)$, the claim follows immediately. The same argument holds for $\mathbb{E}[\frac{Y_i(t)}{M_i}]$ with initial condition (i,D) in the dual process. □

6.4 Dual: Two Particles

We need to find expressions for the second moments appearing in (6.9)–(6.10) in order to fully specify $\mathbb{E}(\Delta_{(i,A),(j,A)}(t))$ and $\mathbb{E}(\Delta_{(i,A),(j,D)}(t))$. This requires us to analyse the dual process starting from two particles. Unlike for the single-particle system, now all parts of the generator L_{dual} (see (6.17)) are nonzero, until the two particles coalesce into a single particle. The two particles repel each other: one particle discourages the other particle to come to the same location. The rates in the two-particle system are:

- **(Migration)** An active particle at site i migrates to site j at rate $a(i, j)$ if there is no active particle at site j; otherwise, at rate $a(i, j)(1 - \frac{1}{N_j})$.
- **($\mathbf{A} \to \mathbf{D}$)** An active particle at site i becomes dormant at site i at rate λ if there is no dormant particle at site i; otherwise, at rate $\lambda(1 - \frac{1}{M_i})$.
- **($\mathbf{D} \to \mathbf{A}$)** A dormant particle at site i becomes active at site i at rate λK_i if there is no active particle at site i; otherwise, at rate $\lambda(K_i - \frac{1}{M_i})$.
- **(Coalescence)** An active particle at site i coalesces with another active particle at site j at rate $\frac{1}{N_i}$ when $j = i$; otherwise, at rate $\frac{a(i,j)}{N_j}$.

Note that after coalescence has taken place, there is only one particle left in the system, which evolves as the single-particle system.

Let $(\xi_1(t), \xi_2(t), c(t)) \in \mathbb{S} = \mathcal{S}^* \times \mathcal{S}^* \times \{0, 1\}$ be the configuration of the two-particle system at time t, where $\mathcal{S}^* = \mathbb{Z}^d \times \{A, D\}$. Here, $\xi_1(t)$ and $\xi_2(t)$ represent the location and state of the two particles. The variable $c(t)$ takes value 1 if the two particles have coalesced into a single particle by time t, and 0 otherwise. It is necessary to add the extra variable $c(t)$ to the configuration in order to make the process Markovian (the rates depend on whether there are one or two particles in the system). To avoid triviality we assume that $c(0) = 0$ with probability 1, i.e. two particles at time 0 are always in a non-coalesced state. We denote the law of the process $(\xi_1(t), \xi_2(t), c(t))_{t\geq 0}$ by $\mathbb{P}^\xi$, where the initial condition is $\xi \in \mathcal{S}^* \times \mathcal{S}^*$. It is to be noted that since the number of active and dormant particles at a site i at any time are limited by N_i and M_i, respectively, the two-particle system is not defined whenever it is started from an initial configuration violating the maximal occupancy of the associated sites. Let τ be the first time at which the coalescence event has occurred, i.e.

$$\tau = \inf\{t \geq 0 \colon\ c(t) = 1\}. \tag{6.24}$$

Note that, conditional on $\tau < t$, $\xi_1(s) = \xi_2(s)$ for all $s \geq t$ with probability 1. Define

$$M_{(i,\alpha),(j,\beta)}(t) = \begin{cases} \frac{X_i(t)(X_i(t)-1)}{N_i(N_i-1)} & \text{if } i = j \text{ and } \alpha = \beta = A, \\ \frac{X_i(t)X_j(t)}{N_iN_j} & \text{if } i \neq j \text{ and } \alpha = \beta = A, \\ \frac{Y_i(t)(Y_i(t)-1)}{M_i(M_i-1)} & \text{if } i = j \text{ and } \alpha = \beta = D, \\ \frac{Y_i(t)Y_j(t)}{M_iM_j} & \text{if } i \neq j \text{ and } \alpha = \beta = D, \\ \frac{X_i(t)Y_j(t)}{N_iM_j} & \text{if } \alpha = A \text{ and } \beta = D, \\ \frac{Y_i(t)X_j(t)}{M_iN_j} & \text{otherwise,} \end{cases} \tag{6.25}$$

where $i, j \in \mathbb{Z}^d$ and $\alpha, \beta \in \{A, D\}$. To avoid ambiguity, we set $M_{(i,\alpha),(j,\beta)}(\cdot) = 0$ when $((i, \alpha), (j, \beta))$ is not a valid initial condition for the two-particle system.

Lemma 6.2 (Second moments) *For every valid initial condition* $((i,\alpha),(j,\beta)) \in (\mathbb{Z}^d \times \{A, D\})^2$ *of the two-particle system,*

$$\begin{aligned}\mathbb{E}\left[M_{(i,\alpha),(j,\beta)}(t)\right] &= Q((i,\alpha),(j,\beta),t) \\ &\quad + \sum_{k\in\mathbb{Z}^d} \frac{X_k(0)}{N_k}\,\mathbb{P}^{((i,\alpha),(j,\beta))}\big(\xi_1(t) = (k,A), \tau < t\big) \\ &\quad + \sum_{k\in\mathbb{Z}^d} \frac{Y_k(0)}{M_k}\,\mathbb{P}^{((i,\alpha),(j,\beta))}\big(\xi_1(t) = (k,D), \tau < t\big),\end{aligned} \tag{6.26}$$

where

$$\begin{aligned}&Q((i,\alpha),(j,\beta),t) \\ &= \sum_{k\in\mathbb{Z}^d} \frac{X_k(0)(X_k(0)-1)}{N_k(N_k-1)}\,\mathbb{P}^{((i,\alpha),(j,\beta))}(\xi_1(t) = \xi_2(t) = (k,A), \tau \geq t) \\ &\quad + \sum_{\substack{k,l\in\mathbb{Z}^d \\ k\neq l}} \frac{X_k(0)X_l(0)}{N_kN_l}\,\mathbb{P}^{((i,\alpha),(j,\beta))}(\xi_1(t) = (k,A), \xi_2(t) = (l,A), \tau \geq t) \\ &\quad + \sum_{k,l\in\mathbb{Z}^d} \frac{X_k(0)Y_l(0)}{N_kM_l}\,\mathbb{P}^{((i,\alpha),(j,\beta))}(\xi_1(t) = (k,A), \xi_2(t) = (l,D), \tau \geq t) \\ &\quad + \sum_{k\in\mathbb{Z}^d} \frac{Y_k(0)(Y_k(0)-1)}{M_k(M_k-1)}\,\mathbb{P}^{((i,\alpha),(j,\beta))}(\xi_1(t) = \xi_2(t) = (k,D), \tau \geq t) \\ &\quad + \sum_{\substack{k,l\in\mathbb{Z}^d \\ k\neq l}} \frac{Y_k(0)Y_l(0)}{M_kM_l}\,\mathbb{P}^{((i,\alpha),(j,\beta))}(\xi_1(t) = (k,D), \xi_2(t) = (l,D), \tau \geq t).\end{aligned} \tag{6.27}$$

Proof Note that $M_{(i,\alpha),(j,\beta)}(t) = D(Z(t); \boldsymbol{\delta}_{i,\alpha}+\boldsymbol{\delta}_{j,\beta})$, where D is the duality function. So, via the duality relation, we have

$$\mathbb{E}\left[M_{(i,\alpha),(j,\beta)}(t)\right] = \mathbb{E}^{\boldsymbol{\delta}_{i,\alpha}+\boldsymbol{\delta}_{j,\beta}}\left[\prod_{k\in\mathbb{Z}^d} \frac{\binom{X_k(0)}{n_k(t)}}{\binom{N_k}{n_k(t)}} \frac{\binom{Y_k(0)}{m_k(t)}}{\binom{M_k}{m_k(t)}} \mathbf{1}_{\{n_k(t)\leq X_k(0), m_k(t)\leq Y_k(0)\}}\right], \tag{6.28}$$

where the expectation in the right-hand side is taken with respect to the dual process when the initial condition has one particle at site i with state α and one particle at site j with state β, which has law $\mathbb{P}^{((i,\alpha),(j,\beta))}$. Depending on the configuration of the process at time t, the right-hand side of (6.28) equals the desired expression. □

The following lemma provides a nice comparison between the one-particle and two-particle system.

Lemma 6.3 (Correlation inequality) *Let* $(\xi(t))_{t\geq 0}$ *and* $(\xi_1(t), \xi_2(t), c(t))_{t\geq 0}$ *be the processes defined in Sects. 6.3 and 6.4, respectively, and* τ *the first time of coalescence defined in* (6.24)*. Then, for any valid initial condition* $((i,\alpha),(j,\beta)) \in (\mathbb{Z}^d \times \{A, D\})^2$ *of the two-particle system and any* $(k,\gamma) \in \mathbb{Z}^d \times \{A, D\}$,

$$\mathbb{P}^{(i,\alpha)}(\xi(t) = (k,\gamma)) \geq \mathbb{P}^{((i,\alpha),(j,\beta))}(\xi_1(t) = (k,\gamma), \tau < t). \tag{6.29}$$

Proof Let $\alpha = A$ and $i, j, k \in \mathbb{Z}^d$ be fixed. Let $\eta = Z(0)$ be the initial configuration defined as

$$(X_n(0), Y_n(0)) = \begin{cases} (N_k, 0) & \text{if } n = k \text{ and } \gamma = A, \\ (0, M_k) & \text{if } n = k \text{ and } \gamma = D, \\ (0,0) & \text{otherwise,} \end{cases} \qquad \forall\, n \in \mathbb{Z}^d. \tag{6.30}$$

Combining Lemmas 6.1 and 6.2, we get

$$\begin{aligned} &\mathbb{E}_\eta\left[\tfrac{X_i(t)}{N_i} - M_{(i,A),(j,\beta)}(t)\right] \\ &= \sum_{n\in\mathbb{Z}^d} \frac{X_n(0)}{N_n}\left[\mathbb{P}^{(i,A)}(\xi(t) = (n,A)) - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t) = (n,A), \tau < t)\right] \\ &\quad + \sum_{n\in\mathbb{Z}^d} \frac{Y_n(0)}{M_n}\left[\mathbb{P}^{(i,A)}(\xi(t)=(n,D)) - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t)=(n,D), \tau < t)\right] \\ &\quad - Q((i,A),(j,\beta),t) \\ &= \left[\mathbb{P}^{(i,A)}(\xi(t)=(k,\gamma)) - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t)=(k,\gamma), \tau<t)\right] - Q((i,A),(j,\beta),t). \end{aligned} \tag{6.31}$$

Since $Q((i,A),(j,\beta),t) \geq 0$ and the left-hand quantity is positive, we get

$$\mathbb{P}^{(i,A)}(\xi(t) = (k,\gamma)) \geq \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t) = (k,\gamma), \tau < t). \tag{6.32}$$

Replacing the left quantity in (6.31) with $\mathbb{E}_\eta[\frac{Y_i(t)}{M_i} - M_{(i,D),(j,\beta)}(t)]$ and using the same arguments, we see that the inequality for $\alpha = D$ follows. □

6.5 Proof of Clustering Criterion

Proof of Theorem 3.17 "$\Longleftarrow$" First we show that, if $((i,A),(j,\beta)) \in (\mathbb{Z}^d \times \{A, D\})^2$ is a valid initial condition for the two-particle system, then

$$\lim_{t\to\infty} \mathbb{E}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(j,\beta)}(t)\right] = 0, \quad \lim_{t\to\infty} \mathbb{E}\left[\frac{Y_j(t)}{M_j} - M_{(i,A),(j,\beta)}(t)\right] = 0. \tag{6.33}$$

Combining Lemmas 6.1 and 6.2, we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(j,\beta)}(t)\right] \\
&= \sum_{k\in\mathbb{Z}^d} \frac{X_k(0)}{N_k}\left[\mathbb{P}^{(i,A)}(\xi(t) = (k, A)) - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t) = (k, A), \tau < t)\right] \\
&\quad + \sum_{k\in\mathbb{Z}^d} \frac{Y_k(0)}{M_k}\left[\mathbb{P}^{(i,A)}(\xi(t) = (k, D)) - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t) = (k, D), \tau < t)\right] \\
&\quad - Q((i, A), (j, \beta), t).
\end{aligned} \tag{6.34}$$

Using Lemma 6.3 and the fact that $Q((i, A), (j, \beta), t) \geq 0$, we have the following:

$$\begin{aligned}
\mathbb{E}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(j,\alpha)}(t)\right] &\leq \sum_{\substack{S\in\{A,D\}\\ k\in\mathbb{Z}^d}} \left|\mathbb{P}^{(i,A)}(\xi(t) = (k, S))\right. \\
&\quad \left. - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t) = (k, S), \tau < t)\right| \\
&= \sum_{\substack{S\in\{A,D\}\\ k\in\mathbb{Z}^d}} \left[\mathbb{P}^{(i,A)}(\xi(t) = (k, S))\right. \\
&\quad \left. - \mathbb{P}^{((i,A),(j,\beta))}(\xi_1(t) = (k, S), \tau < t)\right] \\
&= 1 - \mathbb{P}^{((i,A),(j,\beta))}(\tau < t) = \mathbb{P}^{((i,A),(j,\beta))}(\tau \geq t).
\end{aligned} \tag{6.35}$$

Since, by assumption, $\tau < \infty$ with probability 1 irrespective of the initial configuration of the two-particle system, and since the left-hand quantity is positive, we have $\mathbb{E}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(j,\beta)}(t)\right] \to 0$ as $t \to \infty$. By a similar argument the other part of (6.33) is proved as well.

If $((i, A), (j, A))$ is a valid initial condition for the two-particle system, then by using (6.15)–(6.16) and (6.33), we have

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}\Big(\Delta_{(i,A),(j,A)}(t)\Big) &= \lim_{t\to\infty} \mathbb{E}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(j,A)}(t)\right] \\
&\quad + \lim_{t\to\infty} \mathbb{E}\left[\frac{X_j(t)}{N_j} - M_{(j,A),(i,A)}(t)\right] = 0.
\end{aligned} \tag{6.36}$$

If $((i, A), (j, A))$ is not a valid initial condition, then we must have that $i = j$ and $N_i = 1$, and so $\Delta_{(i,A),(j,A)}(t) = 0$ by definition. Thus, for any $i, j \in \mathbb{Z}^d$,

$$\lim_{t\to\infty} \mathbb{E}\Big[\Delta_{(i,A),(j,A)}(t)\Big] = 0. \tag{6.37}$$

Since $((i, A), (j, D))$ is always a valid initial condition for the two-particle system, we also have

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}\Big[\Delta_{(i,A),(j,D)}(t)\Big] &= \lim_{t\to\infty} \mathbb{E}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(j,D)}(t)\right] \\
&\quad + \lim_{t\to\infty} \mathbb{E}\left[\frac{Y_j(t)}{M_j} - M_{(i,A),(j,D)}(t)\right] = 0,
\end{aligned} \tag{6.38}$$

and hence from (6.5) we have that for any $i, j \in \mathbb{Z}^d$, $\mathbb{E}(\Delta_{i,j}(t)) \to 0$ as $t \to \infty$, which proves the claim.

"$\Longrightarrow$" Suppose that the system clusters for any initial configuration $Z(0) \in \mathcal{X}$. Then, by dominated convergence, the system clusters for any initial distribution of $Z(0)$ as well. Fix $\theta \in (0, 1)$, and let the distribution of $Z(0)$ be ν_θ, where

$$\nu_\theta = \bigotimes_{i \in \mathbb{Z}^d} (\text{Binomial}(N_i, \theta) \otimes \text{Binomial}(M_i, \theta)). \tag{6.39}$$

We will prove via contradiction that in the dual two particles with arbitrary valid initial states coalesce with probability 1, i.e. $\tau < \infty$ with probability 1. Indeed, suppose that this is not true, i.e. for some valid initial configuration $(\xi_1, \xi_2) \in \mathcal{S}^* \times \mathcal{S}^*$ of the two-particle system we have $\mathbb{P}^{(\xi_1,\xi_2)}(\tau = \infty) > 0$, where $\mathcal{S}^* = \mathbb{Z}^d \times \{A, D\}$. Since the dual process with two particles is irreducible (any valid configuration is accessible), we have $\mathbb{P}^{\xi}(\tau = \infty) > 0$ for any valid initial condition $\xi \in \mathcal{S}^* \times \mathcal{S}^*$. Let $\rho := \mathbb{P}^{((i,A),(i,D))}(\tau = \infty) > 0$, where $i \in \mathbb{Z}^d$ is fixed. Note that $((i, A), (i, D))$ is always a valid initial condition for the two-particle system, since $N_i, M_i \geq 1$. Let $\mathbb{P}^{(i,A)}$ be the law of the single-particle process $(\xi(t))_{t \geq 0}$ started with initial condition (i, A).

Since, by assumption, $\mathbb{E}_{\nu_\theta}\left[\Delta_{(i,A),(i,D)}(t)\right] \to 0$ as $t \to \infty$, we must have

$$\lim_{t \to \infty} \mathbb{E}_{\nu_\theta}\left[\frac{X_i(t)(M_i - Y_i(t))}{N_i M_i}\right] = 0. \tag{6.40}$$

Since $((i, A), (i, D))$ is a valid initial condition for the two-particle system, by using (6.34) with ν_θ as initial distribution we get

$$\begin{aligned}
\mathbb{E}_{\nu_\theta}\left[\frac{X_i(t)(M_i - Y_i(t))}{N_i M_i}\right] &= \mathbb{E}_{\nu_\theta}\left[\frac{X_i(t)}{N_i} - M_{(i,A),(i,D)}(t)\right] \\
&= \sum_{n \in \mathbb{Z}^d} \mathbb{E}_{\nu_\theta}\left[\frac{X_n(0)}{N_n}\right]\Big[\mathbb{P}^{(i,A)}(\xi(t) = (n, A)) \\
&\quad - \mathbb{P}^{((i,A),(i,D))}(\xi_1(t) = (n, A), \tau < t)\Big] \\
&\quad + \sum_{n \in \mathbb{Z}^d} \mathbb{E}_{\nu_\theta}\left[\frac{Y_n(0)}{M_n}\right]\Big[\mathbb{P}^{(i,A)}(\xi(t) = (n, D)) \\
&\quad - \mathbb{P}^{((i,A),(i,D))}(\xi_1(t) = (n, D), \tau < t)\Big] \\
&\quad - \mathbb{E}_{\nu_\theta}\big[Q((i, A), (i, D), t)\big],
\end{aligned} \tag{6.41}$$

where $Q(\cdot, \cdot, \cdot)$ is defined as in Lemma (6.2). Since, under ν_θ, $(X_n(0))_{n \in \mathbb{Z}^d}$, $(Y_n(0))_{n \in \mathbb{Z}^d}$ are all independent of each other and $X_n(0)$ and $Y_n(0)$ have distribu-

tions Binomial(N_n, θ) and Binomial(M_n, θ), respectively, we have

$$\begin{aligned}
&\mathbb{E}_{\nu_\theta}\left[\frac{X_n(0)}{N_n}\right] = \mathbb{E}_{\nu_\theta}\left[\frac{Y_n(0)}{M_n}\right] = \theta,\\
&\mathbb{E}_{\nu_\theta}\left[\frac{X_n(0)(X_n(0)-1)}{N_n(N_n-1)}\right] = \theta^2 \qquad \text{if } N_n \neq 1,\\
&\mathbb{E}_{\nu_\theta}\left[\frac{Y_n(0)(Y_n(0)-1)}{M_n(M_n-1)}\right] = \theta^2 \qquad \text{if } M_n \neq 1.
\end{aligned} \tag{6.42}$$

Hence, $\mathbb{E}_{\nu_\theta}[Q((i, A), (i, D), t)] = \theta^2\, \mathbb{P}^{((i,A),(i,D))}(\tau \geq t)$, and thus, (6.41) reduces to

$$\begin{aligned}
\mathbb{E}_{\nu_\theta}\left[\tfrac{X_i(t)(M_i - Y_i(t))}{N_i M_i}\right] &= \theta\left[1 - \mathbb{P}^{((i,A),(i,D))}(\tau < t)\right] - \theta^2\, \mathbb{P}^{((i,A),(i,D))}(\tau \geq t)\\
&= \theta(1-\theta)\, \mathbb{P}^{((i,A),(i,D))}(\tau \geq t).
\end{aligned} \tag{6.43}$$

By (6.40), the left-hand side of (6.43) tends to 0 as $t \to \infty$. Because $\theta \in (0, 1)$, we have

$$\rho = \lim_{t\to\infty} \mathbb{P}^{((i,A),(i,D))}(\tau \geq t) = 0, \tag{6.44}$$

which is a contradiction. □

References

1. Blath, J., Casanova, A.G., Kurt, N., Wilke-Berenguer, M.: A new coalescent for seed-bank models. Ann. Appl. Probab. **26**(2), 857–891 (2016). https://doi.org/10.1214/15-aap1106
2. Blath, J., Eldon, B., Casanova, A.G., Kurt, N.: Genealogy of a wright-fisher model with strong seed bank component. In: XI Symposium on Probability and Stochastic Processes, pp. 81–100. Springer International Publishing (2015). https://doi.org/10.1007/978-3-319-13984-5_4
3. Blath, J., Kurt, N.: Population genetic models of dormancy. In: Probabilistic Structures in Evolution, pp. 247–265. European Mathematical Society Publishing House (2021). https://doi.org/10.4171/ecr/17-1/12
4. Carinci, G., Giardinà, C., Giberti, C., Redig, F.: Dualities in population genetics: a fresh look with new dualities. Stoch. Proc. Appl. **125**, 941–969 (2015). https://doi.org/10.1016/j.spa.2014.10.009
5. Carinci, G., Giardinà, C., Redig, F., Sasamoto, T.: A generalized asymmetric exclusion process with $u_q(\mathfrak{sl}_2)$ stochastic duality. Probab. Theory Relat. Fields **166**(3), 887–933 (2016). https://doi.org/10.1007/s00440-015-0674-0
6. Chen, M.F.: On three classical problems for markov chains with continuous time parameters. J. Appl. Probab. **28**(2), 305–320 (1991). https://doi.org/10.2307/3214868

7. Donnelly, P., Kurtz, T.G.: A countable representation of the Fleming–Viot measure-valued diffusion. Ann. Probab. **24**(2), 698–742 (1996). https://doi.org/10.1214/aop/1039639359
8. Etheridge, A.M., Griffiths, R.C.: A coalescent dual process in a Moran model with genic selection. Theor. Popul. Biol. **75**(4), 320–330 (2009). https://doi.org/10.1016/j.tpb.2009.03.004
9. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence, vol. 282. Wiley, New York (2009)
10. Franceschini, C., Giardinà, C., Groenevelt, W.: Self-duality of Markov processes and intertwining functions. J. Math. Phys. Anal. Geom. **21**, 29 (2018). https://doi.org/10.1007/s11040-018-9289-x
11. Giardinà, C., Kurchan, J., Redig, F., Vafayi, K.: Duality and hidden symmetries in interacting particle systems. J. Stat. Phys. **135**, 25–55 (2009). https://doi.org/10.1007/s10955-009-9716-2
12. Gladstien, K.: The characteristic values and vectors for a class of stochastic matrices arising in genetics. SIAM J. Appl. Math. **34**(4), 630–642 (1978). https://doi.org/10.1137/0134050
13. Greven, A., den Hollander, F., Oomen, M.: Spatial populations with seed-bank: renormalisation on the hierarchical group (2020) (in preparation)
14. Greven, A., den Hollander, F., Oomen, M.: Spatial populations with seed-bank: well-posedness, duality and equilibrium (2020). Preprint arXiv:2004.14137
15. Greven, A., den Hollander, F., Oomen, M.: Spatial populations with seed-bank: finite-systems scheme (2021) (in preparation)
16. Greven, A., Limic, V., Winter, A.: Representation theorems for interacting Moran models, interacting Fisher-Wright diffusions and applications. Electron. J. Probab. **10**, 1286–1358 (2005). https://doi.org/10.1214/EJP.v10-290
17. Groenevelt, W.: Orthogonal stochastic duality functions from lie algebra representations. J. Stat. Phys. **174**(1), 97–119 (2018). https://doi.org/10.1007/s10955-018-2178-7
18. den Hollander, F., Nandan, S.: Spatially inhomogeneous populations with seed-banks: II. clustering regime, and cluster growth versus equilibria (2021) (in preparation)
19. den Hollander, F., Pederzani, G.: Multi-colony Wright-Fisher with seed-bank. Indag. Math. **28**(3), 637–669 (2017). https://doi.org/10.1016/j.indag.2017.02.002
20. Hutzenthaler, M., Alkemper, R.: Graphical representation of some duality relations in stochastic population models. Electron. Commun. Probab. **12**, 206–220 (2007). https://doi.org/10.1214/ecp.v12-1283
21. Jansen, S., Kurt, N.: Graphical representation of certain moment dualities and application to population models with balancing selection. Electron. Commun. Probab. **18**, 1–15 (2013). https://doi.org/10.1214/ecp.v18-2194
22. Jansen, S., Kurt, N.: On the notion(s) of duality for Markov processes. Probab. Surv. **11**, 59–120 (2014). https://doi.org/10.1214/12-ps206
23. Lennon, J.T., den Hollander, F., Wilke-Berenguer, M., Blath, J.: Principles of seed banks and the complexity emerging from dormancy (2020). Preprint arXiv:2012.00072
24. Lennon, J.T., Jones, S.E.: Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat. Rev. Microbiol. **9**, 119–130 (2011). https://doi.org/10.1038/nrmicro2504
25. Liggett, T.M.: Interacting Particle Systems. Springer, Berlin (1985)
26. Liggett, T.M.: Stochastic Interacting Systems: Contact. Voter and Exclusion Processes. Springer, Berlin Heidelberg (1999). https://doi.org/10.1007/978-3-662-03990-8
27. Meyn, S.P., Tweedie, R.L.: Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. Adv. Appl. Probab. **25**(3), 518–548 (1993). https://doi.org/10.2307/1427522
28. Redig, F., Sau, F.: Factorized duality, stationary product measures and generating functions. J. Stat. Phys. **172**(4), 980–1008 (2018). https://doi.org/10.1007/s10955-018-2090-1
29. Shoemaker, W.R., Lennon, J.T.: Evolution with a seed bank: the population genetic consequences of microbial dormancy. Evol. Appl. **11**(1), 60–75 (2018). https://doi.org/10.1111/eva.12557
30. Sturm, A., Swart, J.M., Völlering, F.: The algebraic approach to duality: an introduction. In: Genealogies of Interacting Particle Systems, pp. 81–150. World Scientific (2020). https://doi.org/10.1142/9789811206092_0003
31. Vitalis, R., Glémin, S., Olivieri, I.: When genes go to sleep: the population genetic consequences of seed dormancy and monocarpic perenniality. Am. Nat. **163**(2), 295–311 (2004). https://doi.org/10.1086/381041