

Giulia Pederzani

The Seed-Bank Coalescent

Master thesis

Thesis advisor:
Prof. Dr. W.Th.F. den Hollander

Date master exam:



Mathematisch Instituut, Universiteit Leiden

Contents

1	Introduction	3
2	The Standard Wright-Fisher model and the Kingman coalescent	4
2.1	Definition of the model and Wright-Fisher diffusion limit	4
2.2	Dual process	5
2.3	Kingman coalescent	5
3	The Wright-Fisher model with Geometric Seed-Bank Component	7
3.1	Definition of the model	7
3.2	Scaling and diffusion limit	8
4	The Dual of the Seed-Bank Frequency Process	10
5	The Seed-Bank Coalescent	11
5.1	Definition of the seed-bank coalescent	11
5.2	Properties of the seed-bank coalescent	13
6	A Multi-colony Extension	18
6.1	Migration	18
6.2	Migration and mutation with infinitely many alleles	18
6.3	Fourier analysis	23
6.4	Special choice of parameters	26
6.5	Perturbation	29
6.6	Asymptotic analysis	32
6.7	Perturbation under special choice of parameters	33

1 Introduction

The Wright-Fisher model with seed-bank component aims at modelling the genetic evolution of populations in which individuals are allowed to achieve a dormant state and maintain it for arbitrarily many generations, where dormancy is understood as a reversible rest period, characterized by low metabolic activity and interruption of phenotypic development (see Lennon and Jones [7]). This type of behaviour is observed in many taxa, including plants, bacteria and other micro-organisms, as a typical response to unfavourable environmental conditions. Dormant individuals can be resuscitated, after a varying and possibly substantial number of generations, under more favourable conditions and reprise reproduction. This strategy has been shown to have important implications in population persistence, maintenance of genetic variability and even stability of ecosystem processes, acting as a buffer against such evolutionary forces as genetic drift, selection and environmental variability. The importance of this evolutionary trait has therefore led to attempts at modeling seed-banks from a mathematical perspective.

The mathematical modeling of seed-bank effects has been somewhat challenging. In fact, since individuals can remain dormant for arbitrarily many generations, it is problematic to retain the Markov property. The most successful models to date are extensions of the classical Wright-Fisher model, which is the standard starting point for modeling the genetic evolution of a population from a probabilistic perspective. So far, however, these extensions have either been able to model only weak seed-bank effects, resulting in a delayed Wright-Fisher model, or have led to extreme or elusive behaviours that seem artificial.

The most relevant model was proposed by Kaj, Krone and Lascoux [8], who allowed individuals in a population of fixed size to select a parent from a random number B of generations in the past, where B is an $\mathbb{N} \setminus \{0\}$ -valued random variable assumed to be independent and identically distributed for each individual. The authors prove that, if B can assume a finite number of values $\{1, 2, \dots, m\}$, then, after the usual scaling of time by the population size, the model converges to a delayed Kingman coalescent, where the rates are multiplied by $1/\mathbb{E}[B]^2$. More generally, it was proven by Blath, Casanova, Kurt and Spanò (see [9]) that a sufficient condition for convergence to the Kingman coalescent is $\mathbb{E}[B] < \infty$. This model, however, leads only to a delay in the coalescence process and leaves the coalescent structure unchanged, thus not capturing the power and importance of seed-banks in maintaining genetic variability. We therefore speak of “weak” seed-bank effects.

A further extension of the model, which allows for “strong” seed-bank effects, was proposed in [9], where B is not assumed to be bounded. In particular, the authors studied the case of a heavy-tailed age-distribution, namely, $\mu(\{n, n+1, \dots\}) = L(n) n^{-\alpha}$, where L is a slowly varying function as $n \rightarrow \infty$. It is proved that for $\alpha > 1/2$ the most recent common ancestor (MRCA) of two randomly sampled individuals always exists, but the expected time to the MRCA is infinite if $0 < \alpha < 1$, whereas for $0 < \alpha < 1/2$ a MRCA may not exist at all with positive probability.

Such extreme behaviour may seem artificial and for this reason González Casanova et al. in [10] and Blath et al. in [11] have studied the case in which B scales with the total population size N . In particular the age-distribution μ of B is given by $\mu = (1 - \epsilon) \delta_1 + \epsilon \delta_{N^\beta}$, $\beta > 0$, $\epsilon \in (0, 1)$. For $\beta < 1/3$, this model again shows convergence to the Kingman coalescent, but requires rescaling by the non-classical factor $N^{1+2\beta}$, thus highly increasing the expected time to the MRCA. However, it leaves the coalescent structure unchanged, and results in other parameter regimes remain elusive.

Thus, the mathematical results to date have been effective in modeling weak seed-bank effects, but for stronger components they have so far been unsatisfying in suggesting new limiting coalescent structures. The main advantage of the model proposed and studied by Blath et al. in [1], which we expose and extend here, is that it retains the Markov property and the long-term behaviour of the system is in line with our intuition without the need for artificial scaling assumptions. It also offers a natural interpretation for the scaling limit of both the forward and the backward process, and the limiting genealogy is given by a new coalescent structure.

The remaining sections are organized as follows: in Section 2 we present the standard Wright-Fisher model and discuss its key properties; in Section 3 we define the seed-bank model as presented in [1]; in Section 4 we introduce its dual process, and in Section 5 we define and look at the properties of the seed-bank coalescent. Finally, in Section 6, we present an extension of the model that includes a spatial element: individuals not only reproduce but also migrate.

2 The Standard Wright-Fisher model and the Kingman coalescent

In this section we present the standard Wright-Fisher model, the simplest model for population genetics where the only evolutionary force at play is resampling. The model studied in [1] is an extension of the Wright-Fisher model. Concerning notation, we will denote by \mathbb{N}_0 the set $\{0, 1, 2, \dots\}$ and by \mathbb{N} the set $\{1, 2, 3, \dots\} = \mathbb{N}_0 \setminus \{0\}$.

2.1 Definition of the model and Wright-Fisher diffusion limit

Consider a population of N haploid individuals, with N fixed. For each genetic locus, an individual carries one copy of the gene at that locus and we assume the gene to be one of two types, denoted by A and a . The model is discrete in time. At each time unit, every individual chooses another individual from the population (possibly itself) uniformly at random and adopts its type. This choice is independent of time and of the choices of the other individuals. This type of resampling mechanism is called parallel updating (all individuals choose an ancestor at the same time and independently of each other). Let

$$X_n = \text{number of individuals of type } A \text{ at time } n. \quad (2.1)$$

The sequence $X = (X_n)_{n \in \mathbb{N}_0}$ is the discrete-time Markov chain with state space $\Omega = \{0, 1, 2, \dots, N\}$ and transition kernel

$$p_{ij} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(\frac{N-i}{N}\right)^{N-j}, \quad i, j \in \Omega. \quad (2.2)$$

This formula follows from the fact that, if at time n there are i individuals of type A , then there will be j individuals at time $n+1$ if and only if precisely j individuals choose an ancestor of type A and $N-j$ individuals choose an ancestor of type a . This occurs with the probabilities given by the second and third factor. The first factor expresses the number of ways in which the j individuals can be chosen from the whole population. The initial condition can be any state $X_0 \in \Omega$.

Notice that the states 0 and N are traps, namely, $p_{00} = p_{NN} = 1$, which corresponds to the extinction of one of the two genetic types. Since all other states communicate, the process eventually gets trapped and therefore genetic variability is lost through chance.

It is of interest to consider a space-time rescaling that converges to a limiting process. Let $Y_t^{(N)}$ be the continuous-time version of X where all individuals resample at rate 1 and the number of individuals of type A is divided by N so that the state space is now $\{0, \frac{1}{N}, \frac{2}{N}, \dots, 1\} \subset [0, 1]$. Thus $Y_t^{(N)} = \frac{1}{N} X_{\lceil Nt \rceil}$ corresponds to the fraction of individuals of type A at time t on time scale N .

Theorem 2.1. *In the limit as $N \rightarrow \infty$, if the initial condition scales properly, namely,*

$$w - \lim_{N \rightarrow \infty} Y_0^{(N)} = Y_0, \quad (2.3)$$

then the whole process scales properly, namely,

$$w - \lim_{N \rightarrow \infty} (Y_t^{(N)})_{t \geq 0} = (Y_t)_{t \geq 0}, \quad (2.4)$$

where $Y = (Y_t)_{t \geq 0}$ is the diffusion process on $[0, 1]$ satisfying the stochastic differential equation

$$dY_t = \sqrt{Y_t(1-Y_t)} dW_t, \quad (2.5)$$

where $(W_t)_{t \geq 0}$ is standard Brownian motion. This SDE has a unique strong solution, i.e., there is a unique path $t \mapsto Y_t$ measurable with respect to the canonical filtration associated with Brownian motion.

Notice that in the above statements $w - \lim$ stands for weak limit, namely, convergence in distribution on path space.

The process defined by (2.5) is called the *Wright-Fisher diffusion* and can be thought of as a standard Brownian motion whose local diffusion rate is given by a *diffusion function* $g : [0, 1] \rightarrow [0, \infty)$. In our case $g(y) = y(1-y)$. Notice that $g(0) = g(1) = 0$, i.e., the process stops at states 0 or 1, in accordance with what happens in discrete time.

2.2 Dual process

The Wright-Fisher diffusion given by (2.5) does not give us an explicit formula for Y_t as a function of $(W_s)_{s \in [0, t]}$, but it is easier to manipulate than the original Markov chain. To see why, we present in this section a process that is dual to Y .

Theorem 2.2. *Let $D = (D_t)_{t \geq 0}$ be the death process on \mathbb{N} where transitions $n \rightarrow n - 1$ occur at rate $\binom{n}{2}$. Then*

$$\mathbb{E}[Y_t^n | Y_0 = y] = \mathbb{E}[y^{D_t} | D_0 = n] \quad \forall y \in [0, 1], \quad n \in \mathbb{N}, \quad t \geq 0. \quad (2.6)$$

Because of this duality, we can compute the moments of Y_t given the distribution of D_t and therefore reconstruct the whole distribution of Y_t . Since D is a much simpler process than Y , this is a considerable advantage.

The dual process D is closely related to the backward coalescing random walk obtained by tracing back the lineages of a sample of individuals drawn from the population. In fact, assume $Y_0 = y$ and draw a sample of n individuals from the population at time t . Then the left-hand side of (2.6) is the probability that all the individuals in the sample have type A . This is equivalent to the probability that the ancestors of the n individuals in the sample are of type A , and this probability is given by the right hand side of (2.6), where the number of different ancestors is D_t .

In general, we can think of D_t as

$$D_t = \text{number of ancestral lineages at time } t_0 - t, \quad t \in [0, t_0], \quad (2.7)$$

where $t_0 \gg 1$ is any observation time.

2.3 Kingman coalescent

In order to understand the genealogy of the population, we introduce a coalescent process called the *Kingman coalescent*. This process takes values in the set of partitions of $\{1, 2, \dots, n\}$ and is such that, if there are j sets in the partition the process is currently in, then two sets in the partition merge at rate $\binom{j}{2}$. This process describes the family tree of a large Wright-Fisher population backwards in time, up until the individual called *most recent common ancestor*, from which all individuals descend. As we will make more clear later on, the process is related to the Wright-Fisher diffusion Y and, through duality, to the death process D .

In this section we first construct the coalescent and then state one of its key features, the fact that it comes down from infinity.

Construction of the process In order to construct the Kingman coalescent, we need to begin with the n -coalescent. For $n \in \mathbb{N}$, denote by E_n the set of partitions of $\{1, 2, \dots, n\}$. For a partition $R \in E_n$, $|R|$ denotes the number of sets in the partition. The n -coalescent is the continuous-time Markov process $R^n = (R_t^n)_{t \geq 0}$ with state space E_n , initial state

$$R_0^n = \Delta = (\{1\}, \{2\}, \dots, \{n\}), \quad (2.8)$$

and transition rates

$$p_{RS} = \begin{cases} 1, & \text{if } S \triangleleft R \\ 0, & \text{otherwise,} \end{cases} \quad R, S \in E_n, \quad R \neq S, \quad (2.9)$$

where

$$S \triangleleft R \iff S \in E_n \text{ and } |S| = |R| - 1, \quad (2.10)$$

where R and S have $|R| - 2$ elements in common and the remaining element of S is obtained by merging the remaining two elements of R .

The relationship with the ancestral tree of the genealogy of the population is straightforward. Consider a sample of n individuals at time t_0 and let R_t^n be the partition defined by the condition that two individuals i and j are in the same subset if and only if they have a common ancestor alive at time $t_0 - t$. With this description, the process $R^n = (R_t^n)_{t \geq 0}$ has the structure of an n -coalescent and there is a one-to-one correspondence between the subsets of R^n and the ancestors in the genealogy.

We will now explore the relationship between R^n and $D^n = (D_t^n)_{t \geq 0}$, where the latter is the natural restriction to $\{1, 2, \dots, n\}$ of the death process $D = (D_t)_{t \geq 0}$ defined in Section 2.2. From (2.7) it is clear that

$$D_t^n = |R_t^n|, \quad (2.11)$$

namely, D_t^n is the number of ancestral lineages in the coalescent tree at time $t_0 - t$ or, equivalently, the number of subsets in the partition R_t^n . In fact, the merger of two subsets of R_t^n corresponds to the extinction of one of the lineages from the coalescent tree. From the transition rates of D , we obtain that the total transition rate of R^n is given by

$$p_R = \sum_{S \in E_n} p_{RS} = \lambda_{|R|}, \quad (2.12)$$

where $\lambda_k = \binom{k}{2}$. Therefore the n -coalescent jumps through a sequence of partitions \mathfrak{R}_k with $|\mathfrak{R}_k| = k$ for all $k = 1, \dots, n$ such that

$$\Delta = \mathfrak{R}_n \triangleright \dots \triangleright \mathfrak{R}_2 \triangleright \mathfrak{R}_1 = \Omega, \quad (2.13)$$

where each transition $\mathfrak{R}_k \rightarrow \mathfrak{R}_{k-1}$ occurs at rate λ_k .

Starting from the n -coalescent, we can now define the Kingman coalescent. Let $m, n \in \mathbb{N}$, $2 \leq m < n$, and define the restriction $\rho_{nm} : E_n \rightarrow E_m$ as the map that drops all the labels $m+1, \dots, n$ from the subsets in the partitions in E_n . It is immediate that if $(R_t^n)_{t \geq 0}$ is the n -coalescent, then $(\rho_{nm} R_t^n)_{t \geq 0}$ is the m -coalescent.

Let E be the set of partitions of \mathbb{N} and define $\rho_n : E \rightarrow E_n$ as the restriction that drops all labels $> n$.

Theorem 2.3. *There exists a unique process $R = (R_t)_{t \geq 0}$ on E such that $(\rho_n R_t)_{t \geq 0}$ is the n -coalescent for all $n \geq 2$. This process is called the Kingman coalescent.*

Thus R is the Markov process on E with initial state $R_0 = (\{i\} : i \in \mathbb{N})$ such that each pair of subsets in R_t (equivalently, each pair of lineages in the coalescent tree) coalesces at rate 1 for all $t \geq 0$.

Coming down from infinity The Kingman coalescent describes the ancestry of a countably infinite Wright-Fisher population up to the most recent common ancestor, i.e., the individual from which the whole population has descended. At this point it is natural to wonder how it is possible that an infinite number of lineages decreases to a finite number at positive times. This is expressed by a very interesting property of the Kingman coalescent, namely, the fact that it comes down from infinity.

As before, let $D_t = |R_t|$ be the number of subset in R_t . Then $D_0 = \infty$, i.e., there are infinitely many lineages at the beginning of the coalescent tree.

Theorem 2.4. *The Kingman coalescent comes down from infinity, i.e.,*

$$\mathbb{P}(D_t < \infty \quad \forall t > 0) = 1. \quad (2.14)$$

Proof. We will equivalently show that

$$\forall t > 0 \quad \forall \epsilon > 0 \quad \exists N_{t,\epsilon} \in \mathbb{N} : \quad \mathbb{P}(D_t > N_{t,\epsilon}) < \epsilon. \quad (2.15)$$

Fix $t > 0$ and $\epsilon > 0$ arbitrary. Let $R_t^n = \rho_n(R_t)$ be the natural restriction of the coalescent to E_n and let $D_t^n = |R_t^n|$ be the number of subsets in R_t^n (see (2.11)). Since $(D_t^n)_{t \geq 0}$ is a death process with transitions $k \rightarrow k - 1$ occurring at rate $\lambda_k = \binom{k}{2}$, $k \in \mathbb{N} \setminus \{1\}$, the time τ_k that the process spends in state k is distributed as an exponential random variable with parameter λ_k . Therefore from the Markov inequality we obtain

$$\mathbb{P}(D_t > N_{t,\epsilon}) = \mathbb{P}\left(\sum_{k=N_{t,\epsilon}}^n \tau_k > t\right) \leq \frac{1}{t} \mathbb{E}\left[\sum_{k=N_{t,\epsilon}}^n \tau_k\right] \leq \frac{1}{t} \sum_{k=N_{t,\epsilon}}^{\infty} \mathbb{E}[\tau_k] = \frac{1}{t} \sum_{k=N_{t,\epsilon}}^{\infty} \frac{1}{\lambda_k}. \quad (2.16)$$

Notice that the last sum is independent of n . Since $\sum_{k \in \mathbb{N}} \lambda_k < \infty$, for all $t > 0$ and for all $\epsilon > 0$ we can choose $N_{t,\epsilon}$ large enough to ensure that

$$\frac{1}{t} \sum_{k=N_{t,\epsilon}}^{\infty} \frac{1}{\lambda_k} < \epsilon. \quad (2.17)$$

Therefore

$$\limsup_{n \rightarrow \infty} \mathbb{P}(D_t^n > N_{t,\epsilon}) = \lim_{n \rightarrow \infty} \left(\sup_{m \geq n} \mathbb{P}(D_t^m > N_{t,\epsilon})\right) < \epsilon \quad (2.18)$$

and the statement follows. \square

3 The Wright-Fisher model with Geometric Seed-Bank Component

3.1 Definition of the model

The model introduced in [1] consists of a haploid population of fixed size N that reproduces in discrete generations. Each individual carries a genetic type from a generic type space E . In what follows we will focus on the bi-allelic case, $E = \{A, a\}$. Alongside the active population, we consider a seed-bank of fixed size M containing the dormant individuals.

Given $M, N \in \mathbb{N}$, take $\epsilon \in [0, 1]$ such that $\epsilon N \leq M$ and set $\delta = \epsilon N / M$. We will assume for convenience that $\epsilon N = \delta M$ is a natural number. The dynamics of the model is as follows:

- The N active individuals produce $(1 - \epsilon)N$ active individuals in the next generation, where every new individual randomly chooses a parent from the previous generation and adopts its type.
- The remaining $\epsilon N (= \delta M)$ individuals from the active population produce individuals that become dormant, i.e., seeds in the next generation.
- From the seed-bank, $\delta M (= \epsilon N)$ individuals become active and leave the seed-bank. Therefore in the next generation the active population again consists of $(1 - \epsilon)N + \epsilon N = N$ individuals, where the first term comes from the previously active population and the second term from the previously dormant population.
- The remaining $(1 - \delta)M$ seeds remain inactive and stay in the seed-bank. Therefore in the next generation the population in the seed-bank again consists of $(1 - \delta)M + \delta M = M$ individuals, where the first term comes from the previously dormant population and the second term from the previously active population.

Definition 3.1 (Wright-Fisher model with geometric seed-bank component). *Let M, N, ϵ, δ and E be as above. Given an initial genetic type configuration (ξ_0, η_0) with $\xi_0 \in E^N, \eta_0 \in E^M$, denote by*

$$\xi_k = (\xi_k(i))_{i \in [N]}, \quad \eta_k = (\eta_k(j))_{j \in [M]}, \quad k \in \mathbb{N}, \quad (3.1)$$

the random genetic type configuration of, respectively, the active individuals and the dormant individuals at generation k . The discrete-time Markov chain $(\xi_k, \eta_k)_{k \in \mathbb{N}}$ with state space $E^N \times E^M$ is called the type configuration process of the Wright-Fisher model with geometric seed-bank component.

Notice that the time a seed stays in the seed-bank before becoming active is a random variable with a geometric distribution with parameter δ and the times for every seed are i.i.d. Similarly, the probability that an individual produces a dormant seed is ϵ , and individuals do so independently of one another.

We want to look at the behaviour of the frequency of a alleles in both the active population and the seed-bank. We therefore define

$$X_k^N = \frac{1}{N} \sum_{i \in [N]} \mathbb{1}_{\{\xi_k(i)=a\}}, \quad Y_k^M = \frac{1}{M} \sum_{j \in [M]} \mathbb{1}_{\{\eta_k(j)=a\}}, \quad (3.2)$$

which represent the fraction of individuals having type a at generation k in, respectively, the active population and the dormant population. They together form a discrete-time Markov chain taking values in $I^N \times I^M$, with

$$I^N = \left\{ 0, \frac{1}{N}, \frac{2}{N}, \dots, 1 \right\} \subset [0, 1], \quad I^M = \left\{ 0, \frac{1}{M}, \frac{2}{M}, \dots, 1 \right\} \subset [0, 1]. \quad (3.3)$$

Abbreviate $\mathbb{P}_{x,y}(\cdot) = \mathbb{P}(\cdot \mid X_0^N = x, Y_0^M = y)$, $(x, y) \in I^N \times I^M$. The transition probabilities of the Markov chain can be characterized as in the following proposition.

Proposition 3.2. *Let $c = \epsilon N = \delta M \in \mathbb{N}$. For $(x, y), (\bar{x}, \bar{y}) \in I^N \times I^M$,*

$$\begin{aligned} p_{x,y}(\bar{x}, \bar{y}) &= \mathbb{P}_{x,y}(X_1^N = \bar{x}, Y_1^M = \bar{y}) \\ &= \sum_{i=0}^c \mathbb{P}_{x,y}(Z = i) \mathbb{P}_{x,y}(U = \bar{x}N - i) \mathbb{P}_{x,y}(V = \bar{y}M + i) \end{aligned} \quad (3.4)$$

where Z, U, V are independent under $\mathbb{P}_{x,y}$ and $Z \cong \text{Hyp}_{M,c,yM}$ (hypergeometric distribution), $U \cong \text{Bin}_{N-c,x}$ and $V \cong \text{Bin}_{c,x}$ (binomial distributions).

Proof. The random variables introduced in the previous proposition have a simple interpretation:

- Z is the number of seeds of type a that become active in the next generation or, equivalently, the number of plants at generation 1 that are offspring of a seed of type a at generation 0.
- U is the number of plants that are offspring of plants of type a in the previous generation (and therefore are themselves of type a).
- V is the number of seeds that are offspring of plants of type a in the previous generation

With the interpretation of Z, U and V given above, their distributions are immediate from Definition 1. By construction, we have $X_1^N = \frac{U+Z}{N}$ and $Y_1^M = y + \frac{V-Z}{M}$ and so the claim follows. \square

3.2 Scaling and diffusion limit

In population genetics, it often happens that non-trivial limiting structures are revealed when the parameters describing evolutionary forces (such as mutation, selection and recombination) are scaled in terms of the population size N . In our case, we assume that ϵ, δ and M scale with N , namely, there exist $c, K \in (0, \infty)$ such that

$$\epsilon = \epsilon(N) = \frac{c}{N}, \quad M = M(N) = \frac{N}{K}. \quad (3.5)$$

We may assume without loss of generality that $c \in \mathbb{N}_0$. Under these assumptions, the seed-bank age has a geometric distribution with parameter

$$\delta = \delta(N) = \frac{c}{M(N)} = \frac{cK}{N}, \quad (3.6)$$

and c is the number of active individuals in each generation that moves to the seed-bank, equivalently, the number of seeds that become active. The parameter K determines the relative size of the seed-bank with respect to the size of the active population.

Proposition 3.3. *Assume that (3.5) holds. Consider a test function $f \in C^{(3)}([0, 1]^2)$. For any $(x, y) \in I^N \times I^M$, define the discrete-time generator $A^N = A_{\epsilon, \delta, M}^N$ of the frequency Markov chain $(X_k^N, Y_k^M)_{k \in \mathbb{N}_0}$ by*

$$(A^N f)(x, y) = N \mathbb{E}_{x, y} [f(X_1^N, Y_1^M) - f(x, y)]. \quad (3.7)$$

Then, for all $(x, y) \in [0, 1]^2$,

$$\lim_{N \rightarrow \infty} (A^N f)(x, y) = (Af)(x, y), \quad (3.8)$$

where A is defined by

$$Af(x, y) = c(y - x) \frac{\partial f}{\partial x}(x, y) + cK(x - y) \frac{\partial f}{\partial y}(x, y) + \frac{1}{2}x(1 - x) \frac{\partial^2 f}{\partial x^2}(x, y). \quad (3.9)$$

Since the state space of the frequency Markov chain can be embedded into the compact unit square $[0, 1]^2$, by standard arguments we obtain tightness and convergence on path-space and are able to identify the limit of the frequency Markov chain as a pair of SDEs:

Corollary 3.4. (Wright-Fisher diffusion with seed-bank component). *Under the conditions of Proposition 3.3, if $X_0^N \rightarrow x$ a.s. and $Y_0^M \rightarrow y$ a.s., then*

$$w - \lim_{N \rightarrow \infty} \left(X_{\lceil Nt \rceil}^N, Y_{\lceil Nt \rceil}^N \right)_{t \geq 0} = (X_t, Y_t)_{t \geq 0} \quad (3.10)$$

on $D_{[0, \infty)}([0, 1]^2)$, where $(X_t, Y_t)_{t \geq 0}$ is the two-dimensional diffusion solving

$$\begin{aligned} dX_t &= c(Y_t - X_t)dt + \sqrt{X_t(1 - X_t)}dB_t, \\ dY_t &= xK(X_t - Y_t)dt, \end{aligned} \quad (3.11)$$

where $(B_t)_{t \geq 0}$ is standard Brownian motion.

Proof. The claim follows from standard arguments after showing that A is indeed the generator of a Markov process and using the fact that convergence of the generator on a dense class of test functions implies convergence of the Markov process (cf. [2], Chapter 8, Proposition 2.4). \square

Remark. Two further scaling limits are obtained when we abandon the assumption $N = KM$ and assume instead that either $N/M \rightarrow 0$ or $M/N \rightarrow 0$.

- $N/M \rightarrow 0$

For this case, after rescaling time with N as before, we obtain

$$\lim_{N \rightarrow \infty} (A^N f)(x, y) = c(y - x) \frac{\partial f}{\partial x}(x, y) + \frac{1}{2}x(1 - x) \frac{\partial^2 f}{\partial x^2}(x, y). \quad (3.12)$$

This shows that the limiting process is one-dimensional, namely, the seed-bank frequency process $(Y_t)_{t \geq 0}$ is constantly equal to y and the process $(X_t)_{t \geq 0}$ is a Wright-Fisher diffusion with migration rate c and attracted to the mean y . In this case the seed-bank is much larger than the active population and acts as a reservoir with constant allele frequency y , with which the active individuals interact.

- $M/N \rightarrow 0$

For this case, after rescaling time with M , we obtain

$$\lim_{N \rightarrow \infty} (A^M f)(x, y) = c(y - x) \frac{\partial f}{\partial y}(x, y), \quad (3.13)$$

and constant frequency x in the active population. This tells us that the genetic configuration of the seed-bank converges to the genetic configuration of the active population in a deterministic way.

4 The Dual of the Seed-Bank Frequency Process

Several models in population genetics are paired to a dual process. For example, as we saw in Section 2, the classical Wright-Fisher diffusion has a dual process that equals the block counting process of the Kingman coalescent. In this section we look at the dual process of the Wright-Fisher diffusion with geometric seed-bank component described in Section 3.

Definition 4.1. Let $(N_t, M_t)_{t \geq 0}$ be the continuous-time Markov chain taking values in $\mathbb{N}_0 \times \mathbb{N}_0$ with allowed transitions

$$(n, m) \rightarrow \begin{cases} (n-1, m+1), & \text{at rate } cn, \\ (n+1, m-1), & \text{at rate } cKm, \\ (n-1, m), & \text{at rate } \binom{n}{2}. \end{cases} \quad (4.1)$$

This process is called the block-counting process of the seed-bank coalescent.

Denote by $\mathbb{P}^{n,m}$ the probability distribution under which $(M_0, N_0) \stackrel{a.s.}{=} (n, m)$ and by $\mathbb{E}^{n,m}$ the corresponding expectation. Since the process either exchanges units between N_t and M_t or loses units in N_t , it is easy to see that

$$\lim_{t \rightarrow \infty} N_t + M_t \stackrel{a.s.}{=} 1 \quad \forall n, m \in \mathbb{N}_0. \quad (4.2)$$

We need to show that $(N_t, M_t)_{t \geq 0}$ is the moment dual of $(X_t, Y_t)_{t \geq 0}$.

Theorem 4.2. For all $(x, y) \in [0, 1]^2$, all $n, m \in \mathbb{N}_0$ and all $t \geq 0$,

$$\mathbb{E}_{x,y}[X_t^n Y_t^m] = \mathbb{E}^{n,m}[x^{N_t} y^{M_t}]. \quad (4.3)$$

Proof. The claim follows from applying the generator A of $(X_t, Y_t)_{t \geq 0}$ to the test function $f(x, y, n, m) = x^n y^m$, viewed as a function of x, y for fixed $n, m \in \mathbb{N}_0$. This yields

$$\begin{aligned} Af(x, y) &= c(y-x) \frac{df}{dx}(x, y) + \frac{1}{2}x(1-x) \frac{d^2 f}{dx^2}(x, y) + xK(x-y) \frac{df}{dy}(x, y) \\ &= c(y-x)nx^{n-1}y^m + \frac{1}{2}x(1-x)n(n-1)x^{n-2}y^m + cK(y-x)x^ny^{m-1} \\ &= cn(x^{n-1}y^{m+1} - x^ny^m) + \binom{n}{2}(x^{n-1}y^m - x^ny^m) + cKm(x^{n+1}y^{m-1} - x^ny^m), \end{aligned} \quad (4.4)$$

where the right-hand side is the generator of $(N_t, M_t)_{t \geq 0}$ applied to f , viewed as a function of (n, m) for fixed $x, y \in [0, 1]$. \square

The duality in Theorem 4.2 allows us to derive results on the long-term behaviour of the system. For the classical Wright-Fisher diffusion, which is given by

$$dZ_t = \sqrt{Z_t(1-Z_t)}dW_t, \quad Z_0 = z \in [0, 1], \quad (4.5)$$

we had that the system gets absorbed at the boundaries 0 or 1 almost surely after finite time, hitting 1 with probability z and 0 with probability $1-z$. When adding a seed-bank component, we still have absorbing states $(0, 0)$ and $(1, 1)$, but it is more involved to analyse the time to absorption. However, we can still compute the fixation probabilities.

Proposition 4.3. All mixed moments of $(X_t, Y_t)_{t \geq 0}$ converge to the same finite limit depending only on x, y, K . More precisely, for each fixed $n, m \in \mathbb{N}$,

$$\lim_{t \rightarrow \infty} \mathbb{E}_{x,y}[X_t^n Y_t^m] = \frac{y + xK}{1 + K}. \quad (4.6)$$

Proof. Let $(N_t, M_t)_{t \geq 0}$ be as in Definition 4.1 with initial state $(n, m) \in \mathbb{N}_0 \times \mathbb{N}_0$. Let $T = \inf \{t > 0 \mid N_t + M_t = 1\}$ be the first time at which there is only one block left in the system. Note that $\mathbb{E}_{x,y}[T] < \infty$ for all initial configurations. By Theorem 4.2 we have:

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E}_{x,y}[X_t^n Y_t^m] &= \lim_{t \rightarrow \infty} \mathbb{E}^{n,m}[x^{N_t} y^{M_t}] \\
&= \lim_{t \rightarrow \infty} \mathbb{E}^{n,m}[x^{N_t} y^{M_t} \mid T \leq t] \mathbb{P}^{n,m}(T \leq t) \\
&\quad + \lim_{t \rightarrow \infty} \mathbb{E}^{n,m}[x^{N_t} y^{M_t} \mid T > t] \mathbb{P}^{n,m}(T > t) \\
&= \lim_{t \rightarrow \infty} [x \mathbb{P}^{n,m}(N_t = 1, T \leq t) + y \mathbb{P}^{n,m}(M_t = 1, T \leq t)] \\
&= \lim_{t \rightarrow \infty} [x \mathbb{P}^{n,m}(N_t = 1) + y \mathbb{P}^{n,m}(M_t = 1)] \\
&= \frac{xK}{1+K} + \frac{y}{1+K}, \tag{4.7}
\end{aligned}$$

where the third equality follows from the fact that $\lim_{t \rightarrow \infty} \mathbb{P}^{n,m}(T > t) = 0$ and the last inequality follows from the fact that the invariant distribution of a single individual jumping between two states (active and dormant) at rates c and cK converges to $(\frac{K}{1+K}, \frac{1}{1+K})$. \square

Corollary 4.4 (Fixation in law). *Given c and K , $(X_t, Y_t)_{t \geq 0}$ converges in distribution as $t \rightarrow \infty$ to a two-dimensional random variable (X_∞, Y_∞) whose distribution is given by*

$$\frac{y + xK}{1+K} \delta_{(1,1)} + \frac{1 + (1-x)K - y}{1+K} \delta_{(0,0)}. \tag{4.8}$$

Proof. The proof follows from standard arguments after noting that the only probability distribution on $[0, 1]^2$ such that all moments are constant and equal to $\frac{xK+y}{1+K}$ is given by (4.8). This yields uniqueness, while convergence in distribution follows from the convergence of all moments and the fact that $[0, 1]^2$ is compact. \square

Notice that the above observation is consistent with the classical Wright-Fisher diffusion. In fact, if we let $K \rightarrow \infty$ (i.e., the size of the dormant population becomes small compared to the size of the active population), then the probability of fixation at 1 approaches x , like in the Wright-Fisher diffusion. Conversely, if we let $K \rightarrow 0$ (the size of the active population becomes small compared to the size of the dormant population), then the fixation probability approaches y .

Observing that $(KX_t + Y_t)_{t \geq 0}$ is a bounded martingale with limiting distribution (4.8), we get almost sure convergence of $(X_t, Y_t)_{t \geq 0}$ to (X_∞, Y_∞) . However, we will see in Section 5 that fixation may not happen in finite time since the block-counting process $(N_t, M_t)_{t \geq 0}$ started at $(n, m) = (\infty, 0)$ *does not come down from infinity*. This means that the infinite population does not have a most-recent common ancestor and initial genetic variability is never lost in finite time.

5 The Seed-Bank Coalescent

5.1 Definition of the seed-bank coalescent

The purpose of the coalescent process is to study the genealogy of a sample taken from the population of our Wright-Fisher model with geometric seed-bank component. As we will see, the block-counting process defined in Section 4 will play a central role.

For $k \in \mathbb{N}$, let \mathcal{P}_k be the set of partitions of $[k] = \{1, 2, \dots, k\}$. For $\pi \in \mathcal{P}_k$, let $|\pi|$ denote the number of elements (blocks) in π . We define the space of *marked* partitions to be

$$\mathcal{P}_k^{\{p,s\}} = \left\{ (\zeta, \vec{u}) \mid \zeta \in \mathcal{P}_k, \vec{u} \in \{s,p\}^{|\zeta|} \right\}, \quad (5.1)$$

where each block in a partition is assigned a flag (p for plant, s for seed) that indicates whether the block is currently in the active population or in the seed-bank. An example of an element in $\mathcal{P}_k^{\{p,s\}}$ is $\pi = \{\{1, 2\}^p, \{3, 4\}^p, \{5\}^s\}$. We define the following relations for two marked partitions $\pi, \pi' \in \mathcal{P}_k^{\{p,s\}}$:

- $\pi \succ \pi'$ when π' can be obtained from π by merging two p -blocks and the resulting block again carries a p -flag. For example:

$$\{\{1, 2\}^p, \{3, 4\}^p, \{5\}^s\} \succ \{\{1, 2, 3, 4\}^p, \{5\}^s\}. \quad (5.2)$$

- $\pi \bowtie \pi'$ when π' can be obtained from π by changing the flag of precisely one block of π . For example:

$$\{\{1, 2\}^p, \{3, 4\}^p, \{5\}^s\} \bowtie \{\{1, 2\}^p, \{3, 4\}^s, \{5\}^s\}. \quad (5.3)$$

Definition 5.1 (The seed-bank k -coalescent). *Let $k \geq 2$ and $c, K \in (0, \infty)$. Define the seed-bank k -coalescent $(\Pi_t^{(k)})_{t \geq 0}$ with seed-bank intensity c and relative seed-bank size $1/K$ to be the $\mathcal{P}_k^{\{p,s\}}$ -valued continuous-time Markov chain with the following transition rates:*

$$\pi \rightarrow \pi' \text{ at rate } \begin{cases} 1, & \text{if } \pi \succ \pi', \\ c, & \text{if } \pi \bowtie \pi' \text{ and one } p \text{ is replaced by one } s, \\ cK, & \text{if } \pi \bowtie \pi' \text{ and one } s \text{ is replaced by one } p. \end{cases} \quad (5.4)$$

If $c = K = 1$, then we speak of the standard seed-bank k -coalescent.

Definition 5.2 (The seed-bank coalescent). *The seed-bank coalescent $(\Pi_t)_{t \geq 0} = (\Pi_t^{(\infty)})_{t \geq 0}$ with seed-bank intensity c and relative seed-bank size $1/K$ is defined as the unique Markov process distributed according to the projective limit as $k \rightarrow \infty$ of the laws of the seed-bank k -coalescents.*

Note that if we show consistency, then we obtain that the seed-bank coalescent is well-defined by the Kolmogorov extension theorem. The sketch of the proof is as follows. We define $\overleftarrow{\Pi}_t^{(k)}$ as the projection of $\Pi_t^{(k+1)}$ on $\mathcal{P}_k^{\{p,s\}}$. Since mergers and flag-flips involving the singleton $\{k+1\}$ are only visible in $\Pi_t^{(k+1)}$ and do not affect $\overleftarrow{\Pi}_t^{(k)}$, by the Markov property a change involving the singleton $\{k+1\}$ has no effect on any other transition. Therefore if $\overleftarrow{\Pi}_0^{(k)} = \Pi_0^{(k)}$, then $(\overleftarrow{\Pi}_t^{(k)})_{t \geq 0} = (\Pi_t^{(k)})_{t \geq 0}$ in distribution and, by the Kolmogorov extension theorem, the projective limit exists and is unique.

Observe that, when we distinguish the count of p -blocks and s -blocks, the distribution of the block counting process of the seed-bank coalescent is the same as the distribution of the process $(N_t, M_t)_{t \geq 0}$ in Definition 4.1. Moreover, as the Kingman coalescent describes the genealogy of a sample taken from the Wright-Fisher model, so the seed-bank coalescent describes the genealogy of a sample taken from the Wright-Fisher model with geometric seed-bank component.

In fact, consider a sample of $k \ll N$ individuals sampled from generation 0. Proceeding backwards in time, keeping track of the ancestors of every individual in the sample and their position in the active or dormant population, we can denote the configuration of the genealogy at generation $-i$ by $\Pi_i^{(N,k)} \in \mathcal{P}_k^{\{p,s\}}$, where two individuals are in the same block if and only if their ancestral lines have met before and up to generation $-i$. This means that all individuals in one block have precisely one common ancestor at generation $-i$ and the flag of the block indicates whether this ancestor is currently active or dormant. Looking at the process backwards in time, the possible transitions between a generation and the previous one are the following:

- One (or more) plants become seeds in the previous generation.

- One (or more) seeds become plants in the previous generation.
- Two (or more) ancestral lines merge, namely, two individuals have a common ancestor in the previous generation.
- Any possible combination of the events above.

When we take the limit as $N \rightarrow \infty$, only the first three events will play a role, since combinations will be much rarer.

Proposition 5.3. *Assume the scaling in (3.5), and let $\Pi_0^{(N,k)} \stackrel{\text{a.s.}}{=} \{\{1\}^p, \dots, \{k\}^p\}$ for fixed $k \in \mathbb{N}$. Then, for all $\pi, \pi' \in \mathcal{P}_k^{\{p,s\}}$ and all $i \in \mathbb{N}_0$, the transition probabilities scale as*

$$\mathbb{P}\left(\Pi_{i+1}^{(N,k)} = \pi' \mid \Pi_i^{(N,k)} = \pi\right) = \begin{cases} \frac{1}{N} + O(N^{-2}), & \text{if } \pi \succ \pi', \\ \frac{c}{N} + O(N^{-2}), & \text{if } \pi \bowtie \pi' \text{ and a } p \text{ is replaced by an } s, \\ \frac{cK}{N} + O(N^{-2}), & \text{if } \pi \bowtie \pi' \text{ and an } s \text{ is replaced by a } p, \\ O(N^{-2}), & \text{otherwise.} \end{cases} \quad (5.5)$$

Proof. Because of the way the model is defined, at every new generation precisely c plants (out of N) become seeds and c seeds (out of $M = N/K$) become plants. Therefore, if the current state of the genealogical process contains at least one p -block, then the probability that a given p -block changes flag to s (i.e., becomes active) in the next generation is c/N . Similarly, if there is at least one s -block, then the probability that it changes flag to p is cK/N . When remaining in the active population, the probability that a given p -block chooses a fixed active ancestor is $(1 - \frac{c}{N}) \frac{1}{N}$, since the first term equals the probability that the ancestor is active and the second term equals the probability to choose a particular ancestor out of N . Therefore the probability of coalescence of two given p -blocks is $(1 - \frac{c}{N})^2 \frac{1}{N}$. Since the blocks move independently, the probability of any combination of the three events described above is of order N^{-2} and therefore the claim follows. \square

Corollary 5.4. *Under the scaling in (3.5), for any $k \in \mathbb{N}$, $(\Pi_{\lfloor Nt \rfloor}^{(N,k)})_{t \geq 0}$ converges weakly as $N \rightarrow \infty$ to the seed-bank coalescent $(\Pi_t^{(k)})_{t \geq 0}$ started from k plants.*

Proof. From Proposition 5.3 we see that the generator of $(\Pi_{\lfloor Nt \rfloor}^{(N,k)})_{t \geq 0}$ converges to the generator of $(\Pi_t^{(k)})_{t \geq 0}$, which is defined via (5.5). Weak convergence of the process then follows from standard arguments. \square

5.2 Properties of the seed-bank coalescent

According to the definition given by Pitman (cf. [5]) and Schweinsberg (cf. [6]), a coalescent process *comes down from infinity* if the corresponding block counting process, started with an infinite sample, is such that the number of blocks is almost surely finite for all positive times. We know that this holds for the Kingman coalescent, which indeed comes down from infinity. We will now see that, on the contrary, the seed-bank coalescent does not. In fact, its block counting process *stays infinite*, in a sense we will specify in the next statement.

Theorem 5.5. *The seed-bank coalescent does not come down from infinity. In fact, its block-counting process $(N_t, M_t)_{t \geq 0}$ stays infinite for all $t \geq 0$ \mathbb{P} -a.s., namely, for each starting configuration (n, m) such that $n + m$ is infinite,*

$$\mathbb{P}\left(M_t^{(n,m)} = \infty \quad \forall t \geq 0\right) = 1. \quad (5.6)$$

The proof of Theorem 5.5 is based on a coupling of the block-counting process with a simplified *coloured* seed-bank coalescent introduced below. The coloured seed-bank coalescent is essentially defined as the normal seed-bank coalescent, with the addition that we mark with a colour the individuals who have left the seed-bank at least once. The formal definition is given below.

Analogously to the construction of the seed-bank coalescent, we first define the set of *coloured* marked partitions as

$$\begin{aligned}\mathcal{P}_k^{\{p,s\} \times \{w,b\}} &= \{(\pi, \vec{u}, \vec{v}) \mid (\pi, \vec{u}) \in \mathcal{P}_k^{\{p,s\}}, \vec{v} \in \{w,b\}^k\}, \quad k \in \mathbb{N}, \\ \mathcal{P}^{\{p,s\} \times \{w,b\}} &= \{(\pi, \vec{u}, \vec{v}) \mid (\pi, \vec{u}) \in \mathcal{P}^{\{p,s\}}, \vec{v} \in \{w,b\}^{\mathbb{N}}\}.\end{aligned}\tag{5.7}$$

It is easy to see how these partitions coincide with the ones previously introduced, with the addition that now each element carries a flag indicating its colour: w for *white* and b for *blue*. An important point to notice is that the p -flags and s -flags are assigned to *blocks*, whereas the colour-flags are assigned to *individuals*. As in the previous definitions, we define the following relations:

- $\pi \succ \pi'$ when π' can be obtained from π by merging two p -blocks in π resulting in one p -block in π' while retaining the colours of the individuals,
- $\pi \times \pi'$ if π' can be obtained from π by changing the flag of a block from p to s while leaving the individual colours unchanged,
- $\pi \rtimes \pi'$ if π' can be obtained from π by changing the flag of a block from s to p and colouring all individuals in this block blue.

In other words, after leaving the seed-bank, individuals are coloured blue.

We can now formally define the coloured seed-bank coalescent.

Definition 5.6 (The coloured seed-bank coalescent). *For $k \in \mathbb{N}$ and $c, K \in (0, \infty)$ define the coloured seed-bank k -coalescent with seed-bank intensity c and seed-bank size $1/K$, denoted by $(\underline{\Pi}_t)_{t \geq 0}$, as the continuous-time Markov chain with values in $\mathcal{P}_k^{\{p,s\} \times \{w,b\}}$ and transition rates given by*

$$\pi \mapsto \pi' \text{ at rate } \begin{cases} 1, & \text{if } \pi \succ \pi', \\ c, & \text{if } \pi \times \pi', \\ cK, & \text{if } \pi \rtimes \pi'. \end{cases}\tag{5.8}$$

The coloured seed-bank coalescent with seed-bank intensity c and seed-bank size $1/K$ is the unique Markov process on $\mathcal{P}^{\{p,s\} \times \{w,b\}}$ given by the projective limit of the distributions of the k -coloured seed-bank coalescents as $k \rightarrow \infty$.

Remark

(i) Notice that the coloured seed-bank coalescent is well defined. In fact, the colour of each individual only depends on its path and not on other individuals. Therefore the consistency of the laws of the k -coloured seed-bank coalescents corresponds to the consistency of the seed-bank k -coalescents discussed in the Section 5.1. Analogously, we obtain existence and uniqueness of the coloured seed-bank coalescent from the Kolmogorov extension theorem.

(ii) The (uncoloured) seed-bank (k -)coalescent can be obtained from the coloured one by simply omitting the colour flags. Notice that, if we only consider blocks containing at least one white particle, then we obtain a coalescent similar to the seed-bank coalescent where lineages are discarded once they leave the seed-bank.

For $t \geq 0$, define

- \underline{N}_t is the number of p -blocks containing at least one white plant,
- \underline{M}_t is the number of s -blocks containing at least one white seed.

We will use the superscript (n, m) to denote that the process starts with n plants and m seeds \mathbb{P} -a.s., where $n, m = \infty$ means that we start with countably many individuals. The initial configuration will always only contain white individuals, so all labels are set to w . Notice that \underline{N}_t is non-increasing by construction.

Proposition 5.7. *For any $n, m \in \mathbb{N} \cup \{\infty\}$, the processes $(N_t^{(n,m)}, M_t^{(n,m)})_{t \geq 0}$ and $(\underline{N}_t^{(n,m)}, \underline{M}_t^{(n,m)})_{t \geq 0}$ can be coupled in such a way that*

$$\mathbb{P}\left(\forall t \geq 0: N_t^{(n,m)} \geq \underline{N}_t^{(n,m)}, M_t^{(n,m)} \geq \underline{M}_t^{(n,m)}\right) = 1. \quad (5.9)$$

Proof. By construction, the set of p -block containing at least one white particles is a subset of the set of p -blocks, and the same holds for the s -blocks, so the result is immediate. \square

Proof of Theorem 5.5. Thanks to Proposition 5.7, it suffices to prove the statement for $\underline{N}_t^{(n,m)}$ and $\underline{M}_t^{(n,m)}$. Without loss of generality, we will consider the case with no seeds in the initial configuration, namely, $m = 0$.

For $n \in \mathbb{N} \cup \{\infty\}$ define

$$\tau_j^n := \inf\{t \geq 0 \mid \underline{N}_t^{(n,0)} = j\}, \quad 1 \leq j \leq n-1, \quad j < \infty, \quad (5.10)$$

i.e., the first time the number of active blocks of an n -sample reaches k . Notice that $(\underline{N}_t)_{t \geq 0}$ behaves like the block-counting process of a Kingman coalescent where particles can either coalesce or disappear at a rate proportional to the number of particles alive. Since the corresponding values for the Kingman coalescent are finite \mathbb{P} -a.s., the τ_j^n are too. In fact, for any n , we have that $\tau_{j-1}^n - \tau_j^n$ has an exponential distribution with parameter

$$\lambda_j = \binom{j}{2} + cj, \quad (5.11)$$

where the first term in the right-hand side is the rate of a coalescence event and the second term is the rate of a *deactivation* event, where by the latter we mean a transition of $(\underline{N}_t^n, \underline{M}_t^n)_{t \geq 0}$ of type $(j+1, l) \mapsto (j, l+1)$, i.e., the transition of a lineage from the active population to the seed-bank.

Whenever the number of active blocks decreases, it can be by coalescence or by deactivation. Hence

$$\mathbb{P}(\text{deactivation at } \tau_{j-1}^n) = \frac{cj}{\binom{j}{2} + cj} = \frac{2c}{2c + j - 1}, \quad (5.12)$$

independently of the number of active blocks. We can now define

$$X_j^n = \mathbb{1}_{\{\text{deactivation at } \tau_{j-1}^n\}}, \quad j = 2, \dots, n, \quad (5.13)$$

which are independent Bernoulli random variables $X_j^n \cong \text{Ber}\left(\frac{2c}{2c+j-1}\right)$. Notice that, due to the memoryless property of the exponential distribution, X_j^n depends on j but is independent of τ_{j-1}^n . We now define A_t^n as the number of deactivations up to time $t \geq 0$. Formally, for $n \in \mathbb{N} \cup \{\infty\}$,

$$A_t^n = \sum_{j=2}^n X_j^n \mathbb{1}_{\{\tau_{j-1}^n < t\}}. \quad (5.14)$$

Denote by $(|\tilde{\Pi}_t^n|)_{t \geq 0}$ the block counting process of the Kingman coalescent started from n blocks. Fixed $n \in \mathbb{N}$, since $\lambda_j > \binom{j}{2}$, the death process of the active blocks in the seed-bank coalescent is faster than in the Kingman coalescent (coalescence occurs at the same rate but in the seed-bank coalescent blocks can also move to the seed-bank). Therefore, for all $t \geq 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tau_{\lfloor \log n - 1 \rfloor}^n \leq t) \geq \lim_{n \rightarrow \infty} \mathbb{P}(|\tilde{\Pi}_t^n| \leq \lfloor \log n - 1 \rfloor) \geq \lim_{n \rightarrow \infty} \mathbb{P}(|\tilde{\Pi}_t| \leq \log n - 1) = 1, \quad (5.15)$$

where the last inequality follows from the fact that the Kingman coalescent $(\tilde{\Pi}_t)_{t \geq 0}$ comes down from infinity. Furthermore,

$$\mathbb{P} \left(A_t^n \geq \sum_{j=\log n}^n X_j^n \right) \geq \mathbb{P} \left(\mathbb{1}_{\{\tau_{\log n - 1}^n < t\}} \sum_{j=\log n}^n X_j^n \geq \sum_{j=\log n}^n X_j^n \right) \geq \mathbb{P}(\tau_{\log n - 1}^n < t). \quad (5.16)$$

Together with (5.15), this yields

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(A_t^n \geq \sum_{j=\log n}^n X_j^n \right) = 1. \quad (5.17)$$

Combining (5.12) with the properties of Bernoulli random variables, we get

$$\mathbb{E} \left[\sum_{j=\log n}^n X_j^n \right] = \sum_{j=\log n}^n \frac{2c}{2c+j-1} = 2c(\log n - \log \log n) + R(c, n), \quad (5.18)$$

where $R(c, n)$ converges to a finite value depending on c , the seed-bank intensity, as $n \rightarrow \infty$. By independence of the X_j^n 's, we obtain that the variance is given by

$$\text{Var} \left(\sum_{j=\log n}^n X_j^n \right) = \sum_{j=\log n}^n \text{Var}(X_j^n) = \sum_{j=\log n}^n \frac{2c}{2c+j-1} \left(1 - \frac{2c}{2c+j-1} \right) \leq 2c \log n \quad \text{as } n \rightarrow \infty. \quad (5.19)$$

Therefore for any $\epsilon > 0$ we can choose n large enough so that $\mathbb{E}[\sum_{j=\log n}^n X_j^n] \geq (2c - \epsilon) \log n$ and obtain by Chebyshev's inequality that

$$\begin{aligned} \mathbb{P} \left(\sum_{j=\log n}^n X_j^n < c \log n \right) &\leq \mathbb{P} \left(\sum_{j=\log n}^n X_j^n - \mathbb{E} \left[\sum_{j=\log n}^n X_j^n \right] < -(c - \epsilon) \log n \right) \\ &\leq \mathbb{P} \left(\left| \sum_{j=\log n}^n X_j^n - \mathbb{E} \left[\sum_{j=\log n}^n X_j^n \right] \right| > (c - \epsilon) \log n \right) \\ &\leq \frac{2c}{(c - \epsilon)^2 \log n}. \end{aligned} \quad (5.20)$$

Thus, in particular,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{j=\log n}^n X_j^n < \kappa \right) = 0 \quad \forall \kappa \in \mathbb{N}. \quad (5.21)$$

By construction, the processes $(A_t^n)_{t \geq 0}$ are coupled in such a way that $\mathbb{P}(A_t^\infty < \kappa) \leq \mathbb{P}(A_t^n < \kappa)$ for all $n \in \mathbb{N}$, $t \geq 0$ and $\kappa \in \mathbb{N}$. Therefore $\mathbb{P}(A_t^\infty < \kappa) = 0$, which yields

$$\mathbb{P}(A_t^\infty = \infty) = 1 \quad \forall t \geq 0. \quad (5.22)$$

Moreover, since $(A_t^\infty)_{t \geq 0}$ is non-decreasing in t , we conclude that

$$\mathbb{P}(A_t^\infty = \infty \quad \forall t \geq 0) = 1. \quad (5.23)$$

The above equation means that at any non-negative time t infinitely many blocks have moved from the active population to the seed-bank \mathbb{P} -a.s. We need to prove that this also implies that the number of blocks, i.e., lineages, in the seed-bank is infinite. This happens if a sufficiently large fraction of blocks does not move back to the active population, where it would be “instantly” reduced to a finite number by the coalescence mechanism (since the coalescence process in the active population is faster than the Kingman coalescent, which comes down from infinity).

To this end, we define \mathcal{B}_t as the set of blocks that visited the seed-bank before time $t \geq 0$ and were visible in the white seed-bank coalescent, namely,

$$\mathcal{B}_t = \left\{ B \subseteq \mathbb{N} \mid \exists r \in [0, t] \text{ such that } B^{\{s\}} \in \underline{\Pi}_r^{(\infty, 0)} \text{ and contains at least one } \textit{white} \text{ particle} \right\}. \quad (5.24)$$

Since the coloured coalescent starts in $(\infty, 0)$, the cardinality of \mathcal{B}_t is greater than or equal to A_t^∞ , and so by previous arguments we know that $\mathbb{P}(|\mathcal{B}_t| = \infty) = 1$. Since \mathcal{B}_t is countable, we can index its elements by natural numbers. Write $\mathcal{B}_t = \bigcup_{n \in \mathbb{N}} \{B_t^n\}$ and define the sets $\mathcal{B}_t^n = \{B_t^1, \dots, B_t^n\}$, $n \in \mathbb{N}$. Since \mathcal{B}_t is infinite \mathbb{P} -a.s., these sets \mathcal{B}_t^n exist for all n \mathbb{P} -a.s.

By construction, the following inequalities hold even pointwise:

$$\underline{M}_t^{(\infty, 0)} \geq \sum_{B \in \mathcal{B}_t} \mathbb{1}_{\{B^{\{s\}} \in \underline{\Pi}^{(\infty, 0)}\}} \geq \sum_{B \in \mathcal{B}_t^n} \mathbb{1}_{\{B^{\{s\}} \in \underline{\Pi}^{(\infty, 0)}\}}. \quad (5.25)$$

This implies that, for any $\kappa \in \mathbb{N}$,

$$\mathbb{P}(\underline{M}_t^{(\infty, 0)} \leq \kappa) \leq \mathbb{P} \left(\sum_{B \in \mathcal{B}_t^n} \mathbb{1}_{\{B^{\{s\}} \in \underline{\Pi}^{(\infty, 0)}\}} \leq \kappa \right) \leq \sum_{i=1}^{\kappa} \binom{n}{i} (e^{-ct})^i (1 - e^{-ct})^{n-1} \xrightarrow{n \rightarrow \infty} 0,$$

where in the second inequality we use the fact that for each of the n blocks in \mathcal{B}_t^n we know that $\mathbb{P}(B \in \underline{\Pi}_t^{(\infty, 0)}) \geq e^{-ct}$ and that they leave the seed-bank independently of each other (this implies that the sum in the right-hand side of the first inequality is dominated by a binomial random variable with parameters n and e^{-ct}).

Since the inequality in (5.26) holds for all $\kappa \in \mathbb{N}$ and its left-hand side does not depend on n , it follows that $\mathbb{P}(\underline{M}_t^{(\infty, 0)} = \infty) = 1$ for all $t > 0$. Note that this also implies $\mathbb{P}(\underline{M}_t^{(\infty, 0)} + \underline{N}_t^{(\infty, 0)} = \infty) = 1$. By monotonicity of the sum we deduce that

$$\mathbb{P}(\underline{M}_t^{(\infty, 0)} + \underline{N}_t^{(\infty, 0)} = \infty \quad \forall t > 0) = 1. \quad (5.26)$$

On the other hand, we have seen by comparison with the Kingman coalescent that $\mathbb{P}(\underline{N}_t^{(\infty, 0)} < \infty) = 1$ for all $t > 0$ and, again by monotonicity, we deduce $\mathbb{P}(\underline{N}_t^{(\infty, 0)} < \infty \quad \forall t > 0) = 1$. Putting these results together, we obtain the desired result:

$$\mathbb{P}(\underline{M}_t^{(\infty, 0)} = \infty \quad \forall t > 0) = 1. \quad (5.27)$$

□

6 A Multi-colony Extension

6.1 Migration

In this section we analyse a spatial version of the model introduced in Section 3 and described in Sections 4 and 5. Namely, we add migration, i.e., we allow individuals to choose their ancestors from a colony that is different from their own. To do this, we consider the discrete square torus \mathbb{T}^2 (which may be identified with the lattice $\mathbb{Z}^2 \cap [0, L]^2$, $L \in \mathbb{N}$, after it is folded into a cylinder whose ends are joined, i.e., $(a, 0) = (a, L)$, $(0, a) = (L, a)$, $a \in \mathbb{Z} \cap [0, L]$). At each lattice site there is a colony, which consists of an active population of size N and a seed-bank of size M . At each generation, every individual chooses its ancestor according to a random walk transition kernel $p(x, y)$, $x, y \in \mathbb{T}^2$, where we assume that it can only choose from a given colony individuals that were active in that colony in the previous generation. We assume $p(x, y)$, $x, y \in \mathbb{T}^2$, to depend only on the relative distance between the two colonies x and y , not on their position in the lattice. This will give our system the property of translation invariance. An example for a transition kernel that satisfies this assumption is

$$p(x, y) = (1 - \nu)\delta_{x,y} + \nu q(y - x), \quad x, y \in \mathbb{T}^2, \quad (6.1)$$

with $\delta_{x,y} = \mathbb{1}_{\{x=y\}}$ and $q : \mathbb{T}^2 \rightarrow [0, 1]$ a probability distribution. The parameter $\nu \in (0, 1]$ is the migration probability. An example of a possible distribution q is the uniform nearest-neighbour model

$$q(z) = \begin{cases} 1/4, & \text{if } \|z\| = 1, \\ 0, & \text{otherwise,} \end{cases} \quad (6.2)$$

where $\|\cdot\|$ is the lattice norm. This corresponds to a simple random walk on \mathbb{T}^2 .

We are interested in computing the probability $\psi((x, a), (y, b))$ that two individuals drawn uniformly at random from two colonies $x, y \in \mathbb{T}^2$ in states $a, b \in \{0, 1\}$ are *identical by descent*, i.e., their lineages coalesce. The states $a, b \in \{0, 1\}$ indicate whether the individual is drawn from the dormant population (state 0) or the active population (state 1), e.g. $(x, 0)$ means that an individual is drawn from the seed-bank in colony $x \in \mathbb{T}^2$. If $x = y$, then we require the two individuals to be distinct. We want to find an expression for the 4-vector

$$\Psi_{x,y} = \begin{pmatrix} \psi((x, 0), (y, 0)) \\ \psi((x, 0), (y, 1)) \\ \psi((x, 1), (y, 0)) \\ \psi((x, 1), (y, 1)) \end{pmatrix}, \quad x, y \in \mathbb{T}^2. \quad (6.3)$$

Note that, since $p(x, y)$ depends on $x - y$ only, the same is true for $\Psi_{x,y}$. Naturally, since we are considering a finite lattice and all populations are finite at each lattice site, we have that

$$\Psi_{x,y} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad x, y \in \mathbb{T}^2, \quad (6.4)$$

namely, no matter what colonies the individuals are drawn from, their lineages may move in and out of the seed-bank as well as migrate to other colonies in \mathbb{T}^2 , but eventually, going infinitely far back in time, they will coalesce. The problem becomes more interesting if we modify the dynamics to include mutation.

6.2 Migration and mutation with infinitely many alleles

Assume that at each generation, each individual spontaneously mutates to a new type with probability $\mu \in (0, 1)$, and with probability $1 - \mu$ proceeds to either become active resp. dormant (with probability δ resp. ϵ) or, if it is active

and remains active, to choose an ancestor from any colony $y \in \mathbb{T}^2$ (with probability $(1 - \epsilon)p(x, y)$) and adopt its type. This implies that in this new setting all parameters involved in the dynamics now express probabilities *conditional* on the fact that no mutation has occurred. More precisely, an individual in the seed-bank may

- mutate to a new type, with probability μ ,
- maintain its type and remain in the seed-bank, with probability $(1 - \mu)(1 - \delta)$,
- maintain its type and become active, with probability $(1 - \mu)\delta$.

Similarly, an active individual may

- mutate to a new type, with probability μ ,
- remain active and choose a random ancestor from colony $y \in \mathbb{T}^2$ (possibly $y = x$), with probability $(1 - \mu)(1 - \epsilon)p(x, y)$,
- maintain its type and become dormant, with probability $(1 - \mu)\epsilon$.

Since a lineage is interrupted when a mutation occurs, we now define two individuals to be *identical by descent* if their lineages coalesce before a mutation affects either lineage. We begin by deriving a recursive relation for the family $\{\Psi_{x,y}\}_{x,y \in \mathbb{T}^2}$.

Proposition 6.1. For $x, y \in \mathbb{T}^2$,

$$\Psi_{x,y} = \Phi_{x,y} + A_{x,y}\Psi_{0,0} + \sum_{w,z \in \mathbb{T}^2} B_{w-x,z-y}\Psi_{w,z}, \quad (6.5)$$

where

$$\Phi_{x,y} = (1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ p^2(x, y)\frac{1}{N} \end{pmatrix}, \quad (6.6)$$

$$A_{x,y} = -(1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p^2(x, y)\frac{1}{N} \end{pmatrix}, \quad (6.7)$$

$$B_{w-x,z-y} = C\delta_{x,w}\delta_{y,z} + D_{w-x,z-y}, \quad (6.8)$$

with

$$C = (1 - \mu)^2 \begin{pmatrix} (1 - \delta)^2 & (1 - \delta)\epsilon & (1 - \delta)\epsilon & \epsilon^2 \\ \delta(1 - \delta) & 0 & \delta\epsilon & 0 \\ \delta(1 - \delta) & \delta\epsilon & 0 & 0 \\ \delta^2 & 0 & 0 & 0 \end{pmatrix}, \quad (6.9)$$

$$D_{w-x,z-y} = (1 - \mu)^2(1 - \epsilon) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & (1 - \delta)p(y, z)\delta_{x,w} & 0 & \epsilon p(y, z)\delta_{x,w} \\ 0 & 0 & (1 - \delta)p(x, w)\delta_{y,z} & \epsilon p(x, w)\delta_{y,z} \\ 0 & \delta p(y, z)\delta_{x,w} & \delta p(x, w)\delta_{y,z} & (1 - \epsilon)p(x, w)p(y, z) \end{pmatrix}. \quad (6.10)$$

Proof. We begin by writing a recursion relation for $\Psi_{x,y}^{(n)}$, the probability at time n that two individuals randomly drawn from colonies x, y are identical by descent. This reads as follows:

$$\begin{aligned}
\psi_{n+1}((x, 0), (y, 0)) &= (1 - \mu)^2 \left[(1 - \delta)^2 \psi_n((x, 0), (y, 0)) + \epsilon^2 \psi_n((x, 1), (y, 1)) \right. \\
&\quad \left. + \epsilon(1 - \delta) [\psi_n((x, 1), (y, 0)) + \psi_n((x, 0), (y, 1))] \right], \\
\psi_{n+1}((x, 1), (y, 0)) &= (1 - \mu)^2 \left[\delta(1 - \delta) \psi_n((x, 0), (y, 0)) + \delta\epsilon \psi_n((x, 0), (y, 1)) \right. \\
&\quad \left. + \sum_{w \in \mathbb{T}^2} (1 - \epsilon)(1 - \delta) p(x, w) \psi_n((w, 1), (y, 0)) + \epsilon(1 - \epsilon) p(x, w) \psi_n((w, 1), (y, 1)) \right], \\
\psi_{n+1}((x, 0), (y, 1)) &= (1 - \mu)^2 \left[\delta(1 - \delta) \psi_n((x, 0), (y, 0)) + \delta\epsilon \psi_n((x, 1), (y, 0)) \right. \\
&\quad \left. + \sum_{z \in \mathbb{T}^2} (1 - \epsilon)(1 - \delta) p(y, z) \psi_n((x, 0), (z, 1)) + \epsilon(1 - \epsilon) p(y, z) \psi_n((x, 1), (z, 1)) \right], \\
\psi_{n+1}((x, 1), (y, 1)) &= (1 - \mu)^2 \left[\delta^2 \psi_n((x, 0), (y, 0)) \right. \\
&\quad + \sum_{z \in \mathbb{T}^2} \delta(1 - \epsilon) p(y, z) \psi_n((x, 0), (z, 1)) \\
&\quad + \sum_{w \in \mathbb{T}^2} \delta(1 - \epsilon) p(x, w) \psi_n((w, 1), (y, 0)) \\
&\quad + \sum_{\substack{w, z \in \mathbb{T}^2 \\ w \neq z}} (1 - \epsilon)^2 p(x, w) p(y, z) \psi_n((w, 1), (z, 1)) \\
&\quad \left. + \sum_{z \in \mathbb{T}^2} (1 - \epsilon)^2 p(x, z) p(y, z) \left[\frac{1}{N} + \left(1 - \frac{1}{N}\right) \psi_n((z, 1), (z, 1)) \right] \right]. \tag{6.11}
\end{aligned}$$

The reasoning behind the above expressions comes from considering the possible choices for the ancestors of the two individuals. In the first expression, for example, the drawn individuals are both in the seed-bank and therefore there are three options:

- they were both in the seed bank in the previous generation (with probability $1 - \delta$ each, independently from each other);
- they are both offspring of active individuals (with probability ϵ each);
- one was in the seed-bank in the previous generation and did not become active (with probability $1 - \delta$), and the other is offspring of an active individual (with probability ϵ).

Note that if individuals are in the seed-bank, then their ancestors cannot be from a different colony and therefore the transition kernels do not always appear in the recursive relations. When they appear, they are always multiplied by $1 - \epsilon$, since the transition kernels are probabilities *conditional* on the event that the individuals were not in the

seed bank in the previous generation. In the last term of $\psi_{n+1}((x, 1), (y, 1))$ we are looking at the event in which the ancestors of the two individuals are in the same colony: either they are the same individual (with probability $1/N$ and the iteration ends) or they are two distinct individuals (with probability $1 - \frac{1}{N}$). In this equation we add and subtract the term $(1 - \mu)^2(1 - \epsilon)^2 p(x, w)p(y, z)\psi_n((w, 1), (z, 1))$ (the expression in the fourth line) for $w = z$ to obtain

$$\begin{aligned}
\psi_{n+1}((x, 1), (y, 1)) &= (1 - \mu)^2 \left[\delta^2 \psi_n((x, 0), (y, 0)) \right. \\
&\quad + \sum_{z \in \mathbb{T}^2} \delta(1 - \epsilon)p(y, z)\psi_n((x, 0), (z, 1)) \\
&\quad + \sum_{w \in \mathbb{T}^2} \delta(1 - \epsilon)p(x, w)\psi_n((w, 1), (y, 0)) \\
&\quad + \sum_{w, z \in \mathbb{T}^2} (1 - \epsilon)^2 p(x, w)p(y, z)\psi_n((w, 1), (z, 1)) \\
&\quad \left. + \sum_{z \in \mathbb{T}^2} (1 - \epsilon)^2 p(x, z)p(y, z) \left[\frac{1}{N} - \frac{1}{N}\psi_n((z, 1), (z, 1)) \right] \right]. \tag{6.12}
\end{aligned}$$

By ergodicity we have

$$\lim_{n \rightarrow \infty} \Psi_{x,y}^{(n)} = \Psi_{x,y} \quad \forall x, y \in \mathbb{T}^2. \tag{6.13}$$

Therefore at equilibrium we may drop the time indices from the above expressions, which can be summarized in (4×4) -matrix form as

$$\begin{aligned}
&\begin{pmatrix} \psi((x, 0), (y, 0)) \\ \psi((x, 0), (y, 1)) \\ \psi((x, 1), (y, 0)) \\ \psi((x, 1), (y, 1)) \end{pmatrix} = (1 - \mu)^2 \begin{pmatrix} (1 - \delta)^2 & (1 - \delta)\epsilon & (1 - \delta)\epsilon & \epsilon^2 \\ \delta(1 - \delta) & 0 & \delta\epsilon & 0 \\ \delta(1 - \delta) & \delta\epsilon & 0 & 0 \\ \delta^2 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \psi((x, 0), (y, 0)) \\ \psi((x, 0), (y, 1)) \\ \psi((x, 1), (y, 0)) \\ \psi((x, 1), (y, 1)) \end{pmatrix} \\
&+ \sum_{w, z \in \mathbb{T}^2} (1 - \mu)^2(1 - \epsilon) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & (1 - \delta)p(y, z)\delta_{x,w} & 0 & \epsilon p(y, z)\delta_{x,w} \\ 0 & 0 & (1 - \delta)p(x, w)\delta_{y,z} & \epsilon p(x, w)\delta_{y,z} \\ 0 & \delta p(y, z)\delta_{x,w} & \delta p(x, w)\delta_{y,z} & (1 - \epsilon)p(x, w)p(y, z) \end{pmatrix} \begin{pmatrix} \psi((w, 0), (z, 0)) \\ \psi((w, 0), (z, 1)) \\ \psi((w, 1), (z, 0)) \\ \psi((w, 1), (z, 1)) \end{pmatrix} \\
&- \sum_{z \in \mathbb{T}^2} (1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p(x, z)p(y, z)\frac{1}{N} \end{pmatrix} \begin{pmatrix} \psi((z, 0), (z, 0)) \\ \psi((z, 0), (z, 1)) \\ \psi((z, 1), (z, 0)) \\ \psi((z, 1), (z, 1)) \end{pmatrix} \\
&\quad + \sum_{z \in \mathbb{T}^2} (1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ p(x, z)p(y, z)\frac{1}{N} \end{pmatrix}. \tag{6.14}
\end{aligned}$$

After using that $\sum_{z \in \mathbb{T}^2} p(x, z)p(y, z) = \sum_{z \in \mathbb{T}^2} p(x, z)p(z, y) = p^2(x, y)$, and $\Psi_{z,z} = \Psi_{0,0}$ by translation invariance, we obtain the expression in (6.5). \square

Now that we have a recursive relation for $\Psi_{x,y}$ as expressed in Proposition 6.1, we proceed to solve this relation to find a closed form expression for $\Psi_{x,y}$. To do so, we iterate (6.5) countably many times and, after noting that the last summand tends to 0 (since $1 - \mu < 1$), we obtain the following expression.

Proposition 6.2. For all $x, y \in \mathbb{T}^2$,

$$\Psi_{x,y} = \frac{1 - \Psi_{0,0}^{(4)}}{N} \sum_{n=0}^{\infty} (B * \Gamma)^{(n)}(x, y), \quad (6.15)$$

where $\Psi_{0,0}^{(4)}$ is the fourth entry of $\Psi_{0,0}$ and

$$(B * \Gamma)^{(n)}(x, y) = \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y}^{(n)} \Gamma_{w,z}, \quad (6.16)$$

$$B_{w-x, z-y}^{(n)} = \sum_{w', z' \in \mathbb{T}^2} B_{w'-x, z'-y}^{(n-1)} B_{w-w', z-z'}, \quad (6.17)$$

$$B_{w-x, z-y}^{(0)} = \mathbb{1} \delta_{x,w} \delta_{y,z}, \quad (6.18)$$

$$\Gamma_{w,z} = (1 - \mu)^2 (1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ p^2(w, z) \end{pmatrix}. \quad (6.19)$$

Proof. The claim follows from repeatedly substituting $\Psi_{w,z}$ in (6.5). In fact,

$$\begin{aligned} \Psi_{x,y} &= \Phi_{x,y} + A_{x,y} \Psi_{0,0} + \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y} \Psi_{w,z} \\ &= \Phi_{x,y} + D_{x,y} \Psi_{0,0} + \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y} \left(\Phi_{w,z} + A_{w,z} \Psi_{0,0} + \sum_{w', z' \in \mathbb{T}^2} B_{w'-z, z'-z} \Psi_{w', z'} \right) \\ &= \left(\mathbb{1} \delta_{xw} \delta_{yz} + \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y} \right) \Phi_{w,z} + \left(\mathbb{1} \delta_{xw} \delta_{yz} + \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y} \right) A_{w,z} \Psi_{0,0} + \\ &\quad + \sum_{w,z \in \mathbb{T}^2} \sum_{w', z' \in \mathbb{T}^2} B_{w-x, z-y} B_{w'-z, z'-z} \Psi_{w', z'} \\ &= \sum_{k=0}^1 \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y}^{(k)} (\Phi_{w,z} + A_{w,z} \Psi_{0,0}) \\ &\quad + \sum_{w'', z'' \in \mathbb{T}^2} B_{w''-x, z''-y}^{(2)} \Psi_{w'', z''}, \end{aligned} \quad (6.20)$$

where $B_{w-x, z-y}^{(0)}$ is defined by (6.18) and $B_{w-x, z-y}^{(n)}$ is defined recursively as in (6.17). After n substitutions we obtain

$$\begin{aligned} \Psi_{x,y} &= \sum_{k=0}^n \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y}^{(k)} (\Phi_{w,z} + A_{w,z} \Psi_{0,0}) \\ &\quad + \sum_{w'', z'' \in \mathbb{T}^2} B_{w''-x, z''-y}^{(n+1)} \Psi_{w'', z''}. \end{aligned} \quad (6.21)$$

Letting $n \rightarrow \infty$ and noticing that $\lim_{n \rightarrow \infty} B_{w-x, z-y}^{(n)} = 0$ (since each term is finite and multiplied by $(1 - \mu)^{2n}$ with $0 < \mu < 1$), we see that the summand in the second line of the last equality goes to 0. We therefore obtain

$$\Psi_{x,y} = \sum_{k=0}^{\infty} \sum_{w,z \in \mathbb{T}^2} B_{w-x, z-y}^{(k)} (\Phi_{w,z} + A_{w,z} \Psi_{0,0}). \quad (6.22)$$

We can rewrite

$$\Phi_{w,z} + A_{w,z}\Psi_{0,0} = (1-\mu)^2(1-\epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1-\Psi_{0,0}^{(4)}}{N}p^2(w,z) \end{pmatrix} = \frac{1-\Psi_{0,0}^{(4)}}{N}\Gamma_{w,z}, \quad (6.23)$$

with $\Gamma_{w,z}$ given by (6.19) and the claim follows. \square

Taking $x = y = 0$ in (6.22), we get

$$\Psi_{0,0} = \left(\mathbb{1} - \sum_{n=0}^{\infty} (B * A)^{(n)}(0,0) \right)^{-1} \left(\sum_{n=0}^{\infty} (B * \Phi)^{(n)}(0,0) \right). \quad (6.24)$$

This can be substituted into (6.15) to obtain an explicit expression for $\Psi_{x,y}$. Since this expression contains convolutions, we turn to Fourier analysis to gain more insight into the properties of $\Psi_{x,y}$.

6.3 Fourier analysis

Let $\hat{\mathbb{T}}^2 = \{0, \frac{1}{L}, \dots, \frac{L-1}{L}\}^2$. Define for $f : \mathbb{T}^2 \times \mathbb{T}^2 \mapsto \mathbb{R}$, $x, y \in \mathbb{T}^2$ and $\theta, \eta \in \hat{\mathbb{T}}^2$,

$$\hat{f}(\theta, \eta) = \sum_{x,y \in \mathbb{T}^2} f(x, y) e^{2\pi i(x \cdot \theta + y \cdot \eta)}, \quad (6.25)$$

$$f(x, y) = \frac{1}{|\hat{\mathbb{T}}^2|^2} \sum_{\theta, \eta \in \hat{\mathbb{T}}^2} \hat{f}(\theta, \eta) e^{-2\pi i(x \cdot \theta + y \cdot \eta)}. \quad (6.26)$$

Proposition 6.3. For $\theta, \eta \in \hat{\mathbb{T}}^2$,

$$\hat{\Psi}(\theta, \eta) = \frac{1-\Psi_{0,0}^{(4)}}{N} \left(\mathbb{1} - \hat{B}(\theta, \eta) \right)^{-1} \hat{\Gamma}(\theta, \eta). \quad (6.27)$$

Proof. By linearity of the Fourier transform and the convolution theorem, we get

$$\hat{\Psi}(\theta, \eta) = \frac{1-\Psi_{0,0}^{(4)}}{N} \sum_{n=0}^{\infty} \hat{B}^{(n)}(\theta, \eta) \hat{\Gamma}(\theta, \eta). \quad (6.28)$$

Since $B^{(n)}(x-w, y-z)$ is defined by the recursion in (6.17), we have

$$\hat{B}^{(n)}(\theta, \eta) = \hat{B}^{(n-1)}(\theta, \eta) \hat{B}(\theta, \eta) = \left(\hat{B}(\theta, \eta) \right)^n, \quad (6.29)$$

and therefore

$$\sum_{n=0}^{\infty} \hat{B}^{(n)}(\theta, \eta) \hat{\Gamma}(\theta, \eta) = \sum_{n=0}^{\infty} \left(\hat{B}(\theta, \eta) \right)^n \hat{\Gamma}(\theta, \eta) = \left(\mathbb{1} - \hat{B}(\theta, \eta) \right)^{-1} \hat{\Gamma}(\theta, \eta). \quad (6.30)$$

The claim follows after substitution of (6.30) into (6.28). \square

Our next objective is to compute the right-hand side of (6.27). Recalling the definition of $B_{w-x, z-y}$ through (6.8)–(6.10), we obtain

$$\hat{B}(\theta, \eta) = \sum_{u, v \in \mathbb{T}^2} B_{u, v} e^{2\pi i(\theta \cdot u + \eta \cdot v)} = \sum_{u, v \in \mathbb{T}^2} (C \delta_{u, 0} \delta_{v, 0} + D_{u, v}) e^{2\pi i(\theta \cdot u + \eta \cdot v)} = C + \hat{D}(\theta, \eta), \quad (6.31)$$

where

$$\hat{D}(\theta, \eta) = (1 - \mu)^2 (1 - \epsilon) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & (1 - \delta)\hat{p}(\eta) & 0 & \epsilon\hat{p}(\eta) \\ 0 & 0 & (1 - \delta)\hat{p}(\theta) & \epsilon\hat{p}(\theta) \\ 0 & \delta\hat{p}(\eta) & \delta\hat{p}(\theta) & (1 - \epsilon)\hat{p}(\theta)\hat{p}(\eta) \end{pmatrix}, \quad (6.32)$$

$$\hat{p}(\theta) = \sum_{z \in \mathbb{T}^2} p(0, z) e^{2\pi i \theta \cdot z}. \quad (6.33)$$

Computing $(\mathbb{1} - \hat{B}(\theta, \eta))^{-1}$, we find

$$(\mathbb{1} - \hat{B}(\theta, \eta))^{-1} = \frac{1}{r_0(\theta, \eta)} \begin{pmatrix} r_{1,1}(\theta, \eta) & r_{1,2}(\theta, \eta) & r_{1,3}(\theta, \eta) & r_{1,4}(\theta, \eta) \\ r_{2,1}(\theta, \eta) & r_{2,2}(\theta, \eta) & r_{2,3}(\theta, \eta) & r_{2,4}(\theta, \eta) \\ r_{3,1}(\theta, \eta) & r_{3,2}(\theta, \eta) & r_{3,3}(\theta, \eta) & r_{3,4}(\theta, \eta) \\ r_{4,1}(\theta, \eta) & r_{4,2}(\theta, \eta) & r_{4,3}(\theta, \eta) & r_{4,4}(\theta, \eta) \end{pmatrix}, \quad (6.34)$$

where the 17 functions in the above expression are polynomials of degree ≤ 4 in $\hat{p}(\theta)$ and $\hat{p}(\eta)$ whose coefficients depend on the known parameters δ, ϵ . Since $(\mathbb{1} - \hat{B}(\theta, \eta))^{-1}$ pre-multiplies $\hat{\Gamma}(\theta, \eta)$, whose first three entries are 0, we only need the entries in the fourth column. These are given by the following, where we abbreviate $m = (1 - \mu)^2$:

$$\begin{aligned} r_0(\theta, \eta) &= (1 - m^2 \delta^2 \epsilon^2) \left(1 - m \left(1 - \delta(2 - 2\epsilon - \delta(1 - m\epsilon^2)) \right) \right) \\ &\quad - m^2 (1 - \epsilon)^2 \hat{p}(\eta)^2 (\delta\epsilon - (1 - \delta)(1 - \epsilon)\hat{p}(\theta)) \left(1 - m(1 - \delta(2 - \delta - \epsilon)) - m(1 - m(1 - \delta)^2)(1 - \delta)(1 - \epsilon)\hat{p}(\theta) \right) \\ &\quad - m(1 - \epsilon)(1 - \delta) \left(1 - m \left(1 - \delta(2 - \epsilon + \delta(1 - m\epsilon^2(1 - m\delta\epsilon))) \right) \right) (\hat{p}(\eta) + \hat{p}(\theta)) \\ &\quad + m^2 \delta \epsilon (1 - \epsilon)^2 (1 - m(1 - \delta(2 - \delta - \epsilon))) \hat{p}(\theta)^2 \\ &\quad - m(1 - \epsilon)^2 \hat{p}(\eta) \hat{p}(\theta) \left[(1 - m(1 - \delta)^2)^2 - 2m\delta\epsilon + m^2 \delta^2 \epsilon^2 - m(1 - \delta)(1 - \epsilon)(1 - m(1 - \delta)^2(1 + m\delta\epsilon)) \hat{p}(\theta) \right], \\ r_{1,4}(\theta, \eta) &= m\epsilon^2 ((1 - m\delta\epsilon)^2 - m^2(1 - \delta)^2(1 - \epsilon)^2 \hat{p}(\theta) \hat{p}(\eta)), \\ r_{2,4}(\theta, \eta) &= m\epsilon \left(m\delta\epsilon(1 - \delta)(1 - m\delta\epsilon) + (1 - \epsilon)\hat{p}(\eta)(1 - m(1 - \delta(2 - \delta - \epsilon)) - m(1 - m(1 - \delta)^2)(1 - \delta)(1 - \epsilon)\hat{p}(\theta)) \right), \\ r_{3,4}(\theta, \eta) &= m\epsilon \left(m\delta\epsilon(1 - \delta)(1 - m\delta\epsilon) + (1 - \epsilon)\hat{p}(\theta)(1 - m(1 - \delta(2 - \delta - \epsilon)) - m(1 - m(1 - \delta)^2)(1 - \delta)(1 - \epsilon)\hat{p}(\eta)) \right), \\ r_{4,4}(\theta, \eta) &= -m^2 \delta \epsilon (1 - \delta)^2 (1 - m\delta\epsilon - (1 - \delta)(1 - \epsilon)\hat{p}(\eta)) \\ &\quad + \left(1 - m(1 - \delta(2 - \delta - \epsilon)) - m(1 - m(1 - \delta)^2)(1 - \delta)(1 - \epsilon)\hat{p}(\eta) \right) (1 - m\delta\epsilon - m(1 - \delta)(1 - \epsilon)\hat{p}(\theta)). \end{aligned} \quad (6.35)$$

Looking at $\hat{\Gamma}(\theta, \eta)$, we have

$$\begin{aligned}
\hat{\Gamma}(\theta, \eta) &= (1 - \mu)^2 \sum_{w, z \in \mathbb{T}^2} \Gamma_{w, z} e^{2\pi i(\theta \cdot w + \eta \cdot z)} \\
&= (1 - \mu)^2 \sum_{w, z \in \mathbb{T}^2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ (1 - \epsilon)^2 \frac{1}{N} p^2(w, z) \end{pmatrix} e^{2\pi i(\theta \cdot w + \eta \cdot z)} \\
&= \frac{1}{N} (1 - \mu)^2 (1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sum_{w, z \in \mathbb{T}^2} \sum_{u \in \mathbb{T}^2} p(w, u) p(u, z) e^{2\pi i(\theta \cdot w + \eta \cdot z)} e^{2\pi i(\eta \cdot u - \eta \cdot u + \theta \cdot u - \theta \cdot u)} \end{pmatrix} \\
&= \frac{1}{N} (1 - \mu)^2 (1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sum_{u \in \mathbb{T}^2} \sum_{w \in \mathbb{T}^2} p(w, u) e^{2\pi i\theta \cdot (w - u)} \sum_{z \in \mathbb{T}^2} p(u, z) e^{2\pi i\eta \cdot (z - u)} e^{2\pi i(\theta + \eta) \cdot u} \end{pmatrix} \\
&= \frac{1}{N} (1 - \mu)^2 (1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sum_{u \in \mathbb{T}^2} \sum_{w - u \in \mathbb{T}^2} p(w - u, 0) e^{2\pi i\theta \cdot (w - u)} \sum_{z - u \in \mathbb{T}^2} p(0, z - u) e^{2\pi i\eta \cdot (z - u)} e^{2\pi i(\theta + \eta) \cdot u} \end{pmatrix} \\
&= \frac{1}{N} (1 - \mu)^2 (1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sum_{u \in \mathbb{T}^2} \hat{p}(\theta) \hat{p}(\eta) e^{2\pi i(\theta + \eta) \cdot u} \end{pmatrix} \\
&= \frac{1}{N} (1 - \mu)^2 (1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ \hat{p}(\theta) \hat{p}(\eta) \end{pmatrix} \delta_{\theta, -\eta}. \tag{6.36}
\end{aligned}$$

Therefore we have that

$$\begin{aligned}
(\mathbb{1} - \hat{B}(\theta, \eta))^{-1} \hat{\Gamma}(\theta, \eta) &= \frac{(1 - \mu)^2 (1 - \epsilon)^2}{N} \frac{1}{r_0(\theta, \eta)} \begin{pmatrix} r_{1,4}(\theta, \eta) \\ r_{2,4}(\theta, \eta) \\ r_{3,4}(\theta, \eta) \\ r_{4,4}(\theta, \eta) \end{pmatrix} \hat{p}(\theta) \hat{p}(\eta) \delta_{\theta, -\eta} \\
&= \frac{(1 - \mu)^2 (1 - \epsilon)^2}{N} \frac{1}{r_0} \begin{pmatrix} r_{1,4} \\ r_{2,4} \\ r_{3,4} \\ r_{4,4} \end{pmatrix} \hat{p}(\theta)^2 \delta_{\theta, -\eta}, \tag{6.37}
\end{aligned}$$

where by $r_{i,j}$ we mean $r_{i,j}(\theta, -\theta)$ given in (6.35) and we use that $\hat{p}(\theta) \hat{p}(\eta) \delta_{\theta, -\eta} = \hat{p}(\theta) \hat{p}(-\theta) \delta_{\theta, -\eta} = \hat{p}(\theta)^2 \delta_{\theta, -\eta}$. Indeed,

$$\hat{p}(-\theta) = \sum_{z \in \mathbb{T}^2} p(0, z) e^{2\pi i(-\theta) \cdot z} = \sum_{z \in \mathbb{T}^2} p(0, -z) e^{2\pi i\theta \cdot (-z)} = \hat{p}(\theta), \tag{6.38}$$

where in the second equality we use symmetry of the transition kernel.

Notice that, since $\hat{\Gamma}(\theta, \eta)$ contains the multiplicative factor $\delta_{\theta, -\eta}$, $\hat{\Psi}(\theta, \eta)$ actually depends on θ only. This is in agreement with the fact that $\Psi_{x,y}$ actually depends on $x - y$ only. All the terms in (6.27) are known objects at this

point, except for $\Psi_{0,0}$. In order to express $\Psi_{0,0}$, we take the Fourier transform of (6.22) to obtain

$$\hat{\Psi}(\theta, \eta) = (\mathbb{1} - \hat{B}(\theta, \eta))^{-1} \left(\hat{\Phi}(\theta, \eta) + \hat{A}(\theta, \eta) \Psi_{0,0} \right). \quad (6.39)$$

We then use the Fourier inversion formula (6.26), which gives:

$$\frac{1}{|\hat{\mathbb{T}}^2|^2} \sum_{\theta, \eta \in \hat{\mathbb{T}}^2} \hat{\Psi}(\theta, \eta) = \Psi_{0,0}. \quad (6.40)$$

Notice that both $\hat{\Phi}(\theta, \eta)$ and $\hat{A}(\theta, \eta)$ contain the multiplicative factor $\delta_{\theta, -\eta}$, therefore we can take $\eta = -\theta$. Substitution of (6.39) thus yields

$$\Psi_{0,0} = \left(\mathbb{1} - \frac{1}{|\hat{\mathbb{T}}^2|} \sum_{\theta \in \hat{\mathbb{T}}^2} (\mathbb{1} - \hat{B}(\theta, -\theta))^{-1} \hat{A}(\theta, -\theta) \right)^{-1} \left(\frac{1}{|\hat{\mathbb{T}}^2|} \sum_{\theta \in \hat{\mathbb{T}}^2} (\mathbb{1} - \hat{B}(\theta, -\theta))^{-1} \hat{\Phi}(\theta, -\theta) \right). \quad (6.41)$$

The result is a 4-vector that can be substituted into (6.27) to obtain an explicit expression for $\hat{\Psi}(\theta, -\theta)$. With (6.37) and (6.41) we have obtained an explicit, albeit complicated, expression for $\hat{\Psi}(\theta, -\theta)$. In order to obtain more insight into its properties, we look at how it behaves for a special choice of the parameters, namely, $M = N$, $\delta = \epsilon \ll 1$. By what we have observed above, in the following, we write $\hat{\Psi}(\theta)$ instead of $\hat{\Psi}(\theta, -\theta)$, and similarly for the other terms.

6.4 Special choice of parameters

In order to simplify the polynomials involved in (6.37) and (6.41), we consider the special case $M = N$, for which $\delta = \epsilon \frac{M}{N} = \epsilon$, and suppose that $0 < \delta \ll 1$. In this case, the objects in (6.41) become equal to the expressions in (6.43) and (6.44), with the abbreviations:

$$\begin{aligned} p &= \hat{p}(\theta), & a &= \sum_{\theta \in \hat{\mathbb{T}}^2} \frac{mp^2}{1 - mp^2}, \\ m &= (1 - \mu)^2, & b &= \sum_{\theta \in \hat{\mathbb{T}}^2} \frac{m^2 p^3}{(1 - mp)(1 - mp^2)}, \\ & & c &= \sum_{\theta \in \hat{\mathbb{T}}^2} \frac{mp^2}{(1 - mp^2)^2}, \end{aligned} \quad (6.42)$$

namely,

$$\sum_{\theta \in \hat{\mathbb{T}}^2} (\mathbb{1} - \hat{B}(\theta))^{-1} \hat{\Phi}(\theta) = \frac{1}{N} \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ a \end{pmatrix} + \delta \begin{pmatrix} 0 \\ b \\ b \\ -2c \end{pmatrix} \right] + O(\delta^2), \quad (6.43)$$

$$\sum_{\theta \in \hat{\mathbb{T}}^2} (\mathbb{1} - \hat{B}(\theta))^{-1} \hat{A}(\theta) = -\frac{1}{N} \left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & -2c \end{pmatrix} \right] + O(\delta^2). \quad (6.44)$$

Substituting the above in (6.41) yields

$$\Psi_{0,0} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{a'}{1+a'} \end{pmatrix} + \delta \begin{pmatrix} 0 \\ \frac{b'}{1+a'} \\ \frac{b'}{1+a'} \\ -2 \frac{c'}{(1+a')^2} \end{pmatrix} + O(\delta^2), \quad (6.45)$$

where we renamed

$$a' = \frac{a}{NL^4}, \quad b' = \frac{b}{NL^4}, \quad c' = \frac{c}{NL^4}, \quad d' = \frac{d}{NL^4}. \quad (6.46)$$

For later use we look at the limiting case $\delta = 0$, for which we obtain

$$\sum_{\theta \in \hat{\mathbb{T}}^2} (\mathbb{1} - \hat{B}(\theta))^{-1} \hat{\Phi}(\theta) = \frac{1}{N} \begin{pmatrix} 0 \\ 0 \\ 0 \\ a \end{pmatrix} + O(\delta), \quad (6.47)$$

$$\sum_{\theta \in \hat{\mathbb{T}}^2} (\mathbb{1} - \hat{B}(\theta))^{-1} \hat{A}(\theta) = -\frac{1}{N} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \end{pmatrix} + O(\delta). \quad (6.48)$$

Notice that both a and b depend only on ν and μ . Then

$$\Psi_{0,0} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{a'}{1+a'} \end{pmatrix}. \quad (6.49)$$

This result is consistent with the intuition that, when $\delta = 0$, individuals cannot leave nor enter the seed-bank. In this case, in fact, when we pick two individuals uniformly at random from the population at generation 0 and look backwards in time, the probability that their lineages coalesce is zero if at least one of them is in the seed-bank at generation 0.

We now know all the quantities involved in (6.27) up to the first order of δ and can substitute to obtain the behaviour of $\Psi_{0,x}$ for any $x \in \mathbb{T}^2$.

Proposition 6.4. *For all $x \in \mathbb{T}^2$, $\epsilon = \delta \ll 1$, $\mu \in [0, 1)$, $\nu \in [0, 1]$,*

$$\Psi_{0,x} = \frac{1}{N} c_1 \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha(x) \end{pmatrix} + \delta \begin{pmatrix} 0 \\ \beta(x) \\ \beta(x) \\ 2c_2 \alpha(x) - 2\gamma(x) \end{pmatrix} \right] + O(\delta^2), \quad (6.50)$$

where

$$\begin{aligned} \alpha(x) &= \frac{1}{L^4} \sum_{\theta \in \hat{\mathbb{T}}^2} \frac{(1-\mu)^2 \hat{p}(\theta)^2}{1 - (1-\mu)^2 \hat{p}(\theta)^2} e^{-2\pi i \theta \cdot x}, & c_1 &= \frac{1}{1+a'}, \\ \beta(x) &= \frac{1}{L^4} \sum_{\theta \in \hat{\mathbb{T}}^2} \frac{(1-\mu)^4 \hat{p}(\theta)^3}{(1 - (1-\mu)^2 \hat{p}(\theta))(1 - (1-\mu)^2 \hat{p}(\theta)^2)} e^{-2\pi i \theta \cdot x}, & c_2 &= \frac{c'}{1+a'}, \\ \gamma(x) &= \frac{1}{L^4} \sum_{\theta \in \hat{\mathbb{T}}^2} \frac{(1-\mu)^2 \hat{p}(\theta)^2}{(1 - (1-\mu)^2 \hat{p}(\theta)^2)^2} e^{-2\pi i \theta \cdot x}. \end{aligned} \quad (6.51)$$

Proof. We compute the Taylor expansions of (6.37) in δ and obtain

$$(\mathbb{1} - \hat{B}(\theta))^{-1} \hat{\Gamma}(\theta) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{mp^2}{1-mp^2} \end{pmatrix} + \delta \begin{pmatrix} 0 \\ \frac{m^2 p^3}{(1-mp)(1-mp^2)} \\ \frac{m^2 p^3}{(1-mp)(1-mp^2)} \\ -2 \frac{mp^2}{(1-mp^2)^2} \end{pmatrix} + O(\delta^2), \quad (6.52)$$

(6.53)

Substitution of (6.52) and (6.45) into (6.27) and taking the inverse Fourier transform yields the result. \square

Notice that the only non-zero term in the first vector in (6.50) coincides with the probability of being identical by descent in the standard Wright-Fisher model. In fact, the first vector corresponds to the limiting case $\delta = 0$, namely, the case without any exchange between the active population and the seed-bank. The first three entries are zero because, if at least one individual is in the seed-bank, then there cannot be coalescence of the lineages. The fourth is the only non-zero term and, since active individuals remain active at all times in this limiting case, the system behaves exactly like the standard Wright-Fisher model and thus the probabilities are the same.

To gain more insight into the behaviour of $\alpha(x)$, $\beta(x)$ and $\gamma(x)$, we make the following observation.

Proposition 6.5. *For all $x \in \mathbb{T}^2$,*

$$\begin{aligned} \alpha(x) &= \frac{1}{L^4} \sum_{l=0}^{\infty} c_l^1 q_l(x), \\ \beta(x) &= \frac{1}{L^4} \sum_{l=0}^{\infty} c_l^2 q_l(x), \\ \gamma(x) &= \frac{1}{L^4} \sum_{l=0}^{\infty} c_l^3 q_l(x). \end{aligned} \quad (6.54)$$

where the constants c_l depend on μ and ν only (see below) and $q_l(x)$ is the probability that the simple random walk with transition kernel $q(x)$ visits site x at the l -th step.

The sums involved in (6.54) relate to the average number of visits that a random walk pays to site x . The constants contain information regarding the properties of the random walk, namely, the fact that it does not change position after the previous step with probability $1 - \nu$ and that it is terminated at every step with probability $1 - (1 - \mu)^2$. For example, $\sum_{l=0}^{\infty} (1 - \mu) q_l(x)$ for an arbitrary transition kernel $q(x)$ is the average number of visits paid to site x by a random walker who moves according to a transition kernel $q(x)$ and is terminated with probability μ at every step.

Proof. We will prove the result for $\alpha(x)$. The other equalities follow from similar arguments. Write \hat{q} for $\hat{q}(\theta)$, \hat{p} for

$\hat{p}(\theta)$ and rewrite

$$\begin{aligned}
\frac{\hat{p}^2}{1 - (1 - \mu)^2 \hat{p}^2} &= \hat{p}^2 \frac{1}{1 - (1 - \mu)^2 \hat{p}^2} \\
&= \hat{p}^2 \sum_{k=0}^{\infty} (1 - \mu)^{2k} \hat{p}^{2k} \\
&= \sum_{k=0}^{\infty} (1 - \mu)^{2k} \hat{p}^{2k+2} \\
&= \sum_{k=0}^{\infty} (1 - \mu)^{2k} [(1 - \nu) + \nu \hat{q}]^{2k+2} \\
&= \sum_{k=0}^{\infty} (1 - \mu)^{2k} \sum_{l=0}^{2k+2} \binom{2k+2}{l} (1 - \nu)^{2k+2-l} \nu^l \hat{q}^l \\
&= \sum_{l=0}^{\infty} \sum_{k=\max\{\lceil \frac{l-2}{2} \rceil, 0\}}^{\infty} \binom{2k+2}{l} (1 - \mu)^{2k} (1 - \nu)^{2k+2-l} \nu^l \hat{q}^l \\
&= \sum_{l=0}^{\infty} c_l^1 \hat{q}^l.
\end{aligned} \tag{6.55}$$

By linearity of the Fourier transform, $\mathcal{F}^{-1}(\sum_{l=0}^{\infty} c_l^1 \hat{q}^l(\theta)) = \sum_{l=0}^{\infty} c_l^1 q_l(x)$ and the result follows. The other constants involved in (6.54) are given by

$$\begin{aligned}
c_l^1 &= \sum_{k=\max\{\lceil \frac{l-2}{2} \rceil, 0\}}^{\infty} \binom{2k+2}{l} (1 - \mu)^{2k} (1 - \nu)^{2k+2-l} \nu^l, \\
c_l^2 &= \sum_{k=0}^l c_k^1 \tilde{c}_{l-k}^2, \quad \tilde{c}_l^2 = \sum_{k=\max\{\lceil \frac{l-1}{2} \rceil, 0\}}^{\infty} \binom{2k+1}{l} (1 - \mu)^{2k} (1 - \nu)^{2k+1-l} \nu^l, \\
c_l^3 &= \sum_{k=0}^l \tilde{c}_k^3 \tilde{c}_{l-k}^3, \quad \tilde{c}_l^3 = \sum_{k=\max\{\lceil \frac{l-2}{2} \rceil, 0\}}^{\infty} \binom{2k+2}{l} (1 - \mu)^{2k} (1 - \nu)^{2k+2-l} \nu^l.
\end{aligned} \tag{6.56}$$

□

6.5 Perturbation

In this section we identify the behaviour of $\hat{\Psi}(\theta)$ for $|\theta| \ll 1$, which requires $L \gg 1$. This will allow us to compute the second moment $\sum_{x \in \mathbb{T}^2} |x|^2 \Psi_{0,x}$ of the probability of being identical by descent in closed form for all choices of the parameters. Let

$$\hat{B}(\theta) = \hat{B}(0) + \Delta(\theta) + o(|\theta|^2), \tag{6.57}$$

where $\Delta(\theta)$ is a perturbation with $0 \leq |\theta| \ll 1$. Specifically, for the choice of $p(x, y)$ in (6.1) and (6.2), we have

$$\hat{p}(\theta) = 1 - \nu \pi^2 |\theta|^2 + o(|\theta|^2). \tag{6.58}$$

Indeed,

$$\hat{p}(\theta) = \sum_{z \in \mathbb{T}^2} p(0, z) e^{2\pi i \theta \cdot z} = \sum_{z \in \mathbb{T}^2} [(1 - \nu) \delta_{z,0} + \nu q(z)] e^{2\pi i \theta \cdot z} = (1 - \nu) + \nu \hat{q}(\theta) = 1 - \nu(1 - \hat{q}(\theta)) = 1 - \nu \pi^2 |\theta|^2 + o(|\theta|^2), \quad (6.59)$$

since

$$\hat{q}(\theta) = 1 - \pi^2 |\theta|^2 + o(|\theta|^2), \quad |\theta| \ll 1. \quad (6.60)$$

When we substitute this into (6.8)–(6.10), we get

$$\Delta(\theta) = -\Delta_0 \nu \pi^2 |\theta|^2 + o(|\theta|^2), \quad (6.61)$$

$$\Delta_0 = m(1 - \epsilon) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 - \delta & 0 & \epsilon \\ 0 & 0 & 1 - \delta & \epsilon \\ 0 & \delta & \delta & 2(1 - \epsilon) \end{pmatrix}. \quad (6.62)$$

We therefore obtain the following result on the perturbation of $\hat{\Psi}(\theta)$.

Proposition 6.6. For $0 \leq |\theta| \ll 1$,

$$\partial \hat{\Psi}(\theta) = \frac{1 - \Psi_{0,0}^{(4)}}{N} \left[U \partial \hat{\Gamma}(\theta) + U \Delta(\theta) U \hat{\Gamma}(0) \right] + o(|\theta|^2), \quad (6.63)$$

where

$$\partial \hat{\Psi}(\theta) = \hat{\Psi}(\theta) - \hat{\Psi}(0) \quad (6.64)$$

is the object of interest and

$$\begin{aligned} U &= (\mathbf{1} - \hat{B}(0))^{-1}, \\ \partial \hat{\Gamma}(\theta) &= \hat{\Gamma}(\theta) - \hat{\Gamma}(0), \end{aligned} \quad (6.65)$$

are known quantities.

Proof. Using (6.57), we obtain after a simple computation that

$$\begin{aligned} (\mathbf{1} - \hat{B}(\theta))^{-1} &= \sum_{n=0}^{\infty} \hat{B}(\theta)^n \\ &= \sum_{n=0}^{\infty} (\hat{B}(0) + \Delta(\theta))^n \\ &= \sum_{n=0}^{\infty} \hat{B}(0)^n + \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} (\hat{B}(0))^{n-1-k} \Delta(\theta) (\hat{B}(0))^k + o(|\theta|^2) \\ &= \sum_{n=0}^{\infty} \hat{B}(0)^n + \left(\sum_{n=0}^{\infty} (\hat{B}(0))^n \right) \Delta(\theta) \left(\sum_{k=0}^{\infty} (\hat{B}(0))^k \right) + o(|\theta|^2) \\ &= (\mathbf{1} - \hat{B}(0))^{-1} + (\mathbf{1} - \hat{B}(0))^{-1} \Delta(\theta) (\mathbf{1} - \hat{B}(0))^{-1} + o(|\theta|^2) \\ &= (\mathbf{1} - \hat{B}(0))^{-1} \left(\mathbf{1} + \Delta(\theta) (\mathbf{1} - \hat{B}(0))^{-1} \right) + o(|\theta|^2). \end{aligned} \quad (6.66)$$

Substitution into (6.27) yields

$$\hat{\Psi}(\theta) - \hat{\Psi}(0) = \frac{1 - \Psi_{0,0}^{(4)}}{N} \left(\left(\mathbb{1} - \hat{B}(0) \right)^{-1} \left[\hat{\Gamma}(\theta) - \hat{\Gamma}(0) \right] + \left(\mathbb{1} - \hat{B}(0) \right)^{-1} \Delta(\theta) \left(\mathbb{1} - \hat{B}(0) \right)^{-1} \hat{\Gamma}(\theta) \right) + o(|\theta|^2). \quad (6.67)$$

For the second term on the left-hand side, we use that $\hat{\Gamma}(\theta) \sim \hat{\Gamma}(0)$ for $0 \leq |\theta| \ll 1$. After renaming the different objects according to (6.65), we obtain (6.63). \square

We would like to get some insight into the quantities involved in (6.63), namely, the objects in (6.65). Computing $(\mathbb{1} - \hat{B}(0))^{-1}$, we obtain

$$U = \left(\mathbb{1} - \hat{B}(0) \right)^{-1} = \frac{1}{u_0} \begin{pmatrix} u_{1,1} & u_{1,2} & u_{1,3} & u_{1,4} \\ u_{2,1} & u_{2,2} & u_{2,3} & u_{2,4} \\ u_{3,1} & u_{3,2} & u_{3,3} & u_{3,4} \\ u_{4,1} & u_{4,2} & u_{4,3} & u_{4,4} \end{pmatrix} = \frac{1}{u_0} \begin{pmatrix} t_1 & \epsilon s_1 & \epsilon s_1 & \epsilon^2 s_2 \\ \delta s_1 & t_2 & t_3 & \epsilon s_3 \\ \delta s_1 & t_3 & t_2 & \epsilon s_3 \\ \delta^2 s_2 & \delta s_3 & \delta s_3 & t_4 \end{pmatrix}, \quad (6.68)$$

where,

$$\begin{aligned} u_0 &= (1 - m)(1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon)(1 - m(1 - \delta - \epsilon))(1 - m(1 - \delta - \epsilon)^2), \\ s_1 &= m(1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon)(1 - \delta - m(1 - \epsilon)(1 - \delta - \epsilon)), \\ s_2 &= m(1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon)(1 + m(1 - \delta - \epsilon)), \\ s_3 &= m(1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon)(1 - \epsilon - m(1 - \epsilon)(1 - \delta - \epsilon)), \\ t_1 &= (1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon)(1 - m(1 - \epsilon)^2)(1 - m(1 - \delta - \epsilon)), \\ t_2 &= 1 - m \left[3(1 - \delta - \epsilon) + (\delta + \epsilon)^2 - m(1 - 6\delta + 4\delta^2 - \delta^3 + 1 - 6\epsilon + 4\epsilon^2 - \epsilon^3 \right. \\ &\quad \left. + 1 - \delta^3\epsilon + 5\delta^2\epsilon + 10\delta\epsilon - 5\delta\epsilon^2 + \delta\epsilon^3) - m^2(1 - \delta - \epsilon)(\delta\epsilon + (1 - \delta - \epsilon)^2) \right], \\ t_3 &= m^2\delta\epsilon \left[(1 - \epsilon)^2 + (1 - \delta)^2 - 2m(1 - \delta)(1 - \epsilon)(1 - \delta - \epsilon) \right], \\ t_4 &= (1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon) \left[(1 - m(1 - \delta)^2)(1 - m(1 - \delta - \epsilon)) - 2m\delta\epsilon \right]. \end{aligned} \quad (6.69)$$

$$t_4 = (1 - m(1 - \delta - \epsilon) - 2m\delta\epsilon) \left[(1 - m(1 - \delta)^2)(1 - m(1 - \delta - \epsilon)) - 2m\delta\epsilon \right]. \quad (6.70)$$

Notice that U depends on the parameters δ, ϵ, μ only. We observe that there are several symmetries in the entries of the matrix, which come from the symmetries in the model.

As far as $\partial\hat{\Gamma}(\theta)$ and $\hat{\Gamma}(0)$ are concerned, they are easily computable. Recalling (6.36), we obtain

$$\partial\hat{\Gamma}(\theta) = -(1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 - \hat{p}(\theta)^2 \end{pmatrix}, \quad (6.71)$$

$$\hat{\Gamma}(0) = \Gamma_0 = (1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (6.72)$$

Substitution of (6.58) into (6.73) yields

$$\partial\hat{\Gamma}(\theta) = -(1 - \mu)^2(1 - \epsilon)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \nu\pi^2|\theta|^2 + o(|\theta|^2) = -2\Gamma_0\nu\pi^2|\theta|^2 + o(|\theta|^2), \quad (6.73)$$

Combining (6.61), (6.63), (6.72) and (6.73), we arrive at the following proposition, which is our main result for the perturbation of $\partial\hat{\Psi}(\theta)$.

Proposition 6.7. For $0 \leq |\theta| \ll 1$,

$$\partial \hat{\Psi}(\theta) = -\zeta \pi^2 |\theta|^2 + o(|\theta|^2), \quad (6.74)$$

where

$$\zeta = \frac{1 - \Psi_{0,0}^{(4)}}{N} \nu U (2\mathbb{1} + \Delta_0 U) \Gamma_0 \quad (6.75)$$

is a 4-vector whose entries depend on the parameters $\delta, \epsilon, \mu, \nu, N$ and the constant $\Psi_{0,0}$.

Notice that ζ is a known object given the parameters. In fact, $\delta, \epsilon, \mu, \nu, N$ are known in the model, whereas for $\Psi_{0,0}$ we don't have an explicit formula but it is, in principle, computable. We refer to equation (6.41) for a general formula (namely, valid for all values of the parameters) and to section 6.4 for an explicit computation of $\Psi_{0,0}$ up to the first order of δ , in the special case $\delta = \epsilon$, $\delta \ll 1$.

6.6 Asymptotic analysis

We will now use the results of Propositions 6.6 and 6.7 to recover information on the behaviour of $\Psi_{0,x}$. In particular, we will compute its second moment and analyse its behaviour for large values of $|x|$.

Recalling the definition of the Fourier transform given in (6.25), we have for $0 \leq |\theta| \ll 1$,

$$\begin{aligned} \hat{\Psi}(\theta) &= \sum_{x \in \mathbb{T}^2} \Psi_{0,x} e^{2\pi i(x_1 \theta_1 + x_2 \theta_2)} \\ &= \sum_{x \in \mathbb{T}^2} \Psi_{0,x} \cos(2\pi(x_1 \theta_1 + x_2 \theta_2)) \\ &= \sum_{x \in \mathbb{T}^2} \Psi_{0,x} \left(1 - \frac{4\pi^2(x_1 \theta_1 + x_2 \theta_2)^2}{2} \right) + o(|\theta|^2) \\ &= \hat{\Psi}(0) - 2\pi^2 \sum_{x \in \mathbb{T}^2} \Psi_{0,x} (x_1^2 \theta_1^2 + x_2^2 \theta_2^2 + 2x_1 x_2 \theta_1 \theta_2) + o(|\theta|^2), \end{aligned} \quad (6.76)$$

where $x = (x_1, x_2) \in \mathbb{T}^2$ and $\theta = (\theta_1, \theta_2) \in \hat{\mathbb{T}}^2$. In other words,

$$\hat{\Psi}(\theta) - \hat{\Psi}(0) = -2\pi^2 \sum_{x \in \mathbb{T}^2} \Psi_{0,x} (x_1^2 \theta_1^2 + x_2^2 \theta_2^2 + 2x_1 x_2 \theta_1 \theta_2) + o(|\theta|^2). \quad (6.77)$$

By what we have proved in Proposition 6.7 about the asymptotics of the left-hand side, we have

$$-\zeta \pi^2 (\theta_1^2 + \theta_2^2) + o(|\theta|^2) = -2\pi^2 \sum_{x \in \mathbb{T}^2} \Psi_{0,x} (x_1^2 \theta_1^2 + x_2^2 \theta_2^2 + 2x_1 x_2 \theta_1 \theta_2) + o(|\theta|^2). \quad (6.78)$$

If we choose $\theta \in \hat{\mathbb{T}}^2$ such that $\theta_1 = \theta_2$, then we obtain:

$$\begin{aligned} 2\zeta \theta_1^2 &= 2 \sum_{x \in \mathbb{T}^2} \Psi_{0,x} (x_1^2 \theta_1^2 + x_2^2 \theta_1^2 + 2x_1 x_2 \theta_1^2), \\ \zeta \theta_1^2 &= \sum_{x \in \mathbb{T}^2} \Psi_{0,x} (x_1^2 + x_2^2 + 2x_1 x_2) \theta_1^2, \\ \zeta &= \sum_{x \in \mathbb{T}^2} (x_1^2 + x_2^2 + 2x_1 x_2) \Psi_{0,x}. \end{aligned} \quad (6.79)$$

Notice that

$$\sum_{x \in \mathbb{T}^2} \Psi_{0,x} x_1 x_2 = \begin{cases} 0 & \text{if } L \text{ is odd,} \\ \Psi_{0,(L,L)} L^2 & \text{if } L \text{ is even.} \end{cases} \quad (6.80)$$

It is reasonable to assume that $\Psi_{0,(L,L)}L^2 \rightarrow 0$ as $L \rightarrow \infty$, but since the expression we have for $\Psi_{0,x}$ is too complex to prove it, we will state the following proposition only for odd values of L .

Proposition 6.8. *For all δ, ϵ, μ, N and L odd,*

$$\sum_{x \in \mathbb{T}^2} |x|^2 \Psi_{0,x} = \zeta. \quad (6.81)$$

We have therefore obtained that ζ equals exactly the second moment of $\Psi_{0,x}$. Propositions 6.7 and 6.8 give us a formula for the second moment of $\Psi_{0,x}$, except for the identification of the constant $\Psi_{0,0}$, for all values of the parameters. It is for this reason that in the next section we focus on the regime $0 < \delta, \mu \ll 1$.

First, however, we try to use the information of Proposition 6.7 to recover the asymptotic behaviour of $\Psi_{0,x}$ for large values of $|x|$. To do this, we use the inverse Fourier transform and substitute $\hat{\Psi}(\theta) = \hat{\Psi}(0) - \zeta\pi^2|\theta|^2 + o(|\theta|^2)$ to obtain:

$$\begin{aligned} \Psi_{0,x} &= \frac{1}{|\hat{\mathbb{T}}^2|} \sum_{\theta \in \hat{\mathbb{T}}^2} \hat{\Psi}(\theta) e^{-2\pi i(x \cdot \theta)} \\ &= \frac{1}{|\hat{\mathbb{T}}^2|} \sum_{\theta \in \hat{\mathbb{T}}^2} (\hat{\Psi}(0) - \zeta\pi^2|\theta|^2) e^{-2\pi i(x \cdot \theta)} + h.o.t. \\ &= \hat{\Psi}(0)\delta_{x,0} - \sum_{\theta \in \hat{\mathbb{T}}^2} \zeta\pi^2|\theta|^2 e^{-2\pi i(x \cdot \theta)} + h.o.t. \\ &= \hat{\Psi}(0)\delta_{x,0} - \zeta\pi^2 \sum_{\theta \in \hat{\mathbb{T}}^2} |\theta|^2 e^{-2\pi i(x \cdot \theta)} + h.o.t.. \end{aligned} \quad (6.82)$$

Since we are interested in the behaviour of $\Psi_{0,x}$ for large values of x , we ignore the first term and focus on the sum on the right-hand side. In one dimension, we could proceed by writing:

$$\sum_{\theta \in \hat{\mathbb{T}}} \theta^2 e^{-2\pi i(x\theta)} = \frac{1}{x^2} \sum_{\theta \in \hat{\mathbb{T}}} \theta^2 x^2 e^{-2\pi i(x\theta)} = \frac{1}{x^2} \sum_{\eta \in x \cdot \hat{\mathbb{T}}} \eta^2 e^{-2\pi i\eta} \sim \frac{1}{x^2} \sum_{\eta \in x \cdot \hat{\mathbb{T}}} \eta^2 \cos(-2\pi\eta), \quad (6.83)$$

where $x \cdot \hat{\mathbb{T}} = \{0, \frac{x}{L}, \dots, x \frac{L-1}{L}\}$. Next, we could approximate the sum by the corresponding integral divided by the length of each segment in $\hat{\mathbb{T}}$, thus obtaining

$$\frac{1}{x^2} \sum_{\eta \in x \cdot \hat{\mathbb{T}}} \eta^2 \cos(-2\pi\eta) \sim \frac{L}{x} \frac{1}{x^2} \int_0^x d\eta \eta^2 \cos(-2\pi\eta). \quad (6.84)$$

Integration by parts yields that the right-hand side equals

$$\frac{L}{x^3} \int_0^x d\eta \eta^2 \cos(2\pi\eta) = L \frac{(2\pi^2 x^2 - 1) \sin(2\pi x) + 2\pi x \cos(2\pi x)}{4\pi^3 x^3}. \quad (6.85)$$

The function on the right-hand side goes to zero as $x \rightarrow \infty$, as we would intuitively expect, but it oscillates between positive and negative values, which is in contrast with the fact that it should represent a probability distribution. This means that the approximation of $\hat{\Psi}(\theta)$ by up to second order does not provide enough information about the behaviour of $\Psi_{0,x}$. Its higher order terms are non-negligible and their analysis should be carried out in future research.

6.7 Perturbation under special choice of parameters

We next look at how the matrices and vectors involved in (6.75) change when we consider the special case $M = N$, for which $\delta = \epsilon \frac{M}{N} = \epsilon$, and suppose that $0 < \delta, \mu \ll 1$. We then obtain

$$U = \left(\frac{1}{2\mu} + \frac{1}{4} + \frac{\mu}{8} \right) \mathbb{1}_{4+\delta} \begin{pmatrix} -\frac{1}{2\mu^2} + \frac{1}{2\mu} + \frac{1}{8} & \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} & \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} & 0 \\ \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} & -\frac{1}{2\mu^2} + \frac{1}{2\mu} + \frac{1}{8} & 0 & \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} \\ \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} & 0 & -\frac{1}{2\mu^2} + \frac{1}{2\mu} + \frac{1}{8} & \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} \\ 0 & \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} & \frac{1}{4\mu^2} - \frac{1}{4\mu} - \frac{1}{16} & -\frac{1}{2\mu^2} + \frac{1}{2\mu} + \frac{1}{8} \end{pmatrix} + O(\delta^2, \mu^2). \quad (6.86)$$

$$\Delta_0 = (1 - \mu)^2 \nu (1 - \delta) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 - \delta & 0 & \delta \\ 0 & 0 & 1 - \delta & \delta \\ 0 & \delta & \delta & 2(1 - \delta) \end{pmatrix}, \quad (6.87)$$

$$\Gamma_0 = (1 - \mu)^2 (1 - \delta)^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (6.88)$$

After substitution of (6.86)–(6.88) into (6.75) and expansion of ζ in δ and μ , we obtain

$$\zeta = \frac{1 - \Psi_{0,0}^{(4)}}{N} \nu \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{2\mu^2} \end{pmatrix} + \delta \begin{pmatrix} 0 \\ \frac{3}{8\mu^3} - \frac{15}{16\mu^2} \\ \frac{3}{8\mu^3} - \frac{15}{16\mu^2} \\ -\frac{1}{\mu^3} + \frac{3}{2\mu^2} \end{pmatrix} \right] + o\left(\delta, \frac{1}{\mu^2}\right). \quad (6.89)$$

We can identify two regimes: $\frac{1}{\mu^2} \gg \frac{\delta}{\mu^3}$ and $\frac{1}{\mu^2} \ll \frac{\delta}{\mu^3}$, which correspond to, respectively, $\mu \gg \delta$ and $\mu \ll \delta$. In the first regime, which corresponds to individuals feeling the effect of mutation more strongly than the effect of the seed-bank, we obtain

$$\zeta \sim \frac{1 - \Psi_{0,0}^{(4)}}{N} \nu \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{2\mu^2} \end{pmatrix}, \quad (6.90)$$

where the first three entries are 0 because the exchange rate with the seed-bank is too small.

In the second regime, which corresponds to individuals interacting with the seed-bank many times before a mutation occurs, we obtain

$$\zeta \sim \frac{1 - \Psi_{0,0}^{(4)}}{N} \nu \delta \begin{pmatrix} 0 \\ \frac{3}{8\mu^3} - \frac{15}{16\mu^2} \\ \frac{3}{8\mu^3} - \frac{15}{16\mu^2} \\ -\frac{1}{\mu^3} + \frac{3}{2\mu^2} \end{pmatrix}. \quad (6.91)$$

Notice that the fourth entry of ζ is negative for $\mu < \frac{2}{3}$, which is our case since we require $\mu \ll 1$. In (6.89) we also have that the fourth entry is negative for $\mu < \frac{2\delta}{1+3\delta}$. This seems to contradict the required non-negativity of ζ , since it equals the second moment of $\Psi_{0,x}$. However, looking closely at its general expression, this turns out to be non-negative for all values of the parameters, and equals zero only for $\mu = 1$. When taking Taylor expansions of its expression up to a finite order $n \in \mathbb{N}$ in both δ and μ , however, these turn out to be negative for values of μ in an interval $(a_n(\delta), 1]$, where $\lim_{n \rightarrow \infty} a_n(\delta) = 1$. In the limit as n goes to infinity, however, all entries of ζ are non-negative and the results are consistent.

References

- [1] J. Blath, A. González Casanova, N. Kurt, M. Wilke-Berenguer, *A new coalescent for seed-bank models*, preprint, arXiv:1411.4747, 2014.
- [2] S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and Convergence (2nd ed.)*, Wiley, New York, 2005.
- [3] F. den Hollander, *Stochastic Models for Genetic Evolution*, Lecture Notes for a Master Course, 2012.
- [4] R. Durrett, *Probability Models for DNA Sequence Evolution (2nd ed.)*, Springer, New York, 2008.
- [5] J. Pitman, *Coalescents with multiple collisions*, Ann. Probab. **27**, 1870-1902, 1999.
- [6] J. Schweinsberg, *A necessary and sufficient condition for the Λ -coalescent to come down from infinity*, Electron. Comm. Probab. **5**, 1-11, 2000.
- [7] J. T. Lennon, S. E. Jones, *Microbial seed banks: the ecological and evolutionary implications of dormancy*, Nature Reviews Microbiology **9**, 119-130, 2011.
- [8] I. Kaj, S. Krone, M. Lascoux, *Coalescent theory for seed bank models*, Journals of Applied Probability **38**, 285-300, 2001.
- [9] J. Blath, A. González Casanova, N. Kurt, D. Spanò, *The ancestral process of long-range seedbank models*, Journals of Applied Probability **50**, 741-759, 2013.
- [10] A. González Casanova, E. Aguirre-von Wobeser, G. Espin, L. Servin-González, N. Kurt, D. Spanò, J. Blath, G. Soberón-Chávez, *Strong seed-bank effects in bacterial evolution*, J. Theor. Biol. **356**, 62-70, 2014.
- [11] J. Blath, A. González Casanova, B. Eldon, N. Kurt, *Genealogy of a Wright Fisher model with strong seed-bank component*, Submitted, arXiv:1403.2925, 2014.