Structural properties of the seed bank and the two island diffusion

Jochen Blath\textsuperscript{1} · Eugenio Buzzoni\textsuperscript{1} · Adrián González Casanova\textsuperscript{2,3} · Maite Wilke-Berenguer\textsuperscript{4}

Received: 30 August 2018 / Revised: 5 February 2019 / Published online: 9 May 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract
We investigate various aspects of the (biallelic) Wright–Fisher diffusion with seed bank in conjunction with and contrast to the two-island model analysed e.g. in Kermany et al. (Theor Popul Biol 74(3):226–232, 2008) and Nath and Griffiths (J Math Biol 31(8):841–851, 1993), including moments, stationary distribution and reversibility, for which our main tool is duality. Further, we show that the Wright–Fisher diffusion with seed bank can be reformulated as a one-dimensional stochastic delay differential equation, providing an elegant interpretation of the age structure in the seed bank also forward in time in the spirit of Kaj et al. (J Appl Probab 38(2):285–300, 2001). We also provide a complete boundary classification for this two-dimensional SDE using martingale-based reasoning known as McKean’s argument.

Keywords Wright–Fisher diffusion · Seed bank coalescent · Two island model · Boundary classification · Duality · Reversibility · Stochastic delay differential equation

Mathematics Subject Classification Primary 60K35 · Secondary 92D10

1 Introduction
The Wright–Fisher diffusion is a classical probabilistic object in mathematical population genetics. It describes the scaling limit of the neutral allele frequencies in a large
haploid population on a macroscopic timescale, known as *evolutionary timescale*. This diffusion process and its generalizations have been thoroughly investigated, starting with the pioneering work of Wright (1931). See also Ethier and Kurtz (1992), Fu et al. (2003) and Etheridge (2011) for more recent mathematical accounts, in particular regarding characterizations of the stationary distribution in the presence of weak neutral mutation and the corresponding boundary classification, and for further references. It is well known that the scaling limit of the genealogy of a sample from this model is given by a *Kingman coalescent* (with mutation), cf. Wakeley (2009) for an overview of coalescent theory.

In the presence of population structure, e.g. in the guise of the two-island model (Wright 1931; Moran 1959), many new effects appear. Here, the classical Wright–Fisher model is modified by dividing the population into two separate subpopulations, e.g. living on *two islands*, with migration between them. As a result, for example, the genealogy of a sample taken from the subdivided population will now be described by a *structured coalescent* in which two lines may merge only at times when both are in the same island. Yet, other qualitative features remain unchanged, including the fact that the structured coalescent with two islands still “comes down from infinity”. Interestingly, already in this simple and well-studied model, the description of important quantities becomes much more involved or remained open. For example, for the Wright–Fisher diffusion with two islands (tracing the limiting neutral allele frequencies on both islands), under weak mutation, there seems to be no explicit closed-form characterization of its stationary distribution, though recursion formulas can still be found, see e.g. Nath and Griffiths (1993), Fu et al. (2003) and Kermany et al. (2008) for results in this direction. Further, we are also not aware of a full boundary classification. The standard Feller approach via speed measure and scale function cannot be employed here, since the two island model leads to a two-dimensional diffusion.

The situation changes again when a seed bank is being added to a classical Wright–Fisher model. Such seed banks can be thought of as reservoirs of dormant individuals that can potentially be resuscitated in the future. They are ubiquitous for example in bacterial communities, where they extend the persistence of genotypes and are important for the diversity and functioning of populations. See Lennon and Jones (2011) and Shoemaker and Lennon (2018) for recent overviews. Models for so-called “strong” seed banks, that is, where individuals may stay dormant for long periods of time, are currently under investigation, and in fact only recently, in Blath et al. (2016), the *Wright–Fisher diffusion with strong seed bank* and its dual, the *seed bank coalescent*, have been introduced as mathematical objects (see also Lambert and Ma 2015, in which the same dual has been obtained as scaling limit of the genealogy in a metapopulation model with peripatric speciation). While at first glance similar to the two-island diffusion and the structured coalescent, the seed bank diffusion and its dual exhibit some remarkable qualitative differences. For example, the seed bank coalescent does not come down from infinity, and its expected time to the most recent common ancestor is unbounded as the sample size increases (see Blath et al. 2016 for details).

Given these similarities and differences, it is a natural task to jointly investigate the properties of these models and thereby the relation between them.
The paper is organized as follows: the basic models under consideration are introduced in Sect. 1.1 followed by the characterization of the seed bank diffusion as a \textit{stochastic delay differential equation} in Sect. 1.2. This provides an elegant manifestation in a \textit{forward-in-time} model of the age-structure introduced by seed banks similar to that seen in the \textit{backward-in-time} considerations in the classical modeling of seed banks in Kaj et al. (2001). In Sect. 2, we observe a non-standard dual processes for our models with mutation that allows us to characterize the moments of the unique stationary distribution with the help of recursions and show that the seed bank diffusion is non-reversible. Finally, in Sect. 3, we investigate the boundary behavior of the seed bank and the two island diffusion using a technique called McKean’s argument, which is based on the martingale convergence theorem on stochastic intervals and is suitable also in multi-dimensional settings. A complete analysis of both models is possible, since they are instances of so-called \textit{polynomial diffusions}, which have recently drawn considerable attention in the finance literature (see e.g. Filipović and Larsson 2016). We think that the flexibility of McKean’s argument should potentially make it widely applicable in population genetics, beyond the two models discussed above. We conclude with a short discussion of the biological interpretation of our results.

1.1 The model(s)

The \textit{Wright–Fisher diffusion with seed bank} was recently introduced in Blath et al. (2016) as the forward in time scaling limit of a bi-allelic Wright–Fisher model (with type space \{a, A\}) that describes a population where individuals may stay inactive in a \textit{dormant form} such as seeds or spores (in the \textit{seed bank}), essentially “jumping” a significant (geometrically distributed) number of generations, before rejoining the \textit{active} population. For an active population of size \(N\) and a seed bank size \(M = \lfloor KN \rfloor, K > 0\), under the classical scaling of speeding time by a factor \(N\) the \(a\)-allele frequency process \((X_N(t))_{t \geq 0}\) in the active and \((Y_N(t))_{t \geq 0}\) in the dormant population converge to the (unique strong) solution \((X_t, Y_t)_{t \geq 0}\) of a two-dimensional SDE. In Blath et al. (2015) the model was extended to include mutation in both the active and the dormant population in which case the limiting process is the solution to the SDE given in Definition 1.1 below. Since the population model and limiting result are completely analogous to the case without mutation we refrain from details and instead refer to Blath et al. (2016, Section 2).

Definition 1.1 (Seed bank diffusion) Let \((B(t))_{t \geq 0}\) be a standard Brownian motion, \(u_1, u_2, u'_1, u'_2\) be finite, non-negative constants and \(c, K\) finite, positive constants. The \textit{Wright–Fisher diffusion with seed bank} with parameters \(u_1, u_2, u'_1, u'_2, c, K\), starting in \((x, y) \in [0, 1]^2\), is given by the \([0, 1]^2\)-valued continuous strong Markov process \((X(t), Y(t))_{t \geq 0}\) that is the unique strong solution of the initial value problem

\[
dX(t) = \left[-u_1X(t) + u_2(1 - X(t)) + c(Y(t) - X(t))\right]dt + \sqrt{X(t)(1 - X(t))}dB_t,
\]

\[
dY(t) = \left[-u'_1Y(t) + u'_2(1 - Y(t)) + Kc(X(t) - Y(t))\right]dt,
\]

with \((X(0), Y(0)) = (x, y) \in [0, 1]^2\).
The first coordinate process \((X(t))_{t \geq 0}\) can be interpreted as describing the fraction of \(a\)-alleles in the limiting active population, while \((Y(t))_{t \geq 0}\) gives the fraction of \(a\)-alleles in the limiting dormant population, i.e. in the seed bank. The parameters \(u_1, u_2\) describe the mutation rates from \(a\) to \(A\), respectively from \(A\) to \(a\), in the active population, and \(u_1', u_2'\) the corresponding values in the seed bank. Note that the mutation rates may differ for active and dormant individuals. \(K\) fixes the so-called relative seed bank size \((M = [KN])\) and \(c\) is the rate of migration between the active population and seed bank, i.e. initiation of dormancy and resuscitation. For more details on the biological background see Blath et al. (2015, 2016).

**Remark 1.2** (General model and two-island diffusion) A natural extension of this model can be obtained by (potentially) adding noise in the second coordinate. For parameters \(u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0, c, c' > 0\) and independent standard Brownian motions \((B(t))_{t \geq 0}, (B'(t))_{t \geq 0}\) consider the initial value problem

\[
\begin{align*}
dX(t) &= \left[-u_1 X(t) + u_2 (1 - X(t)) + c (Y(t) - X(t))\right] dt + \alpha \sqrt{X(t) (1 - X(t))} dB(t), \\
y(t) &= \left[-u_1' Y(t) + u_2' (1 - Y(t)) + c' (X(t) - Y(t))\right] dt + \alpha' \sqrt{Y(t) (1 - Y(t))} dB'(t),
\end{align*}
\]

with \((X(0), Y(0)) = (x, y) \in [0, 1]^2\).

For \(\alpha = 1, \alpha' = 0\) and \(c' = c K\) (for a \(K > 0\)) this is the seed bank diffusion. For \(\alpha' > 0\) we obtain the diffusion of Wright’s two-island model initially introduced in Wright (1931) and considered in this form for example in Kermany et al. (2008).

As is standard, Theorem 3.2 in Shiga and Shimizu (1980) provides the existence of a unique strong solution for every (possibly random) initial condition \((X(0), Y(0)) = (x, y) \in [0, 1]^2\), both initial value problems (1) and (2) admit a unique strong semimartingale solution which is a two-dimensional continuous strong Markov diffusion. Denote by \(A^{u_1, u_2, u_1', u_2', \alpha, \alpha', c, c'}\) its Markov generator and note that its domain contains \(C^2([0, 1]^2)\), the space of twice continuously differentiable functions inside \([0, 1]^2\) (continuous at the boundary). Since the model referred to will be clear from the context, we will usually omit the superscripts on the generator and simply write \(A\). The action of \(A\) on any test function \(f \in C^2([0, 1]^2)\) is then described by

\[
Af(x, y) = \left[-u_1 x + u_2 (1 - x) + c (y - x)\right] \frac{\partial f}{\partial x}(x, y) + \frac{\alpha^2}{2} x (1 - x) \frac{\partial^2 f}{\partial x^2}(x, y) \\
+ \left[-u_1' y + u_2' (1 - y) + c' (x - y)\right] \frac{\partial f}{\partial y}(x, y) + \frac{(\alpha')^2}{2} y (1 - y) \frac{\partial^2 f}{\partial y^2}(x, y)
\]

Note that there is no ambiguity in the definition of the process at the boundaries, since the diffusion part vanishes at 0 and 1, so that no further conditions on the domain of the generator are required.

**Remark 1.3** (Extension to multiple seed banks) It is straightforward to extend system (1) to several (e.g. geographically) subdivided seed banks. Indeed, let \(k \geq 1\) and denote
the frequency process for the active population by \((X(t))_{t \geq 0}\). Assume there are \(k\) seed banks. For each seed bank \(i \in \{1, \ldots, k\}\) we consider specific parameters \(c_i, K_i\) as well as mutation rates \(u_1^i, u_2^i\) and denote its frequency process by \((Y_i(t))_{t \geq 0}\). Then the seed bank diffusion with \(k\) seed banks is given by the following \(k + 1\) interacting SDEs

\[
\begin{align*}
    dX(t) &= \left[ -u_1 X(t) + u_2 (1 - X(t)) + \sum_{i=1}^{k} c_i (Y_i(t) - X(t)) \right] dt + \sqrt{X(t)(1 - X(t))} dB(t), \\
    dY_i(t) &= \left[ - u_1^i Y_i(t) + u_2^i (1 - Y_i(t)) + K_i c_i (X(t) - Y_i(t)) \right] dt, \quad i \in \{1, \ldots, k\},
\end{align*}
\]

with initial value \((X(0), Y_1(0), \ldots, Y_k(0)) = (x, y_1, \ldots, y_k) \in [0, 1]^{k+1}\). Existence and uniqueness are again standard with Theorem 3.2 in Shiga and Shimizu (1980). Of course, this can be further generalized to multiple islands with multiple seed banks, see den Hollander and Pederzani (2017).

### 1.2 A stochastic delay differential equation

Note that the only source of randomness in the two-dimensional system (1) and also in its generalization (4) is the one-dimensional Brownian motion \((B(t))_{t \geq 0}\) driving the fluctuations in the active population. This and the special form of the seed bank(s) allow us to reformulate this system as an essentially one-dimensional stochastic delay differential equation. Recall the notation from Remark 1.3 and abbreviate, for convenience, \(u^i := u_1^i + u_2^i\), \(i \in \{1, \ldots, k\}\).

**Proposition 1.4** The solution to (4) with initial values \(x, y_1, \ldots, y_k \in [0, 1]\) is a.s. equal to the unique strong solution of the system of stochastic delay differential equations

\[
\begin{align*}
    dX(t) &= \sum_{i=1}^{k} c_i \left( y_i e^{-(u^i + K_i c_i) t} + \int_{0}^{t} e^{-(u^i + K_i c_i) (t-s)} (u_2^i + K_i c_i X(s)) ds - X(t) \right) dt \\
    &+ \left[ - u_1 X(t) + u_2 (1 - X(t)) \right] dt + \sqrt{X(t)(1 - X(t))} dB(t), \\
    dY_i(t) &= \left( - y_i (u^i + K_i c_i) e^{-(u^i + K_i c_i) t} \right. \\
    &- \left( u^i + K_i c_i \right) \int_{0}^{t} e^{-(u^i + K_i c_i) (t-s)} (u_2^i + K_i c_i X(s)) ds + u_2^i + K_i c_i X(t) \right) dt,
\end{align*}
\]

for \(i \in \{1, \ldots, k\}\) with the same initial condition.

Note that the first equation in (5) does not depend on \(Y_i, i = 1, \ldots, k\), and that the latter equations for the \(Y_i\) are in turn deterministic functions of \(X\), so that the system of SDDDEs is essentially one-dimensional.

**Remark 1.5** Let us consider a simple special case of the above result to reveal the underlying “age structure”: it is an immediate corollary from the above that the seed
bank diffusion solving (1) with parameters $c = 1$, $K = 1$, $u_1 = u_2 = u'_1 = u'_2 = 0$, started in $X(0) = \mathbf{x} = \mathbf{y} = Y(0) \in [0, 1]$ is a.s. equal to the unique strong solution of the stochastic delay differential equations

\begin{align*}
\mathrm{d}X(t) &= \left( xe^{-t} + \int_0^t e^{-(t-s)} X(s) \mathrm{d}s - X(t) \right) \mathrm{d}t + \sqrt{X(t)(1-X(t))} \mathrm{d}B(t), \\
\mathrm{d}Y(t) &= \left( - ye^{-t} - \int_0^t e^{-(t-s)} X(s) \mathrm{d}s + X(t) \right) \mathrm{d}t,
\end{align*}

(6)

with the same initial condition. This now provides an elegant interpretation of the delay in the SDDE as the seed bank. Indeed, it shows that the type (a or A) of any “infinitesimal” resuscitated individual, is determined by the active population present an exponentially distributed time ago (with a cutoff at time 0), which the individual spent dormant in the seed bank. The net effect is positive if the frequency of a-alleles at that time was higher than the current frequency, and negative if it was lower. This is the forward-in-time equivalent of the model for seed banks or dormancy in the coalescent context as formulated in Kaj et al. (2001), where the seed bank is modelled by having individuals first choose a generation in the past according to some measure $\mu$ and then choosing their ancestor uniformly among the individuals present in that generation. The seed bank model given in Blath et al. (2016) is obtained when $\mu$ is chosen to be geometric, i.e. memoryless, like the exponential distribution. This indicates that a forward-in-time model for more general dormancy models are to be searched among SDDEs rather than among SDEs.

Such a reformulation is of course not feasible for the two island model, which is driven by two independent sources of noise.

**Proof of Proposition 1.4** Recall e.g. from Revuz and Yor (1999, Proposition 3.1) that for continuous semimartingales $Z$, $W$ we have the integration by parts formula

\[ \int_0^t W(s) \mathrm{d}Z(s) = W(t) Z(t) - Z(0) W(0) - \int_0^t Z(s) \mathrm{d}W(s) - \langle Z, W \rangle(t), \]

where $\langle \cdot, \cdot \rangle$ denotes the covariance process and $t \geq 0$. Note that for every differentiable deterministic function $f$, since $\langle Z, f \rangle \equiv 0$, this reduces to

\[ f(t) Z(t) - f(0) Z(0) = \int_0^t f(s) \mathrm{d}Z(s) + \int_0^t f'(s) Z(s) \mathrm{d}s. \]

Substituting the expression for $\mathrm{d}Y_i(t)$ from (4), we obtain that

\[ \begin{align*}
f(t)Y_i(t) - f(0)Y_i(0) &= \int_0^t f(s) \left[ - u'_1 Y_i(s) + u'_2 (1 - Y_i(s)) + K_i c_i (X(s) - Y_i(s)) \right] \mathrm{d}s \\
&\quad + \int_0^t f'(s) Y_i(s) \mathrm{d}s.
\end{align*} \]  

(7)
Letting \( f(t) := e^{(u^i + K_i c_i)t}, t \geq 0 \), Eq. (7) further simplifies to

\[
 f(t)Y_i(t) - f(0)Y_i(0) = \int_0^t e^{(u^i + K_i c_i)s} (K_i c_i X(s) + u_2^i) \, ds.
\]

This can be rewritten, given the initial value \( y_i = Y_i(0) \), as

\[
e^{(u^i + K_i c_i)t} Y_i(t) = y_i + \int_0^t e^{(u^i + K_i c_i)s} (K_i c_i X(s) + u_2^i) \, ds.
\]

By dividing on both sides by \( e^{(u^i + K_i c_i)t} \) we finally get

\[
 Y_i(t) = y_i e^{-(u^i + K_i c_i)t} + \int_0^t e^{(u^i + K_i c_i)(s-t)} (K_i c_i X(s) + u_2^i) \, ds.
\]

Plugging this into the first line of the system in (4) proves that the unique strong solution of (4) is a strong solution to (5). On the converse, let \((X(t), Y_1(t), \ldots, Y_k(t))_{t \geq 0}\) now be a solution to (5). (We already know that there exists at least one.) Using (8) we immediately see that \((X(t))_{t \geq 0}\) solves the first equation in (4). Likewise, using (8) in the right-hand-side of the last \(k\) equations in (5), we obtain the last \(k\) equations of (4). Since (4) has a unique solution, this must then hold for (5), too, and the two solutions coincide \(\mathbb{P}\)-almost surely.

\[\square\]

2 Duality, moments and stationary distribution

A convenient way to study the behavior of diffusions in population genetics has proven to be the usage of (moment-) duality for Markov processes, cf. Jansen and Kurt (2014) for an overview of the technique and Blath et al. (2016), Etheridge (2011) for examples of applications. The art lies in finding a suitable dual that can serve as such a tool. In Blath et al. (2016) the moment dual of the seed bank diffusion (without mutation) \((N(t), M(t))_{t \geq 0}\) (known therein as block-counting process of the seed bank coalescent) is given as the continuous time Markov chain with values in \(E := N_0 \times N_0\) equipped with the discrete topology, with conservative generator \(\tilde{A}\), i.e. \(\tilde{A}\) is the matrix of transition rates given by:

\[
\tilde{A}_{(n,m), (\tilde{n}, \tilde{m})} = \begin{cases} 
\binom{n}{2} & \text{if } (\tilde{n}, \tilde{m}) = (n-1, m), \\
 cn & \text{if } (\tilde{n}, \tilde{m}) = (n-1, m+1), \\
 cKm & \text{if } (\tilde{n}, \tilde{m}) = (n+1, m-1), 
\end{cases}
\]

when \((\tilde{n}, \tilde{m}), (n, m) \in N_0 \times N_0\), zero otherwise off the diagonal and with the elements on the diagonal chosen such that the rows add up to 0.

A difficulty arises when adding mutation to the model. There are several ways of incorporating this mechanism into a dual and we comment on this as well as the motivation behind our strategy—adding a death state to the state-space—below, but will first formally introduce our dual.
Fig. 1 The coalescent corresponding to the process defined in Definition 2.1. The dashed and black lines represent the two islands. When a forward-mutation of type $A \mapsto a$ occurs, the line is ended, since it ensures all its leaves to be of type $a$. A forward-mutation of type $a \mapsto A$ renders it impossible to have all leaves of type $a$ and the process jumps to the death state $(\partial, \partial)$.

Definition 2.1 (Moment dual of the general diffusion) Consider the space $E := \mathbb{N}_0 \times \mathbb{N}_0 \cup \{ (\partial, \partial) \}$ equipped with the discrete topology. Let $u_1, u_2, u_1', u_2', a, a' \geq 0, c, c' > 0$. Define $(N(t), M(t))_{t \geq 0}$ to be the continuous time Markov chain with values in $E$ with conservative generator $\bar{A}$ given by:

$$
\bar{A}_{(n,m), (\tilde{n}, \tilde{m})} = \begin{cases} 
\alpha^2 \binom{n}{2} + nu_2 & \text{if } (\tilde{n}, \tilde{m}) = (n - 1, m), \\
(a')^2 \binom{m}{2} + mu_2' & \text{if } (\tilde{n}, \tilde{m}) = (n, m - 1), \\
nu_1 + mu_1' & \text{if } (\tilde{n}, \tilde{m}) = (\partial, \partial), \\
cn & \text{if } (\tilde{n}, \tilde{m}) = (n - 1, m + 1), \\
c'm & \text{if } (\tilde{n}, \tilde{m}) = (n + 1, m - 1), 
\end{cases}
$$

when $(\tilde{n}, \tilde{m}), (n, m) \in \mathbb{N}_0 \times \mathbb{N}_0$ and zero otherwise off the diagonal. We call this process the moment dual of the diffusion associated with the diffusion given in (2).

The name of the process will be justified in Lemma 2.2 below. This dual, illustrated in Fig. 1, arises in the context of a sampling duality. See González Casanova and Spanò (2018) for a thorough introduction to the concept. It is based on the idea that the question of “what is the probability of sampling $n$ individuals of type $a$ at time $t$, if the frequency of type $a$ is $x$ at time 0?” can be answered in two ways: one, looking forward in time at the diffusion which will give the frequency of type $a$ individuals at time $t$ precisely, but also two, tracing back the genealogy to the number of ancestors of the sample present at time 0 and using the frequency $x$. It is precisely in this latter view that one realizes the need of an artificial death-state. In order for all $n$ individuals in the sample to be of type $a$ at time $t$ it is imperative that we do not encounter a mutation from type $a$ to $A$ in the forward sense, i.e. a mutation from $A$ to $a$ in the coalescent-time, on their ancestral lines. Hence, the process $(N(t), M(t))_{t \geq 0}$ is killed off as soon as this happens, since the probability for the sample to be of type $a$ only is now 0. At the same time, if we encounter a mutation of type $A$ to $a$ in the forward sense, i.e. a mutation from type $a$ to $A$ tracing backwards, we are assured all descendants of that
Fig. 2 The second moments $M_{2,0}$ (left) and $M_{0,2}$ (right) for a varying mutation rate $u_1 = u_2$. The parameters $c = K = 1$ are fixed. Red: seed bank ($\alpha' = 0$) with no mutation in the seeds ($u_1' = u_2' = 0$). Magenta: seed bank ($\alpha' = 0$) with mutation in the seed bank ($u_1' = u_2' = u_1$). Dark blue: two-island ($\alpha' = 1$) with no mutation in the second island ($u_1' = u_2' = 0$). Light blue: two-island ($\alpha' = 1$) with mutation in the second island ($u_1' = u_2' = u_1$) (colour figure online).

line are of type a with probability 1 and we can stop tracing it, whence the process is reduced by one line. See Fig. 2 for an illustration.

It is trivial to extend the dual process in Definition 2.1 to a general structured coalescent. A structured-mutation moment dual is new in the literature, as far as we know. These moment duals differ from the weighted moment dual for the Wright–Fisher diffusion with mutation introduced in Etheridge and Griffiths (2009) and studied extensively in Griffiths et al. (2016) and Etheridge et al. (2010). The small difference between our construction for mutation and the construction in Etheridge and Griffiths (2009), namely the addition of the extra state $\partial$, makes our dual compatible with the presence of selection as in Krone and Neuhauser (1997).

The following are straightforward, but important observations on the duals: note that in the case of $u_1 + u_2 + u_1' + u_2' > 0$, the moment dual of the general diffusion will reach either $\{(0, 0)\}$ or $\{((\partial, \partial))\}$ in finite time (for any starting point $(n, m) \in E$), whereas for $u_1 + u_2 + u_1' + u_2' = 0$ it will reach the set $\{(1, 0), (0, 1)\}$ in finite time ($\mathbb{P}$-a.s.) and then alternate between these two states. Furthermore observe that, whenever the dual of the general diffusion is started in some $(n, m) \in E$, it will stay in $\{0, \ldots, n + m\} \times \{0, \ldots, n + m\} \cup \{((\partial, \partial))\}$, hence the state space in this case is, indeed, finite.

**Lemma 2.2** Let $S: [0, 1]^2 \times E \to [0, 1]$ be defined as

$$S((x, y), (n, m)) := x^n y^m \mathbb{1}_{N_0 \times M_0}((n, m))$$

for any $(x, y) \in [0, 1]^2$ and $(n, m) \in E$ and let $u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0$, $c, c' > 0$. Then for every $(x, y) \in [0, 1]^2$, $(n, m) \in E$ and for any $t \geq 0$

$$\mathbb{E}^{x, y}[S((X(t), Y(t)), (n, m))] = \mathbb{E}_{n,m}[S((x, y), (N(t), M(t)))],$$

where $(N(t), M(t))_{t \geq 0}$ is defined in Definition 2.1 and $(X(t), Y(t))_{t \geq 0}$ is the solution to the SDE in Eq. (2).
Since $S: [0, 1]^2 \times E \rightarrow [0, 1]$ is continuous (in the product topology of $[0, 1]^2 \times E$), the result follows by proving the assumptions of Proposition 1.2 in Jansen and Kurt (2014), which consist of certain requirements on the respective generators $A$ of $(X(t), Y(t))_{t \geq 0}$ and $\tilde{A}$ of $(N(t), M(t))_{t \geq 0}$.

Recall the generator $A$ of $(X(t), Y(t))_{t \geq 0}$ from (3) and observe that for any bounded function $h: E \rightarrow \mathbb{R}$, the generator of $(N(t), M(t))_{t \geq 0}$ is given by $Ah((\partial, \partial)) = 0$ and

$$\tilde{A}h(n, m) = \left[\frac{\alpha^2}{2} + nu_2\right]h(n, m) - \left[\frac{(\alpha')^2}{2} + mu'_2\right]h(n, m)(n, m) + [n \alpha_1 + mu'_1]h(\partial, \partial)$$

for any $(n, m) \in \mathbb{N}_0 \times \mathbb{N}_0$ with the convention that $\left(\frac{1}{2}\right) = 0$. Let $P$ and $\tilde{P}$ be the semigroups corresponding to $A$ and $\tilde{A}$ respectively. Since $(N(t) + M(t))_{t \geq 0}$ is monotonically non-increasing, the assumptions that $S((x, y), (n, m))$, $P_tS((x, y), (n, m))$ are in the domain of $A$ and $S((x, y), (n, m))$, $\tilde{P}_tS((x, y), (n, m))$ are in the domain of $A$ are readily verified.

As $S((x, y), (\partial, \partial)) = 0$ for all $(x, y) \in [0, 1]^2$, we immediately see that

$$AS((x, y), (\partial, \partial)) = 0 = \tilde{A}S((x, y), (\partial, \partial))$$

for any $(x, y) \in [0, 1]^2$. Furthermore, if we fix $(x, y) \in [0, 1]^2$ and $(n, m) \in \mathbb{N}_0 \times \mathbb{N}_0$,

$$AS((x, y), (n, m)) = [-u_1x + u_2(1 - x) + c(y - x)]n x^{n-1} y^{m}$$

$$+ \frac{\alpha^2}{2} x(1 - x)n(n - 1)x^{n-2} y^{m}$$

$$+ [-u_1' y + u_2'(1 - y) + c'(x - y)]m x^{n} y^{m-1}$$

$$+ \frac{(\alpha')^2}{2} y(1 - y)m(m - 1)x^{n} y^{m-2}$$

$$= \left[\frac{\alpha^2(n)}{2} + nu_2\right]x^{n-1} y^{m} - x^{n} y^{m}$$

$$+ [(\alpha')^2(m) + mu'_2]x^{n} y^{m-1} - x^{n} y^{m}$$

$$+ c[x^{n-1} y^{m+1} - x^{n} y^{m}] + c'[x^{n+1} y^{m-1} - x^{n} y^{m}]$$

$$+ (nu_1 + mu'_1)[0 - x^{n} y^{m}]$$

$$= \tilde{A}S((x, y), (n, m)).$$

\[\square\]
This duality now allows us to use the process \((N(t), M(t))_{t \geq 0}\) to study the mixed moments of \((X(t), Y(t))_{t \geq 0}\) from which we can draw conclusions on the limiting behavior of the diffusions itself. The case with and without mutation differs strongly in this behavior.

**Remark 2.3** In the absence of mutation in the general diffusion given in (2) with \(\alpha = \alpha' = 1\)

\[
\lim_{t \to \infty} \mathbb{E}^{x,y}[X(t)^n Y(t)^m] = \frac{yc + xc'}{c + c'}
\]

for all \((n, m) \in \mathbb{N}_0 \times \mathbb{N}_0\) \(\setminus\{(0, 0)\}\) and all \((x, y) \in [0, 1]^2\). From this we can conclude that \((X(t), Y(t))_{t \geq 0}\) converges \(\mathbb{P}\)-a.s. to a random variable \((X_\infty, Y_\infty)\) with values in \([0, 1]^2\) whose distribution is given by

\[
\frac{yc + xc'}{c + c'} \delta_{(1,1)} + \frac{(1 - y)c + (1 - x)c'}{c + c'} \delta_{(0,0)}
\]

as can easily be seen by the same arguments as in Proposition 2.9 and Corollary 2.10 in Blath et al. (2016).

**Proposition 2.4** Let \(u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0\), \(c, c' > 0\) and assume that at least one mutation rate among \(u_1, u_2, u_1', u_2'\) is non-zero. Then, for every \((n, m) \in \mathbb{N}_0 \times \mathbb{N}_0\) and for every \((x, y) \in [0, 1]^2\)

\[
\lim_{t \to \infty} \mathbb{E}^{x,y}[X(t)^n Y(t)^m] = \mathbb{P}_{n,m} \left\{ \lim_{t \to \infty} (N(t), M(t)) = (0, 0) \right\}.
\]

**Proof** Fix \((x, y) \in [0, 1]^2\) and \((n, m) \in \mathbb{N}_0 \times \mathbb{N}_0\). Then

\[
\lim_{t \to \infty} \mathbb{E}^{x,y}[X(t)^n Y(t)^m] = \lim_{t \to \infty} \mathbb{E}^{x,y}[X(t)^n Y(t)^m \mathbb{1}_{N_0 \times N_0}(n, m)] = S((X(t), Y(t)), (n, m))
\]

\[
= \lim_{t \to \infty} \mathbb{E}_{n,m} \left[ x^{N(t)} y^{M(t)} \mathbb{1}_{N_0 \times N_0}(N(t), M(t)) \right]
\]

\[
= \mathbb{P}_{n,m} \left\{ \lim_{t \to \infty} (N(t), M(t)) = (0, 0) \right\},
\]

where the last three equalities follow from the duality in Lemma 2.2, bounded convergence and the fact that \((N(t), M(t))_{t \geq 0}\) is absorbed in \((0, 0)\) or \((\partial, \partial)\) in finite time \(\mathbb{P}\)-a.s., respectively. (We use the convention that \(0^0 = 1\).)

We can now use this to characterize the long-term behavior of the diffusion \((X(t), Y(t))_{t \geq 0}\) solving (2) with mutation. In order to do this, note that Proposition 2.4 implies that the following is well-defined.
Definition 2.5 Let $u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0, c, c' > 0$ and assume $u_1 + u_2 + u_1' + u_2' > 0$ in (2). For any $(n, m) \in \mathbb{N}_0 \times \mathbb{N}_0$ define

$$M_{n,m} := \lim_{t \to \infty} \mathbb{P}^{x,y}[X^n(t)Y^m(t)] \quad \text{(for any } (x, y) \in [0,1]^2).$$

Proposition 2.6 Let $u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0, c, c' > 0$. Assume $u_1 + u_2 + u_1' + u_2' > 0$ in (2). Then there exists a unique invariant distribution $\mu$ for $(X(t), Y(t))_{t \geq 0}$ and the diffusion is ergodic in the sense that

$$\mathbb{P}^{x,y} \{(X(t), Y(t)) \in \cdot \} \xrightarrow{w} \mu, \quad \text{for } t \to \infty,$$

for all starting points $(x, y) \in [0,1]^2$, where $\xrightarrow{w}$ denotes weak convergence of measures. Furthermore $\mu$ is characterized by

$$\forall n, m \in \mathbb{N}_0: \int_{[0,1]^2} x^n y^m d\mu(x, y) = M_{n,m}. \quad (9)$$

Proof A simple calculation shows $(M_{n,m})_{n,m \in \mathbb{N}_0}$ to be a so-called completely monotonic sequence. The multidimensional Hausdorff-Theorem (Thm. 4 in Hildebrandt and Schoenberg 1933 and preceding corollary) then yields the existence of a unique measure $\mu$ such that (9) holds. The latter in particular implies

$$1 = M_{0,0} = \int_{[0,1]^2} x^0 y^0 d\mu(x, y) = \mu([0,1]^2)$$

and therefore $\mu$ is a distribution. From the definition of the $M_{n,m}$, $n, m \in \mathbb{N}_0$, we know that

$$\lim_{t \to \infty} \int_{[0,1]^2} p(x, y) d\mathbb{P}^{\tilde{x},\tilde{y}} \{(X(t), Y(t)) \in \cdot \} = \lim_{t \to \infty} \mathbb{P}^{\tilde{x},\tilde{y}}[p(X(t), Y(t))] = \int_{[0,1]^2} p(x, y) d\mu(x, y)$$

for any polynomial $p$ on $[0,1]^2$ (and any $(\tilde{x}, \tilde{y}) \in [0,1]^2$). Since the polynomials are dense in the set of continuous (and bounded) functions on $[0,1]^2$ we can conclude that

$$\mathbb{P}^{x,y} \{(X(t), Y(t)) \in \cdot \} \xrightarrow{w} \mu.$$

It is now easy to check that $\mu$ is the unique invariant distribution of $(X(t), Y(t))_{t \geq 0}.$ \hfill \Box

Unfortunately, we cannot calculate $\mu$ explicitly, but we can give the following characterization of its mixed moments:
Lemma 2.7  Let \( u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0, c, c' > 0 \) and assume \( u_1 + u_2 + u_1' + u_2' > 0 \). Then \( M_{0,0} = 1 \) and the following recursion holds for all \( (n, m) \in \mathbb{N}_0 \times \mathbb{N}_0 \setminus \{(0, 0)\} \)

\[
M_{n,m} = \frac{1}{D_{n,m}} \left( a_n M_{n-1,m} + a'_m M_{n,m-1} + cn M_{n-1,m+1} + c' m M_{n+1,m-1} \right),
\]

where

\[
a_n := \alpha^2 \left( \frac{n}{2} \right) + nu_2, \quad a'_m := (\alpha')^2 \left( \frac{m}{2} \right) + mu_2,
\]

\[
D_{n,m} := \alpha^2 \left( \frac{n}{2} \right) + (\alpha')^2 \left( \frac{m}{2} \right) + (u_2 + u_1)n + (u_1' + u_2')m + cn + c'm.
\]

We use the notational convention that \((\frac{1}{2}) = 0\) and \(M_{-1,k} = M_{k,-1} = 0\) for any \( k \in \mathbb{N} \).

**Proof**  For the process \((N(t), M(t))_{t \geq 0}\) let

\[
\tau := \inf \{ t \geq 0 : (N(t), M(t)) \neq (n, m) \}.
\]

For any \((n, m) \neq (0, 0), \tau\) is a \(\mathbb{P}_{n,m}\)-a.s. finite stopping time. Using Proposition 2.4 in the first and third, and the strong Markov property in the second equality we see

\[
M_{n,m} = \mathbb{P}_{n,m} \left\{ \lim_{t \to \infty} (N(t), M(t)) = (0, 0) \right\}
= \sum_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0} \mathbb{P}_{i,j} \left\{ \lim_{t \to \infty} (N(t), M(t)) = (0, 0) \right\} \mathbb{P}_{n,m} \{(N(\tau), M(\tau)) = (i, j)\}
= \sum_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0} M_{i,j} \mathbb{P}_{n,m} \{(N(\tau), M(\tau)) = (i, j)\}.
\]

Writing out the values of \(\mathbb{P}_{n,m} \{(N(\tau), M(\tau)) = (i, j)\}\) finishes the proof. \(\square\)

**Remark 2.8**  Given the existence of an invariant distribution, the question of reversibility arises naturally. The classical Wright–Fisher frequency process with mutation is reversible. However, the diffusion process of the two-island model is not, as shown in Kermany et al. (2008). It turns out, that the seed bank diffusion with mutation is not reversible in general, either. Assume for example that \(c, u_1, u_2 \neq 0\) and \(u_1' = u_1, u_2' = u_2\) and recall that for the diffusion to be reversible we would need

\[
\mathbb{E}^\mu \left[ f(X(t), Y(t)) Ag(X(t), Y(t)) \right] = \mathbb{E}^\mu \left[ g(X(t), Y(t)) Af(X(t), Y(t)) \right],
\]

for all \(f, g \in C^2([0, 1]^2)\). This, however, fails for \(f(x, y) = x\) and \(g(x, y) = y\) as can be checked calculating recursively the values for the mixed moments.
3 Boundary classification

We begin with the simple observation that in the presence of mutation the marginals of the stationary distribution $\mu$ of the general diffusion (2) have no atoms at the boundaries. To be precise, if we let $(X, Y) = (X(t), Y(t))_{t \geq 0}$ be the solution to (2), assume $u_1, u_2, u_1', u_2' > 0$ and recall that $\mu$ denotes the unique invariant distribution of $(X(t), Y(t))_{t \geq 0}$, then, for any $t > 0$, we have

$$P^\mu \{X(t) \in \{0, 1\}\} = P^\mu \{Y(t) \in \{0, 1\}\} = 0. \tag{10}$$

This is a straightforward extension of the corresponding observation for the two-island model in Kermany (2008, Proposition 1) with an entirely analogous proof, which is therefore omitted here.

In the above, each of the parameters $u_1, u_2, u_1', u_2'$ is responsible for the value of exactly one of the probabilities in (10). This correspondence will be further clarified in the following more detailed description of the boundary behavior of the solution to (2). Define the first hitting time of $X$ of the boundary 0 by

$$\tau^X_0 := \inf \{t \geq 0 \mid X(t) = 0\},$$

and define $\tau^X_1$, $\tau^Y_0$ and $\tau^Y_1$ in the same manner. We say that started from the interior $X$ will never hit 0, if for every initial distribution $\mu_0$ such that $\mu_0((0, 1) \times \{0\}) = 1$

$$P^{\mu_0} \left( \tau^X_0 < \infty \right) = 0.$$  

Using analogous formulations for the other cases, we state the main result of this section:

**Theorem 3.1** Let $(X(t), Y(t))_{t \geq 0}$ be the solution to (2) with parameters satisfying $u_1, u_2, u_1', u_2', \alpha, \alpha' \geq 0$ and $c, c' > 0$.

(i) Started from the interior $X$ will never hit 0 if and only if $2u_2 \geq \alpha^2$.

(ii) Started from the interior $X$ will never hit 1 if and only if $2u_1 \geq \alpha^2$.

(iii) Started from the interior $Y$ will never hit 0 if and only if $2u_2' \geq (\alpha')^2$.

(iv) Started from the interior $Y$ will never hit 1 if and only if $2u_1' \geq (\alpha')^2$.

**Remark 3.2** Note that the theorem simply states that for, say, $2u_2 < \alpha^2$, there must exist some initial distribution $\tilde{\mu}_0$ such that $P^{\tilde{\mu}_0} \{\tau^X_0 < \infty\} > 0$. However, we will in fact prove the following, more informative statement [and their respective analogs for (ii)–(iv)]:

Let $2u_2 < \alpha^2$. Then, for any $s > 0$ there exists an $\varepsilon > 0$ such that

$$P \{\|(X(0), Y(0)) - (0, 0)\| < \varepsilon\} = 1 \quad \Rightarrow \quad P \left\{ \tau^X_0 \leq s \right\} > 0.$$  

To obtain this result, we regard our SDE as a *polynomial diffusion*. These are solutions to (multidimensional) SDEs whose generator maps the set of polynomials.
of degree $n$ into itself (for any $n \in \mathbb{N}_0$), see for example Filipović and Larsson (2016) and Larsson and Pulido (2017). As explained in the remark after Definition 2.1 in Filipović and Larsson (2016), these are SDES of the form

$$dZ(t) = b(Z(t))dt + \sigma(Z(t))dW(t)$$

(11)

where $W$ is a (multidimensional) Brownian motion, $b$ consists of polynomials of degree at most 1 and $a(x, y) := \sigma(x, y)\sigma(x, y)^T$ of polynomials of degree at most 2.

A quick glance immediately allows the observation that our generalized SDE (2) can be rewritten in the form of (11) with

$$b(x, y) := \begin{pmatrix} -u_1 x + u_2 (1 - x) + c(y - x) \\ -u_1' y + u_2'(1 - y) + c'(x - y) \end{pmatrix}$$

and

$$\sigma(x, y) := \begin{pmatrix} \alpha \sqrt{x(1 - x)} \\ 0 \\ \alpha' \sqrt{y(1 - y)} \end{pmatrix}$$

on $[0, 1]^2$ and a two-dimensional Brownian motion $W$. Since $b$ consists of polynomials of degree 1 and

$$a(x, y) := \sigma(x, y)\sigma(x, y)^T = \begin{pmatrix} \alpha^2 x(1 - x) \\ 0 \\ (\alpha')^2 y(1 - y) \end{pmatrix}$$

of polynomials of degree 2, our SDE is indeed a polynomial diffusion on $[0, 1]^2$ in the sense (and notation) of Filipović and Larsson (2016) and we are free to use the results found therein as well as techniques from this area.

Indeed the ‘if’ direction is proven using one such technique known as “McKean’s argument”—a martingale method that can be applied in any dimension and should therefore have the potential to be applicable for a large class of processes, see Mayerhofer et al. (2011, Section 4) for an overview and further references.

The ‘only if’ direction on the other hand is essentially a direct application of Theorem 5.7 (iii) in Filipović and Larsson (2016). The following proposition therefore mainly paraphrases the abovementioned result in our notation for the reader’s convenience and the proof consists of assuring we consider a suitable set-up.

Define $P = \{x, 1 - x, y, 1 - y\}$ where we abuse notation using $x$ for the map $(x, y) \mapsto x$ and similar for the other polynomials.

**Proposition 3.3** (Theorem 5.7 (iii) in Filipović and Larsson 2016) Let $(X(t), Y(t))_{t \geq 0}$ be the $[0, 1]^2$-valued solution to (2). Recall that it can be written in the form of (11) and that $A$ denotes the corresponding generator given in (3).

For every polynomial $p \in P$, assume there exists a vector of polynomials $h_p$ such that $a\nabla p = h_p p$. Furthermore assume the initial distribution $\mu_0$ to be such that $\mu_0((0, 1)) = 1$.

Finally, let $\bar{z} \in [0, 1]^2 \cap \{(x, y) \in [0, 1]^2 \mid p(x, y) = 0\}$ be such that

$$Ap(\bar{z}) \geq 0 \quad \text{and} \quad 2Ap(\bar{z}) - h_p(\bar{z})^T \nabla p(\bar{z}) < 0.$$
Then for any \( s > 0 \), there exists \( \varepsilon > 0 \) such that

\[
\mathbb{P}^{\mu_0}\{\|(X(0), Y(0)) - \tilde{z}\| < \varepsilon\} = 1 \Rightarrow \mathbb{P}^{\mu_0}\{\inf\{t \geq 0 \mid p(X(t), Y(t)) = 0\} \leq s\} > 0.
\]

As mentioned, the proof of this proposition consists mainly of ensuring the set-up is as in Theorem 5.7 in Filipović and Larsson (2016).

**Proof** We have already observed that the SDE (2) is indeed a polynomial diffusion on \([0, 1]^2\). The set of polynomials \(P\) describes the state space of our diffusion by

\[
[0, 1]^2 = \{(x, y) \in [0, 1]^2 \mid \forall p \in P: p(x, y) \geq 0\}
\]
as required. The only further assumption in Theorem 5.7 is the requirement that \(\{t \geq 0 \mid p(X(t), Y(t)) = 0\}\) be a Lebesgue null set. However, as is immediate from the last paragraph of the proof, this is only required in order to allow the process to start on the boundary. Since we assume our process to start in the interior of \([0, 1]^2\) a.s., this requirement is not necessary. Hence the proposition follows directly from Theorem 5.7, (iii), in Filipović and Larsson (2016).

We now turn to the proof of the theorem.

**Proof of Theorem 3.1** We will use the notation from Eq. (11) and begin with a short observation helpful for both parts of the proof.

Let \(p_0(x, y) := x \in P\). For \(h_{p_0}(x, y) := (\alpha^2(1 - x), 0)^T\) we have

\[
a \nabla p_0(x, y) = \begin{pmatrix} \alpha^2 x(1 - x) & 0 \\ 0 & (\alpha')^2 y(1 - y) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = x \begin{pmatrix} \alpha^2 (1 - x) \\ 0 \end{pmatrix} = p_0(x, y) h_{p_0}(x, y).
\]

Similarly, let \(p_1(x, y) := 1 - x \in P\). For \(h_{p_1}(x, y) := (-\alpha^2 x, 0)^T\) we have

\[
a \nabla p_1(x, y) = \begin{pmatrix} \alpha^2 x(1 - x) & 0 \\ 0 & (\alpha')^2 y(1 - y) \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = (1 - x) \begin{pmatrix} -\alpha^2 x \\ 0 \end{pmatrix} = p_1(x, y) h_{p_1}(x, y).
\]

**Part 1** “\(\Rightarrow\)” We begin proving the ‘only if’ statements, as they rely on the Proposition 3.3 we just introduced. Let \(\tilde{z} := (0, 0)\). Then \(p_0(\tilde{z}) = 0\),

\[
Ap_0(\tilde{z}) = u_2 \geq 0 \quad \text{and} \quad 2Ap_0(\tilde{z}) - h_{p_0}(\tilde{z})^T \nabla p_0(\tilde{z}) = 2u_2 - \alpha^2 < 0
\]
where the latter holds if, and only if \(2u_2 < \alpha^2\). Hence the ‘only if’ in (i) follows by Proposition 3.3.

In the same way consider instead \(\bar{z} := (1, 1)\). Then \(p_1(\bar{z}) = 0\),

\[
Ap(\bar{z}) = u_1 \geq 0 \quad \text{and} \quad 2Ap(\bar{z}) - h_p(\bar{z})^T \nabla p(\bar{z}) = 2u_1 - \alpha^2 < 0
\]

and again, the latter holds, if, and only if \(2u_1 < \alpha^2\). Therefore, the ‘only if’ in (ii) follows from Proposition 3.3 as well.

The analogous statements in (iii) and (iv) hold by symmetry.

**Part 2 “\(\Leftarrow\)”:** We now turn to the proof of the ‘if’ statements, which is more involved and uses McKean’s argument as it similar to the proof of Proposition 2.2 in Larsson and Pulido (2017). The approach is the same for all four cases, whence we can start with general observations and only check the different cases in the very end. Recall that we assumed the initial distribution \(\mu_0\) to be such that \(\mu_0((0, 1)^2) = 1\).

Take \(p \in \mathcal{P}\) and let \(h_p\) be a vector of polynomials such that \(a \nabla p = h_p p\) (we saw in Part 1 that this always exists). Choose \(z_p \in \{(x, y) \in [0, 1]^2 \mid p(x, y) \neq 0\}\) and define

\[
\tau_p := \inf\{t \geq 0 \mid p(X(t), Y(t)) = 0\}.
\]

Note that each of the \(\tau_p\) corresponds to one of the stopping times defined before Theorem 3.1, hence, we want to prove that \(\mathbb{P}^{\mu_0}(\tau_p < \infty) = 0\).

Itô’s formula and the identity \(a \nabla p = h_p p\) yield

\[
\log p(X(t), Y(t)) = \log p(X(0), Y(0)) + \int_0^t \left( \frac{Ap(X(s), Y(s))}{p(X(s), Y(s))} - \frac{1}{2} \frac{\nabla p(T) a \nabla p(X(s), Y(s))}{p(X(s), Y(s))^2} \right) \, ds
\]

\[
+ \int_0^t \nabla p(T) \sigma(X(s), Y(s)) \frac{\sigma(X(s), Y(s))}{p(X(s), Y(s))} \, dW_s
\]

\[
= \log p(X(0), Y(0)) + \int_0^t \frac{2Ap(X(s), Y(s)) - h_p^T h_p(X(s), Y(s))}{2p(X(s), Y(s))} \, ds
\]

\[
=: P(t)
\]

\[
+ \int_0^t \nabla p(T) \sigma(X(s), Y(s)) \frac{\sigma(X(s), Y(s))}{p(X(s), Y(s))} \, dW_s
\]

\[
=: M(t)
\]

for any \(t < \tau_p\). Suppose now, we find a constant \(\kappa_p > 0\) such that

\[
2Ap(x, y) - h_p^T \nabla p(x, y) \geq -2\kappa_p p(x, y) \quad \text{for all } (x, y) \in [0, 1]^2.
\]

Then \(P\) is adapted and càdlàg, defined on \([0, \tau_p)\) and

\[
\inf_{t \in [0, \tau_p \wedge T]} P(t) \geq -\kappa_p T > -\infty
\]
Since $M$ is a continuous local martingale on $[0, \tau_p)$ with $M(0) = 0$, Proposition 4.3 in Mayerhofer et al. (2011) implies $\tau_p = \infty$, $\mathbb{P}^{\mu_0}$-almost surely (for details concerning stochastic processes on stochastic intervals see for example Maisonneuve 1977).

Recall $p_0(x, y) = x$ and the assumption in (i) that $2u_2 - \alpha^2 \geq 0$. Set

$$\kappa_0 := u_1 + u_2 + c - \alpha^2/2 > 0$$

(since $c > 0$) and observe that then

$$2Ap_0(x, y) - h_{p_0}^T \nabla p_0(x, y) = x(-2u_1 - 2u_2 - 2c + \alpha^2) + y2c + 2u_2 - \alpha^2 \geq x(-2u_1 - 2u_2 - 2c + \alpha^2) + 2u_2 - \alpha^2 \geq -2\kappa_0 x = -2\kappa p_0(x, y) \quad \text{for all } (x, y) \in [0, 1]^2.$$

Hence (12) holds for $p_0$ and since $\tau_{p_0} = \tau_0^X$, the proof of (i) is completed.

For (ii) we assumed $2u_1 - \alpha^2 \geq 0$ and will use $p_1(x, y) = 1 - x$. Set

$$\kappa_1 := u_2 + c > 0,$$

since then

$$2Ap_1(x, y) - h_{p_1}^T \nabla p_1(x, y) = x(2u_1 + 2u_2 + 2c - \alpha^2) - y2c - 2u_2 \geq x(2u_1 + 2u_2 + 2c - \alpha^2) - 2c - 2u_2 \geq -2\kappa_1(1 - x) = -2\kappa_1 p_1(x, y) \quad \text{for all } (x, y) \in [0, 1]^2.$$

Again, (12) holds for $p_1$ and the equality $\tau_{p_1} = \tau_1^X$ completes the proof of (ii).

As before, the remaining statements follow by symmetry. \qed

**Theorem 3.4** In the set-up of Theorem 3.1, assume $2u_2 < \alpha^2$. Then for every $s > 0$ and every distribution $\mu_0$ with $\mu_0((0, 1)^2) = 1$

$$\mathbb{P}^{\mu_0} \left\{ \tau_0^X \leq s \right\} > 0 \quad \text{and} \quad \mathbb{P}^{\mu_0} \left\{ \tau_{0,1}^X < \infty \right\} = 1,$$

where $\tau_{0,1}^X := \min\{\tau_0^X, \tau_1^X\}$. Analogous results hold for the cases (ii), (iii) and (iv).

**Remark 3.5** Note that a result of the type $\mathbb{P}^{\mu_0} \left\{ \tau_0^X < \infty \right\} = 1$ cannot hold without further assumptions. In the case without mutation, for example, i.e. when $u_1 = u_1' = u_2 = u_2' = 0$, we have almost sure convergence of the diffusion to a random variable $(X_\infty, Y_\infty)$ whose distribution is given by

$$\frac{\delta_{(1,1)}}{c^+c^-} + \frac{(1-y)c+(1-x)c'}{c^+c^-} \delta_{(0,0)}$$

(see Remark 2.3) and therefore a positive probability of never reaching $[0] \times [0, 1]$.

**Proof** The proof is based on a coupling argument and the strong Markov property of the diffusion $(X(t), Y(t))_{t \geq 0}$. Using the same independent standard Brownian motions
(B(t))_{t \geq 0} and (B'(t))_{t \geq 0} from (2), define the diffusions (\tilde{X}(t))_{t \geq 0} and (\tilde{Y}(t))_{t \geq 0} with 
\tilde{X}(0) := X(0), \tilde{Y}(0) := Y(0) and

\[ d\tilde{X}(t) = \left[-u_1 \tilde{X}(t) + (u_2 + c)(1 - \tilde{X}(t))\right]dt + \alpha \sqrt{\tilde{X}(t)(1 - \tilde{X}(t))}dB_t, \]
\[ d\tilde{Y}(t) = \left[-u'_1 \tilde{Y}(t) + (u'_2 + c')(1 - \tilde{Y}(t))\right]dt + \alpha' \sqrt{\tilde{Y}(t)(1 - \tilde{Y}(t))}dB'_t. \]

Then for any initial distribution \( \mu_0 \) we have

\[ \mathbb{P}^{\mu_0} \left\{ \forall \ t \geq 0: X(t) \leq \tilde{X}(t) \text{ and } Y(t) \leq \tilde{Y}(t) \right\} = 1. \]

For any \( z \in [0, 1] \) define

\[ \bar{\tau}^X_z := \inf \left\{ t \geq 0 \mid X(t) \leq z \right\} \]

and likewise for \( \bar{\tau}^X_z, \bar{\tau}^Y_z \) and \( \bar{\tau}^\tilde{Y}_z \). Now assume \( \mu_0((0, 1)^2) = 1 \).

Since \( (\tilde{X}(t))_{t \geq 0} \) and \( (\tilde{Y}(t))_{t \geq 0} \) are simply Wright–Fisher diffusions (with mutation), it is easy to verify calculating their speed measure and scale function (cf. Chapter 3.3, Etheridge 2011), that for every \( s > 0 \) and every \( z \in (0, 1) \)

\[ \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^\tilde{X}_z \leq s \right\} > 0 \quad \text{and} \quad \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^\tilde{Y}_z \leq s \right\} > 0. \quad (13) \]

Furthermore, define

\[ \bar{\tau}^{(X,Y)}_z := \inf \left\{ t \geq 0 \mid X(t) \leq z, Y(t) \leq z \right\} \]

and likewise \( \bar{\tau}^{(X,Y)}_z \) for any \( z \in (0, 1) \). Since \( (\tilde{X}(t))_{t \geq 0} \) and \( (\tilde{Y}(t))_{t \geq 0} \) are independent, we get, for any \( s > 0 \) and \( z \in (0, 1) \)

\[ \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^{(X,Y)}_z \leq s \right\} \geq \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^\tilde{X}_z \leq s, \bar{\tau}^\tilde{Y}_z \leq s \right\} \]
\[ = \mathbb{P}^{\mu_0} \left[ \mathbb{1}_{\{\bar{\tau}^\tilde{X}_z \leq s\}} \mathbb{1}_{\{\bar{\tau}^\tilde{Y}_z \leq s\}} \theta_{\bar{\tau}^\tilde{X}_z} \mathbb{1}_{\{\bar{\tau}^\tilde{Y}_z \leq s\}} \right] > 0 \]
\[ = q(\bar{\tau}^\tilde{X}_z) \]

where \( q(t) := \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^\tilde{Y}_z \leq s \right\} \) and positivity follows from (13). This in turn implies for any \( s > 0 \) and \( z \in (0, 1) \)

\[ \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^{(X,Y)}_z \leq s \right\} \geq \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^{(\tilde{X},\tilde{Y})}_z \leq s \right\} > 0. \]

Now assume \( 2u_2 < \alpha^2 \), fix \( s > 0 \) and let \( \varepsilon = \varepsilon(s/2) > 0 \) be as in Remark 3.2. Then choose \( z \in (0, 1) \) such that \( \| (z, z) - (0, 0) \| < \varepsilon \). Finally, using the strong Markov
property we conclude that

\[
\mathbb{P}^{\mu_0} \left\{ \tau_0^X \leq s \right\} \geq \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^X_0 \leq s, \bar{\tau}^{(X,Y)}_C \leq s/2 \right\} = \mathbb{P}^{\mu_0} \left\{ \tau_0^X \leq s \mid \bar{\tau}^{(X,Y)}_C \leq s/2 \right\} \mathbb{P}^{\mu_0} \left\{ \bar{\tau}^{(X,Y)}_C \leq s/2 \right\} > 0 \text{ by (13)}
\]

where we used the strong Markov property and the observation that \(\mathbb{P}^{(z,z)} \left\{ \tau_0^X \leq s/2 \right\} \leq \mathbb{P}^{(x,y)} \left\{ \tau_0^X \leq s/2 \right\} \) for any \(x \leq z\) and \(y \leq z\) together with Theorem 3.1. This proves the first claim.

Recall that \(s\) is fixed. In order to see the second, let \(g(x, y) := \mathbb{P}^{x,y} \left\{ \tau_{0,1}^X \leq s \right\}\) and observe that \(g: [0, 1]^2 \rightarrow [0, 1]\) is a continuous function, since we have a Feller diffusion. Therefore, there exists a point \((x_{\min}, y_{\min}) \in [0, 1]^2\) such that \(g(x_{\min}, y_{\min}) \leq g(x, y)\) for all \(x, y \in [0, 1]\). Since \(g(1, y) = g(0, y) = 1\) for all \(y \in [0, 1]\), we can assume \((x_{\min}, y_{\min}) \in (0, 1) \times [0, 1]\). We want to show that \(g(x_{\min}, y_{\min}) > 0\). If \((x_{\min}, y_{\min}) \in (0, 1)^2\), this follows immediately by the first claim we proved and

\[
g(x_{\min}, y_{\min}) \geq \mathbb{P}^{x_{\min},y_{\min}} \left\{ \tau_0^X \leq s \right\} > 0.
\]

If, on the other hand, \((x_{\min}, y_{\min}) \in (0, 1) \times [0, 1]\), choose \(\varepsilon \in (0, s)\) and use the Markov property to see

\[
g(x_{\min}, y_{\min}) \geq \mathbb{E}^{x_{\min},y_{\min}} \left[ \mathbb{E}^{x_{\min},y_{\min}} \left[ \mathbb{1}_{\{\varepsilon < \tau_{0,1}^X \leq s\}} \mid \mathcal{F}_\varepsilon \right] \right] \\
\geq \mathbb{E}^{x_{\min},y_{\min}} \mathbb{E}^{X(\varepsilon),Y(\varepsilon)} \left[ \mathbb{1}_{\{\tau_{0,1}^X \leq s-\varepsilon\}} \right] \\
\geq \mathbb{E}^{x_{\min},y_{\min}} \mathbb{E}^{X(\varepsilon),Y(\varepsilon)} \left[ \mathbb{1}_{\{(X(\varepsilon), Y(\varepsilon)) \in (0,1)^2\}} \right] > 0
\]

using the first claim, since \(\mathbb{P}^{x_{\min},y_{\min}} \left\{ (X(\varepsilon), Y(\varepsilon)) \in (0,1)^2 \right\} > 0\). Using the Markov property yet again, we conclude that

\[
\mathbb{P}^{\mu_0} \left\{ \tau_{0,1}^X \leq i s \right\} \geq 1 - (1 - g(x_{\min}, y_{\min}))^i
\]

which then implies \(\mathbb{P}^{\mu_0} \left\{ \tau_{0,1}^X < \infty \right\} = 1\) as desired by continuity of measures. \(\Box\)

### 4 Discussion

Both the seed bank as well as the two-island model describe a population divided into two subpopulations with migration between them. The respective diffusions model the
scaling limit as the population size $N$ grows to infinity (with time measured in units of $N$) of the proportion of $a$-alleles in a Wright–Fisher model with two subpopulations, in a bi-allelic set-up, i.e. with a type-space \{a, A\}, with migration between them and possibly mutation between the two types.

The so-called strong seed bank is modelled by dividing the total population into the “active” and the “dormant” subpopulation. See Blath et al. (2015) for a thorough discussion of the scaling and modelling assumptions and the biological motivation. The key feature is that reproduction is turned off for the dormant individuals in the seed bank. Therefore the formally two-dimensional seed bank diffusion, describing the allele frequencies in the active and dormant population, is driven by an only one-dimensional “reproductive noise”, and this is precisely what allowed us the reformulation in terms of a one-dimensional stochastic delay differential equation (Proposition 1.4). This delay equation provides an intuitive interpretation of the seed bank mechanism as determining the type of a sampled individual according to an individual alive an exponential amount of time ago (on the macroscopic timescale of order $N$), and provides a link to classical “weak” seed bank models as in Kaj et al. (2001). This type of representation appears to be entirely new in population genetics models.

In the two-island model each subpopulation is considered to live on a separate island with migration allowed between them. The difference between the corresponding two-island diffusion and the seed bank diffusion lies in the fact that reproduction in both subpopulations is possible. Otherwise, many structural properties are similar, in particular those describing the mutation mechanism and migration between sub-populations. These formal similarities of both systems have indeed allowed for the application of similar mathematical tools, despite the fact that the genealogical processes predict several qualitative differences, and that the underlying population genetic models have different biological motivations.

Both models have been analyzed with the help of certain dual processes before. However, in Sect. 2, we apply a special variant of a moment duality derived from the so-called “sampling-dualities” as discussed in González Casanova and Spanò (2018) to analyze the dynamical behavior. In the presence of mutation, we show the convergence of the expected allele frequencies (as well as of all higher moments) to expectation (or the respective moments) of a non-trivial equilibrium (Propositions 2.4, 2.6). The equilibrium is characterized through its moments for which we obtain a recursive formula in Lemma 2.7. This can be interpreted as a sampling formula for two-allelic populations that have evolved under stable conditions for a long time, and whose allele frequencies are thus “almost in equilibrium”. This is useful for practical purposes because if it is assumed that the frequency of individuals of certain type is close to equilibrium, different parameters can be estimated from biological data by comparing the moments of the empirical data with those of the stationary distribution obtained using this theoretical model.

The recursions also show how the two models behave differently regarding their allele frequencies at stationarity. For example, while the first moments, i.e. the expectations, coincide, the second moments, which then correspond to the variance, are smaller in the seed bank model, as shown in Figs. 2 and 3. The difference is most prominent in the second island/seed population if we don’t allow mutation there: The
Fig. 3 The second moments $M_{2,0}$ (left) and $M_{0,2}$ (right) for a varying migration rate $c$. The parameters $K = 1$ and $u_1 = u_2 = 0.1$ are fixed. Red: seed bank ($\alpha' = 0$) with no mutation in the seeds ($u'_1 = u'_2 = 0$). Magenta: seed bank ($\alpha' = 0$) with mutation in the seed bank $u'_1 = u'_2 = u_1$. Dark blue: two-island ($\alpha' = 1$) with no mutation in the second island ($u'_1 = u'_2 = 0$). Light blue: two-island ($\alpha' = 1$) with mutation in the second island ($u'_1 = u'_2 = u_1$) (colour figure online).

The variance in the seed bank case is pushed to 0 by increased mutation rate whereas it tends to a positive constant for the two-island case, see the right box in Fig. 2. However, there is a delicate interplay of parameters that require a careful statistical analysis to differentiate them. In order to exemplify how the effect of different forces can cancel out, observe that the moments $M_{2,0}$ of a seed bank model without mutation in the seeds and a two-island model with mutation in both islands coincide exactly for several combinations of parameters (see the left box of Fig. 2 and more clearly Fig. 3), if $K = 1$. An upcoming statistical article inspired in some ideas discussed in this paper will provide a deeper discussion of the statistical methods to distinguish between models based on biological data.

Finally, we have investigated the boundary behaviour of both allele frequency processes. By this we mean giving an answer to the question whether, depending on the model parameters (in particular the mutation rates), the allele frequency process of a subpopulation will always stay away from boundary 1 (resp. 0). If the process never reaches 1, this means that the $a$-alleles will never completely dominate the whole population. In such a case one speaks of an unattainable boundary. Such boundary classifications have been investigated with the help of Feller’s diffusion theory in mathematics for many one-dimensional diffusions, but this formalism is not available in our two-dimensional case. However, we can make use of the polynomial diffusion framework recently developed for mathematical finance (cf. Filipović and Larsson 2016; Larsson and Pulido 2017) to answer the above question, which had been left open in Kermany et al. (2008) for the two-island model several years ago. Our criterion states that in the active population of the seed bank model, boundary 1 can never been reached if and only if the mutation rate $u_2$ exceeds $\frac{\alpha}{2}$. This means that the mutation rate $u_2$ (producing type $A$ individuals and thus driving the allele frequency away from boundary 1) needs to be large enough in comparison to the effective population size (resp. diffusion coefficient) given by $\alpha$. In biological terms, if the mutation rate is big enough compared to the random genetic drift, the allele frequencies of both type of individuals will be positive at all times.
Acknowledgements The authors would like to thank S. Pulido for pointing out the connection to polynomial diffusions and the reviewers for helpful suggestions. JB and MWB were supported by DFG Priority Programme 1590 “Probabilistic Structures in Evolution”, Project BL 1105/5-1, MWB also by Project KU 2886/1-1 awarded to N. Kurt. MWB would like to thank the TU Berlin for their hospitality. EB received support from the Berlin Mathematical School and the DFG RTG 1845 “Stochastic Analysis with applications in biology, finance and physics”. AGC was supported by CONACYT, Project FC-2016-1946.

References

Blath J, González Casanova A, Eldon B, Kurt N, Wilke-Berenguer M (2015) Genetic variability under the seedbank coalescent. Genetics 200(3):921–934
Blath J, González Casanova A, Kurt N, Wilke-Berenguer M (2016) A new coalescent for seed-bank models. Ann Appl Probab 26(2):857–891
den Hollander F, Pederzani G (2017) Multi-colony Wright–Fisher with seed-bank. Indag Math 28(3):637–669
Etheridge A (2011) Some mathematical models from population genetics, volume 2012 of lecture notes in mathematics. Springer, Heidelberg. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009. École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]
Etheridge AM, Griffiths RC (2009) A coalescent dual process in a Moran model with genic selection. Theor Popul Biol 75(4):320–330 (Sam Karlin: Special Issue)
Etheridge AM, Griffiths RC, Taylor JE (2010) A coalescent dual process in a Moran model with genic selection, and the lambda coalescent limit. Theor Popul Biol 78(2):77–92
Ethier SN, Kurtz TG (1992) On the stationary distribution of the neutral diffusion model in population genetics. Ann Appl Probab 2(1):24–35
Filipović D, Larsson M (2016) Polynomial diffusions and applications in finance. Finance Stoch 20(4):931–972
Fu R, Gelfand AE, Holsinger KE (2003) Exact moment calculations for genetic models with migration, mutation, and drift. Theor Popul Biol 63(3):231–243 (Uses of DNA and genetic markers for forensics and population studies)
González Casanova A, Spanò D (2018) Duality and fixation for \( \Xi \)-Wright-Fisher processes with frequency-dependent selection. Ann Appl Probab 28(1):250–284
Griffiths RC, Jenkins PA, Lessard S (2016) A coalescent dual process for a Wright–Fisher diffusion with recombination and its application to haplotype partitioning. Theor Popul Biol 112(Supplement C):126–138
Hildebrandt TH, Schoenberg JJ (1933) On linear functional operations and the moment problem for a finite interval in one or several dimensions. Ann Math 34(2):317–328
Jansen S, Kurt N (2014) On the notion(s) of duality for Markov processes. Probab Surv 11:59–120
Kaj I, Krone SM, Lascoux M (2001) Coalescent theory for seed bank models. J Appl Probab 38(2):285–300
Kermany ARR, Zhou X, Hickey DA (2008) Joint stationary moments of a two-island diffusion model of population subdivision. Theor Popul Biol 74(3):226–232
Krone SM, Neuhauser C (1997) Ancestral processes with selection. Theor Popul Biol 51(3):210–237
Lambert A, Ma C (2015) The coalescent in peripatric metapopulations. J Appl Probab 52(2):538–557
Larsson M, Pulido S (2017) Polynomial diffusions on compact quadric sets. Stoch Process Appl 127(3):901–926
Lennon JT, Jones SE (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 9(2):119–130
Maisonneuve B (1977) Une mise au point sur les martingales locales continues définies sur un intervalle stochastique. In: Lecture notes in mathematics, vol 581, pp 435–445
Mayerhofer E, Pfaffel O, Stelzer R (2011) On strong solutions for positive definite jump diffusions. Stoch Process Appl 121(9):2072–2086
Moran PAP (1959) The theory of some genetical effects of population subdivision. Austral J Biol Sci 12(2):109–116
Nath HB, Griffiths RC (1993) The coalescent in two colonies with symmetric migration. J Math Biol 31(8):841–851
Revuz D, Yor M (1999) Continuous martingales and Brownian motion, volume 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental principles of mathematical sciences], 3rd edn. Springer, Berlin
Shiga T, Shimizu A (1980) Infinite-dimensional stochastic differential equations and their applications. J Math Kyoto Univ 20(3):395–416
Shoemaker WR, Lennon JT (2018) Evolution with a seed bank: the population genetic consequences of microbial dormancy. Evol Appl. 11(1):60–75
Wakeley J (ed) (2008) Coalescent theory: an introduction. Roberts and Company, Greenwood Village, CO
Wright S (1931) Evolution in Mendelian populations. Genetics 16(2):97–159

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.