

Slow-fast stochastic diffusion dynamics and quasi-stationarity for diploid populations with varying size

Camille Coron

the date of receipt and acceptance should be inserted later

Abstract

We are interested in the long-time behavior of a diploid population with sexual reproduction and randomly varying population size, characterized by its genotype composition at one bi-allelic locus. The population is modeled by a 3-dimensional birth-and-death process with competition, cooperation and Mendelian reproduction. This stochastic process is indexed by a scaling parameter K that goes to infinity, following a large population assumption. When the individual birth and natural death rates are of order K , the sequence of stochastic processes indexed by K converges toward a new slow-fast dynamics with variable population size. We indeed prove the convergence toward 0 of a fast variable giving the deviation of the population from Hardy-Weinberg equilibrium, while the sequence of slow variables giving the respective numbers of occurrences of each allele converges toward a 2-dimensional diffusion process that reaches $(0, 0)$ almost surely in finite time. The population size and the proportion of a given allele converge toward a Wright-Fisher diffusion with stochastically varying population size and diploid selection. We insist on differences between haploid and diploid populations due to population size stochastic variability. Using a non trivial change of variables, we study the absorption of this diffusion and its long time behavior conditioned on non-extinction. In particular we prove that this diffusion starting from any non-trivial state and conditioned on not hitting $(0, 0)$ admits a unique quasi-stationary distribution. We give numerical approximations of this quasi-stationary behavior in three biologically relevant cases: neutrality, overdominance, and separate niches.

Keywords Diploid populations · demographic Wright-Fisher diffusion processes · stochastic slow-fast dynamical systems · quasi-stationary distributions · allele coexistence.

Camille Coron
CMAP, École Polytechnique, CNRS UMR 7641, Route de Saclay, 91128 Palaiseau Cedex, France
E-mail: coron@cmmap.polytechnique.fr

Mathematics Subject Classification (2000) 60J10 · 60J80 · 60J27 · 92D25 · 92D15

1 Introduction

We study the diffusion limit and quasi-stationary behavior of a population of diploid individuals modeled by a non-linear 3-type birth-and-death process with competition, cooperation and Mendelian reproduction. Individuals are characterized by their genotype at one locus for which there exist 2 alleles, A and a . We study the genetic evolution of the population, i.e. the dynamics of the respective numbers of individuals with genotype AA , Aa , and aa . Following an infinite population size approximation (also studied in Fournier and Méléard (2004) and Champagnat (2006) for instance, to study the adaptive dynamics of haploid populations) we assume that the initial number of individuals is of order K where K is a scale parameter that will go to infinity. The population is then modeled by a 3-type birth-and-death process denoted by $\nu^K = (\nu_t^K, t \geq 0)$ and we consider the sequence of stochastic processes $Z^K = \nu^K / K$. At each time t and for all K , we define the deviation Y_t^K of the population Z_t^K from a so-called Hardy-Weinberg structure. We are interested in the convergence of the sequence of stochastic processes Z^K under a weak-selection regime: the individual birth and natural death rates are assumed to be both equivalent to γK , with $\gamma > 0$, which corresponds to a diffusive scaling (a biological interpretation is also given in Champagnat et al. (2006)). Such scalings have been studied notably in Papanicolaou et al. (1977), Ethier and Nagylaki (1980), Ethier and Nagylaki (1988) and Katzenberger (1991), and here we consider models and asymptotic dynamics with interactions and randomly varying population size, which rises both new complications and perspectives, in particular linked to population size explosion and population extinction. In Section 3 we first establish some conditions on the competition and cooperation parameters so that the sequence of population sizes satisfies a moment propagation property. Next, we prove the convergence of the sequence of stochastic processes Z^K toward a slow-fast dynamics (see Méléard and Tran (2012) or Ball et al. (2006) for other examples of such dynamics and Kurtz (1992) and Berglund and Gentz (2006) for treatments of slow-fast scales in diffusion processes). More precisely, we prove that for all $t > 0$, the sequence of random variables $(Y_t^K)_{K \in \{1, 2, \dots\}}$ goes to 0 as K goes to infinity, while the sequence of processes $(N_t^K, X_t^K)_{t \geq 0}$ giving respectively the

population size and the proportion of allele A converges in law toward a 2-dimensional diffusion process $(N_t, X_t)_{t \geq 0}$. Under this large population and weak-selection regime, even with a stochastic dynamics of population size, the convergence toward Hardy-Weinberg structure obtained for deterministic models for instance in Norton (1928), Hofbauer et al. (1982), Hofbauer and Sigmund (1998), Hoppensteadt (1975) and for stochastic models with constant population size for instance in Nagylaki (1992) (Chap. 4.10), remains true. The limiting diffusion $(N_t, X_t)_{t \geq 0}$ can be seen as a Wright-Fisher diffusion with competition and general diploid selection, associated to a randomly varying population size. An important remark is that, due to the stochasticity in population size dynamics, this diffusion cannot be obtained by a time change of the corresponding haploid one obtained from Cattiaux and Méléard (2010). In Section 4, we first find an appropriate change of variables $S = (f_1(N, X), f_2(N, X))$ such that S is a Kolmogorov diffusion process, (i.e. a diffusion process with unit diffusion coefficient and gradient drift), evolving in a subset \mathcal{D} of \mathbb{R}^2 . The stochastic process S is absorbed in the set $\mathbf{A} \cup \mathbf{a} \cup \mathbf{0}$ almost surely in finite time, where \mathbf{A} , \mathbf{a} and $\mathbf{0}$ correspond respectively to the sets where $X = 1$ (fixation of allele A), $X = 0$ (fixation of allele a), and $N = 0$ (extinction of the population). This extinction event is an important feature of our model and we study the quasi-stationary behavior of the diffusion process $(S_t)_{t \geq 0}$ conditioned on the non extinction of the population, i.e. on not reaching $\mathbf{0}$. First, the diffusion process $(S_t)_{t \geq 0}$ conditioned on not reaching $\mathbf{A} \cup \mathbf{a} \cup \mathbf{0}$ admits a Yaglom limit, i.e. the law of S_t conditioned on $\{S_t \notin \mathbf{A} \cup \mathbf{a} \cup \mathbf{0}\}$ converges as t goes to infinity toward a distribution which is independent from S_0 . Second, the law of S_t conditioned on $\{S_t \notin \mathbf{0}\}$ also converges as t goes to infinity toward a distribution which is independent from S_0 , if $S_0 \notin \mathbf{A} \cup \mathbf{a} \cup \mathbf{0}$. Finally in Section 5, we present numerical applications and study the long-time coexistence of the two alleles for three biologically relevant cases: a pure competition neutral case, a case in which each genotype has its own ecological niche, and an overdominance case. In particular, we show that a long-term coexistence of alleles is possible even in some full competition cases, which is not true for haploid clonal reproduction (Cattiaux and Méléard (2010)). Note that for the sake of simplicity, most proofs of this article are given in the main text for the neutral case where demographic parameters do not depend on the types of individuals, and the calculations for the non-neutral case are given in Appendix A.

2 Model

We consider a population of diploid hermaphroditic individuals characterized by their genotype at one bi-allelic locus, whose alleles are denoted by A and a . Individuals can then have one of the three possible genotypes AA , Aa , and aa , (also called types 1, 2, and 3). The population at any time t is represented by a 3-dimensional birth-and-death process giving the respective numbers of individuals with each genotype. As in Fournier and Méléard (2004), Champagnat and Méléard (2007) or Collet et al. (2013), we consider an infinite population size approximation. To this end we index the population by a scaling parameter $K \in \{1, 2, \dots\}$ that will go to infinity. The initial numbers of individuals of each type will be of order K and we then consider the sequence of rescaled stochastic processes

$$(Z_t^K)_{t \geq 0} = (Z_t^{1,K}, Z_t^{2,K}, Z_t^{3,K})_{t \geq 0}$$

that give at each time t the respective numbers of individuals weighted by $1/K$, with genotypes AA , Aa , and aa . The rescaled population size at time t is denoted by

$$N_t^K = Z_t^{1,K} + Z_t^{2,K} + Z_t^{3,K} \in \frac{\mathbb{Z}_+}{K}, \quad (2.1)$$

and the proportion of allele A at time t is denoted by

$$X_t^K = \frac{2Z_t^{1,K} + Z_t^{2,K}}{2(Z_t^{1,K} + Z_t^{2,K} + Z_t^{3,K})}. \quad (2.2)$$

As in Coron (2014) or Collet et al. (2013), the jump rates of Z model Mendelian panmictic reproduction. More precisely, if we set $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$, then for all $i \in \{1, 2, 3\}$, the rates $\lambda_t^K(Z)$ at which the stochastic process Z^K jumps from $z = (z_1, z_2, z_3) \in \left(\frac{\mathbb{Z}_+}{K}\right)^3$ to $z + e_i/K$, as long as $z_1 + z_2 + z_3 = n \neq 0$, are given by:

$$\begin{aligned}
\lambda_1^K(z) &= \frac{Kb_1^K}{n} \left(z_1 + \frac{z_2}{2} \right)^2, \\
\lambda_2^K(z) &= \frac{Kb_2^K}{n} 2 \left(z_1 + \frac{z_2}{2} \right) \left(z_3 + \frac{z_2}{2} \right), \\
\lambda_3^K(z) &= \frac{Kb_3^K}{n} \left(z_3 + \frac{z_2}{2} \right)^2.
\end{aligned} \tag{2.3}$$

These birth rates are naturally set to 0 if $n = 0$ and the demographic parameters $b_i^K \in \mathbb{R}_+$ are called birth demographic parameters. Now individuals can die naturally and either compete or cooperate with other individuals, depending on the genotype of each individual. More precisely, for all $i \in \{1, 2, 3\}$, the rates $\mu_i^K(z)$ at which the stochastic process Z^K jumps from $z = (z_1, z_2, z_3) \in (\mathbb{Z}_+)^3/K$ to $z - e_i/K$ for $i \in \{1, 2, 3\}$ are given by:

$$\begin{aligned}
\mu_1^K(z) &= Kz_1(d_1^K + K(c_{11}^K z_1 + c_{21}^K z_2 + c_{31}^K z_3))^+, \\
\mu_2^K(z) &= Kz_2(d_2^K + K(c_{12}^K z_1 + c_{22}^K z_2 + c_{32}^K z_3))^+, \\
\mu_3^K(z) &= Kz_3(d_3^K + K(c_{13}^K z_1 + c_{23}^K z_2 + c_{33}^K z_3))^+.
\end{aligned} \tag{2.4}$$

where the interaction (competition or cooperation) demographic parameters c_{ij}^K are arbitrary real numbers and $(x)^+ = \max(x, 0)$ for any $x \in \mathbb{R}$. If $c_{ij}^K > 0$ (resp. $c_{ij}^K < 0$), then individuals with type i have a negative (resp. positive) influence on individuals of type j . The demographic parameter $d_i^K \in \mathbb{R}_+$ is called the intrinsic death rate of individuals of type i . From now on, we say that the stochastic process Z^K is neutral if its demographic parameters do not depend on the types of individuals, i.e.

$$b_i^K = b^K, \quad d_i^K = d^K \quad \text{and} \quad c_{ij}^K = c^K \quad \forall i, j \in \{1, 2, 3\}. \tag{2.5}$$

Note that for any fixed $K \in \{1, 2, \dots\}$, the pure jump process Z^K is well defined for all $t \in \mathbb{R}_+$. Indeed, N^K is stochastically dominated by a rescaled pure birth process \overline{N}^K that jumps from $n \in \mathbb{Z}_+/K$ to $n + 1/K$ at rate $(\max_i b_i^K)Kn$ and, from Theorem 10 in Méléard and Villemonais (2012), \overline{N}^K does not explode almost surely. The stochastic process Z^K is then a $\frac{(\mathbb{Z}_+)^3}{K}$ -valued pure jump Markov process absorbed at $(0, 0, 0)$, defined for all $t \geq 0$ by

$$Z_t^K = Z_0^K + \sum_{i \in \{1,2,3\}} \left[\int_0^t \frac{e_i}{K} \mathbf{1}_{\{\theta \leq \lambda_i^K(Z_{s-}^K)\}} \eta_1^i(ds, d\theta) - \int_0^t \frac{e_i}{K} \mathbf{1}_{\{\theta \leq \mu_i^K(Z_{s-}^K)\}} \eta_2^i(ds, d\theta) \right]$$

where the measures η_j^i for $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$ are independent Poisson point measures on $(\mathbb{R}_+)^2$ with intensity $dsd\theta$. For any K , the law of Z^K is therefore a probability measure on the trajectory space $\mathbb{D}(\mathbb{R}_+, (\mathbb{Z}_+)^3/K)$ which is the space of left-limited and right-continuous functions from \mathbb{R}_+ to $(\mathbb{Z}_+)^3/K$, endowed with the Skorohod topology. Finally, the infinitesimal generator L^K of Z^K satisfies for every bounded measurable function f from $(\mathbb{Z}_+)^3/K$ to \mathbb{R} and for every $z \in \frac{(\mathbb{Z}_+)^3}{K}$:

$$L^K f(z) = \sum_{i \in \{1,2,3\}} \left[\lambda_i^K(z) \left(f\left(z + \frac{e_i}{K}\right) - f(z) \right) + \mu_i^K(z) \left(f\left(z - \frac{e_i}{K}\right) - f(z) \right) \right]. \quad (2.6)$$

To complete the model description, let us introduce for all $K \in \{1, 2, \dots\}$ the stochastic processes Y^K such that for every $t \geq 0$, as long as $N_t^K > 0$,

$$Y_t^K = \frac{4Z_t^{1,K} Z_t^{3,K} - (Z_t^{2,K})^2}{4N_t^K}. \quad (2.7)$$

If $N_t^K = 0$, we set $Y_t^K = 0$ as $|Y_t^K| \leq N_t^K$ for all $t \geq 0$. This stochastic process will play a main role in the article; note first that:

$$Y_t^K = Z_t^{1,K} - \frac{(2Z_t^{1,K} + Z_t^{2,K})^2}{4N_t^K} = \left(p_t^{AA,K} - (p_t^{A,K})^2 \right) N_t^K$$

if $p_t^{A,K}$ (resp. $p_t^{AA,K}$) is the proportion of allele A (resp. genotype AA) in the population at time t . Similarly,

$$Y_t^K = \left(p_t^{aa,K} - (p_t^{a,K})^2 \right) N_t^K = \left(2p_t^{A,K} p_t^{a,K} - p_t^{Aa,K} \right) N_t^K.$$

Then if $Y_t^K = 0$, the proportion of each genotype in the population Z_t^K is equal to the proportion of pairs of alleles forming this genotype. By an abuse of language, if $Y_t^K = 0$ we say that the population Z_t^K is at Hardy-Weinberg equilibrium (classically this equilibrium also requires constant allele proportions which is not satisfied here, as presented notably in Crow and Kimura (1970), p. 34). In the rest of the article, we will see that the quantities of interest in this model are the population size N^K , the deviation from Hardy-

Weinberg equilibrium Y^K and the proportion X^K of allele A . More precisely, the following lemma gives the change of variable:

Lemma 1 *Let us set for all $z = (z_1, z_2, z_3) \in (\mathbb{R}_+)^3 \setminus \{(0, 0, 0)\}$,*

$$\phi_1(z) = z_1 + z_2 + z_3, \quad \phi_2(z) = \frac{2z_1 + z_2}{2(z_1 + z_2 + z_3)} \quad \text{and} \quad \phi_3(z) = \frac{4z_1z_3 - (z_2)^2}{4(z_1 + z_2 + z_3)}.$$

Then the function

$$\begin{aligned} \phi : (\mathbb{R}_+)^3 \setminus \{(0, 0, 0)\} &\rightarrow E \\ z &\mapsto \phi(z) = (\phi_1(z), \phi_2(z), \phi_3(z)) \end{aligned}$$

where $E = \{(n, x, y) | n \in \mathbb{R}_+^, x \in [0, 1], -n \min(x^2, (1-x)^2) \leq y \leq nx(1-x)\}$ is a bijection.*

Note that $\phi(Z^K) = (N^K, X^K, Y^K)$, where N^K , X^K , and Y^K have been respectively defined in Equations (2.1), (2.2) and (2.7).

Proof We easily obtain that $(n, x, y) = \phi(z_1, z_2, z_3)$ if and only if

$$z_1 = nx^2 + y, \quad z_2 = 2nx(1-x) - 2y \quad \text{and} \quad z_3 = n(1-x)^2 + y, \quad (2.8)$$

which gives the injectivity. Next, for any (n, x, y) such that $n \in \mathbb{R}_+^*$, $x \in [0, 1]$ and $-n \min(x^2, (1-x)^2) \leq y \leq nx(1-x)$, $z = (z_1, z_2, z_3)$ defined by Equation (2.8) is in $(\mathbb{R}_+)^3 \setminus \{(0, 0, 0)\}$ which gives the surjectivity. \square

3 Convergence toward a slow-fast stochastic dynamics

In this section, we investigate a diffusive scaling under which the population size and the proportion of allele a evolve stochastically with time (in particular the population can get extinct and one of the two alleles can eventually get fixed), while the population still converges rapidly toward Hardy-Weinberg equilibrium. The results we obtain then provide a rigorous justification for populations with interactions and stochastically varying size, of the assumption of Hardy-Weinberg equilibrium which is often made when studying large

diploid populations and was studied for deterministic models in Norton (1928) and for stochastic models with constant population size in Nagylaki and Crow (1974). As we will see later, the limiting dynamics obtained here is never the same as the haploid population dynamics studied notably in Cattiaux and Méléard (2010), contrarily to what is commonly observed for populations with constant size and additive diploid selection (Depperschmidt et al. (2012)). In the model considered in the present article, due to the (stochastic) evolution of population size, even when the studied diploid population is at Hardy-Weinberg equilibrium and experiences additive selection, it behaves differently than haploid populations studied in Cattiaux and Méléard (2010): Indeed, the respective numbers of alleles A and a are directed by correlated Brownian motions in a diploid population, whereas they are directed by uncorrelated Brownian motions in a haploid population (see also Remark 3).

We assume that individual birth and natural death rates are of order K , while Z_0^K converges in law toward a random vector Z_0 as $K \rightarrow \infty$. More precisely, we set for $\gamma > 0$:

$$b^{i,K} = \gamma K + \beta_i \in [0, \infty[$$

$$d^{i,K} = \gamma K + \delta_i \in [0, \infty[$$

$$c_{ij}^K = \frac{\alpha_{ij}}{K} \in \mathbb{R}$$

$$Z_0^K \xrightarrow{K \rightarrow \infty} Z_0 \quad \text{in law,}$$

where Z_0 is a $(\mathbb{R}_+)^3$ -valued random variable. This means that the birth and natural death events are now happening faster and compensate each other, which will introduce some stochasticity in the limiting process. A biological interpretation for the scaling of the interaction coefficients c_{ij}^K is that a given quantity of resources is shared by individuals with biomass of order $1/K$ (as presented in Champagnat et al. (2006)); these coefficients will only step in limiting drift terms. Under this scaling of demographic parameters, Y^K will be a "fast" variable that converges directly toward the long time equilibrium of \mathcal{Y} (equal to 0, as studied in Ethier and Nagylaki (1980), Ethier and Nagylaki (1988) and Katzenberger (1991)), while X^K and N^K will be "slow" variables, converging toward a non-deterministic process.

Population size stochasticity induces complications linked to both population extinction and population size control, the latter being linked to interaction rates. For any $z = (z_1, z_2, z_3) \in (\mathbb{R}_+)^3$, let $g(z) =$

$\sum_{i,j \in \{1,2,3\}} \alpha_{ij} z_i z_j$. An important tool in this article is the following moment propagation property.

Proposition 1 *If $g(z) > 0$ for all $z \in (\mathbb{R}_+)^3 \setminus \{(0,0,0)\}$ and if, for any $k \in \mathbb{Z}_+$, there exists a constant C_0 such that for all $K \in \{1, 2, \dots\}$, $\mathbb{E}((N_0^K)^k) \leq C_0$, then*

(i) *There exists a constant C such that $\sup_K \sup_{t \geq 0} \mathbb{E}((N_t^K)^k) \leq C$.*

(ii) *For all $T < +\infty$, there exists a constant C_T such that $\sup_K \mathbb{E} \left(\sup_{t \leq T} (N_t^K)^k \right) \leq C_T$.*

Note that these properties of the model are not valid for all values of the interaction parameters α_{ij} and in particular when $\alpha_{ii} \leq 0$ for any $i \in \{1, 3\}$, i.e. when homozygous individuals that have the same genotype cooperate or at least do not compete. Indeed, from Coron (2014) (proof of Proposition 3.4) allele A has a strictly positive fixation probability and from this fixation time, if $\alpha_{11} \leq 0$, the population size is a one-type birth-and-death process with cooperation which does not satisfy this property.

Proof Assume that $g(z) > 0$ for all $z \in (\mathbb{R}_+)^3 \setminus \{(0,0,0)\}$. Note that g can be written as

$$g(z) = \phi_1(z)^2 \sum_{i,j \in \{1,2,3\}} \alpha_{ij} p_i p_j := \phi_1(z)^2 f(p_1, p_2, p_3)$$

where ϕ_1 has been defined in Lemma 1 and $p_i = z_i/\phi_1(z)$ for all $i \in \{1, 2, 3\}$. If $g(z) > 0$ for all $z \in (\mathbb{R}_+)^3 \setminus \{(0,0,0)\}$, the function f is then non-negative and continuous on $\{(p_1, p_2, p_3) \in [0, 1]^3 | p_1 + p_2 + p_3 = 1\}$ which is a compact set. Then f reaches its minimum $m \geq 0$. Now if there exists (p_1, p_2, p_3) such that $f(p_1, p_2, p_3) = 0$ then for any $n > 0$, $g(np_1, np_2, np_3) = 0$ and $(np_1, np_2, np_3) \in (\mathbb{R}_+)^3 \setminus \{(0,0,0)\}$ which is impossible. Then $m > 0$, $g(z) \geq m\phi_1(z)^2$, and $\mu_1^K(z) + \mu_2^K(z) + \mu_3^K(z) \geq (\gamma K + \inf_i \delta_i + m\phi_1(z))K\phi_1(z)$ for all $z \in (\mathbb{R}_+)^3$. Then, for all K , N^K is stochastically dominated by the logistic birth-and-death process \bar{N}^K jumping from $n \in \mathbb{Z}_+/K$ to $n + 1/K$ at rate $(\gamma K + \sup_{i \in \{1,2,3\}} \beta_i)Kn$ and from n to $n - 1/K$ at rate $(\gamma K + \inf_{i \in \{1,2,3\}} \delta_i + mn)Kn$. Finally, the sequence of stochastic processes \bar{N}^K satisfies (i) and (ii), which gives the result (see

respectively Lemma 1 of Champagnat (2006) and the proof of Theorem 5.3 of Fournier and Méléard (2004)). \square

From now we assume the following hypotheses:

$$g(z) > 0 \quad \text{for all } z \in (\mathbb{R}_+)^3 \setminus \{(0, 0, 0)\}, \quad (\text{H1})$$

and a 3-rd-order moments conditions:

$$\text{there exists } C < \infty \quad \text{such that} \quad \sup_K \mathbb{E}((N_0^K)^3) \leq C. \quad (\text{H2})$$

In Section 4 we will consider only the symmetrical case where $\alpha_{ij} = \alpha_{ji}$ for all i, j and give some explicit sufficient conditions on the parameters α_{ij} so that (H1) is true.

The following proposition gives that $(Y_t^K, t \geq 0)$ is a fast variable that converges toward the deterministic value 0 as K goes to infinity.

Proposition 2 *Under (H1) and (H2), for all $s, t > 0$, $\sup_{t \leq u \leq t+s} \mathbb{E}((Y_u^K)^2) \rightarrow 0$ as K goes to infinity.*

Proof Let us fix $z = (z_1, z_2, z_3) \in (\mathbb{R}_+)^3$ and set $(n, x, y) = \phi(z)$ where ϕ is defined in Lemma 1. The infinitesimal generator L^K of the jump process Z^K applied to a measurable real-valued function f (see Equation (2.6)) is decomposed as follows in z :

$$\begin{aligned}
L^K f(z) &= \gamma K^2 y \left[f\left(z - \frac{e_1}{K}\right) - 2f\left(z - \frac{e_2}{K}\right) + f\left(z - \frac{e_3}{K}\right) \right] \\
&\quad + \gamma K^2 n(x)^2 \left[f\left(z + \frac{e_1}{K}\right) + f\left(z - \frac{e_1}{K}\right) - 2f(z) \right] \\
&\quad + \gamma K^2 2nx(1-x) \left[f\left(z + \frac{e_2}{K}\right) + f\left(z - \frac{e_2}{K}\right) - 2f(z) \right] \\
&\quad + \gamma K^2 n(1-x)^2 \left[f\left(z + \frac{e_3}{K}\right) + f\left(z - \frac{e_3}{K}\right) - 2f(z) \right] \\
&\quad + \beta_1 K n(x)^2 \left[f\left(z + \frac{e_1}{K}\right) - f(z) \right] + \beta_2 K 2nx(1-x) \left[f\left(z + \frac{e_2}{K}\right) - f(z) \right] \\
&\quad + \beta_3 K n(1-x)^2 \left[f\left(z + \frac{e_3}{K}\right) - f(z) \right] \\
&\quad + K \sum_{i \in \{1,2,3\}} \left(\delta_i + \sum_{j \in \{1,2,3\}} \alpha_{ji} z_j \right) z_i \left[f\left(z - \frac{e_i}{K}\right) - f(z) \right] \\
&\quad + K \sum_{i \in \{1,2,3\}} \left(\gamma K + \delta_i + \sum_{j \in \{1,2,3\}} \alpha_{ji} z_j \right)^- z_i \left[f\left(z - \frac{e_i}{K}\right) - f(z) \right]
\end{aligned} \tag{3.1}$$

where $(x)^- = \max(-x, 0)$. Now if $f = (\phi_3)^2$, then there exist functions g_1^K , g_2^K , and g_3^K and a constant C_1 such that for all $z \in (\mathbb{R}_+)^3$,

$$\begin{aligned}
f\left(z + \frac{e_1}{K}\right) - f(z) &= -\frac{2f(z)}{Kn} + \frac{2z_3 y}{Kn} + g_1^K(z) \\
f\left(z + \frac{e_2}{K}\right) - f(z) &= -\frac{2f(z)}{Kn} - \frac{z_2 y}{Kn} + g_2^K(z) \\
f\left(z + \frac{e_3}{K}\right) - f(z) &= -\frac{2f(z)}{Kn} + \frac{2z_1 y}{Kn} + g_3^K(z)
\end{aligned}$$

with $|g_i^K(z)| \leq \frac{C_1}{K^2}$ for all $i \in \{1, 2, 3\}$. Finally, note that since $\gamma > 0$, there exists a positive constant C_2 such that for all $i \in \{1, 2, 3\}$ and all $z \in (\mathbb{R}_+)^3$,

$$\begin{aligned}
\mathbf{1} \left\{ \left(\gamma K + \delta_i + \sum_{j \in \{1,2,3\}} \alpha_{ji} z_j \right)^- \neq 0 \right\} &= \mathbf{1} \left\{ \sum_{j \in \{1,2,3\}} \alpha_{ji} z_j \leq -\gamma K - \delta_i \right\} \\
&\leq \mathbf{1} \left\{ \exists j \in \{1,2,3\} : \alpha_{ji} < 0, \alpha_{ji} z_j \leq -\frac{\gamma}{3} K - \frac{\delta_i}{3} \right\} \leq \mathbf{1} \{ \phi_1(z) \geq C_2 K \}.
\end{aligned}$$

Therefore, there exists a positive constant C_3 such that

$$L^K (\phi_3)^2(z) \leq -2\gamma K (\phi_3)^2(z) + C_3 [(\phi_1(z))^2(\phi_1(z) + K \mathbf{1}_{\{\phi_1(z) \geq C_2 K\}}) + 1].$$

Now from Proposition 1 and Markov's inequality, under (H1) and (H2), there exists a constant C such that $\sup_K \sup_{t \geq 0} \mathbb{E} \left(C_3 \left[(N_t^K)^2 (N_t^K + K \mathbf{1}_{\{N_t^K \geq C_2 K\}}) + 1 \right] \right) \leq C$. Therefore from the Kolmogorov forward equation, since $0 \leq (Y_t^K)^2 \leq (N_t^K)^2$ for all t and from Proposition 1,

$$\frac{d\mathbb{E}((Y_t^K)^2)}{dt} \leq -2\gamma K \mathbb{E}((Y_t^K)^2) + C.$$

This gives for all $t \geq 0$, $\frac{d}{dt} (e^{2\gamma K t} \mathbb{E}((Y_t^K)^2)) \leq C e^{2\gamma K t}$. Then by integration,

$$\begin{aligned} \mathbb{E}((Y_t^K)^2) &\leq \mathbb{E}((Y_0^K)^2) e^{-2\gamma K t} + \frac{C}{2\gamma K} - \frac{C}{2\gamma K} e^{-2\gamma K t} \\ &\leq \mathbb{E}((N_0^K)^2) e^{-2\gamma K t} + \frac{C}{2\gamma K} - \frac{C}{2\gamma K} e^{-2\gamma K t}, \quad \text{which gives the result.} \end{aligned}$$

□

In particular, under (H1) and (H2) and for all $t > 0$, Y_t^K converges in L^2 to 0. We say that Y^K is a fast variable compared to the vector (N^K, X^K) whose behavior is now studied. Let us introduce the following notation for all $z = (z_1, z_2, z_3) \in (\mathbb{R}_+)^3$:

$$\psi_1(z) = 2z_1 + z_2 \quad \text{and} \quad \psi_2(z) = 2z_3 + z_2.$$

Note that $\psi_1(Z_t^K) = 2N_t^K X_t^K$ (resp. $\psi_2(Z_t^K) = 2N_t^K (1 - X_t^K)$) is the rescaled number of allele A (resp. a) in the rescaled population Z^K at time t . For any $K \in \{1, 2, \dots\}$, $(\psi_1(Z^K), \psi_2(Z^K))$ is a pure jump Markov process with trajectories in $\mathbb{D}(\mathbb{R}_+, (\mathbb{Z}_+)^2 / K)$ and for all $i \in \{1, 2\}$, the process $\psi_i(Z^K)$ admits the following semi-martingale decomposition: for all $t \geq 0$,

$$\psi_i(Z_t^K) = \psi_i(Z_0^K) + M_t^{i,K} + \int_0^t L^K \psi_i(Z_s^K) ds$$

where $M^K = (M^{1,K}, M^{2,K})$ is, under (H1) and (H2), a square integrable \mathbb{R}^2 -valued càdlàg martingale (from Proposition 1) and is such that for all $i, j \in \{1, 2\}$, the predictable quadratic variation is given for all

$t \geq 0$ by:

$$\langle M^{i,K}, M^{j,K} \rangle_t = \int_0^t L^K \psi_i \psi_j(Z_s^K) - \psi_i(Z_s^K) L^K \psi_j(Z_s^K) - \psi_j(Z_s^K) L^K \psi_i(Z_s^K) ds.$$

Using this decomposition we prove the

Theorem 1 *Under (H1) and (H2), if the sequence $\{(\psi_1(Z_0^K), \psi_2(Z_0^K))\}_{K \in \mathbb{N}^*}$ of random variables converges in law toward a random variable (N_0^A, N_0^a) as K goes to infinity, then for all $T > 0$, the sequence of stochastic processes $(\psi_1(Z^K), \psi_2(Z^K))$ converges in law in $\mathbb{D}([0, T], (\mathbb{R}_+)^2)$ as K goes to infinity, toward the diffusion process (N^A, N^a) starting from (N_0^A, N_0^a) and satisfying the following diffusion equation, where $B = (B^1, B^2)$ is a 2-dimensional Brownian motion:*

$$\begin{aligned} dN_t^A &= \frac{N_t^A}{N_t^A + N_t^a} \left[\left[\beta_1 - \delta_1 - \frac{\alpha_{11}(N_t^A)^2 + \alpha_{21}2N_t^A N_t^a + \alpha_{31}(N_t^a)^2}{2(N_t^A + N_t^a)} \right] N_t^A \right. \\ &\quad \left. + \left[\beta_2 - \delta_2 - \frac{\alpha_{12}(N_t^A)^2 + \alpha_{22}2N_t^A N_t^a + \alpha_{32}(N_t^a)^2}{2(N_t^A + N_t^a)} \right] N_t^a \right] dt \\ &\quad + \sqrt{\frac{4\gamma}{N_t^A + N_t^a}} N_t^A dB_t^1 + \sqrt{2\gamma \frac{N_t^A N_t^a}{N_t^A + N_t^a}} dB_t^2 \\ dN_t^a &= \frac{N_t^a}{N_t^A + N_t^a} \left[\left[\beta_3 - \delta_3 - \frac{\alpha_{33}(N_t^a)^2 + \alpha_{23}2N_t^A N_t^a + \alpha_{13}(N_t^A)^2}{2(N_t^A + N_t^a)} \right] N_t^a \right. \\ &\quad \left. + \left[\beta_2 - \delta_2 - \frac{\alpha_{32}(N_t^a)^2 + \alpha_{22}2N_t^A N_t^a + \alpha_{12}(N_t^A)^2}{2(N_t^A + N_t^a)} \right] N_t^A \right] dt \\ &\quad + \sqrt{\frac{4\gamma}{N_t^A + N_t^a}} N_t^a dB_t^1 - \sqrt{2\gamma \frac{N_t^A N_t^a}{N_t^A + N_t^a}} dB_t^2 \end{aligned} \tag{3.2}$$

In particular in the neutral case where $\beta_i = \beta$, $\delta_i = \delta$ and $\alpha_{ij} = \alpha$ for all i, j ,

$$\begin{aligned} dN_t^A &= N_t^A \left(\beta - \delta - \alpha \frac{N_t^A + N_t^a}{2} \right) dt + \sqrt{\frac{4\gamma}{N_t^A + N_t^a}} N_t^A dB_t^1 + \sqrt{2\gamma \frac{N_t^A N_t^a}{N_t^A + N_t^a}} dB_t^2 \\ dN_t^a &= N_t^a \left(\beta - \delta - \alpha \frac{N_t^A + N_t^a}{2} \right) dt + \sqrt{\frac{4\gamma}{N_t^A + N_t^a}} N_t^a dB_t^1 - \sqrt{2\gamma \frac{N_t^A N_t^a}{N_t^A + N_t^a}} dB_t^2 \end{aligned} \tag{3.3}$$

Note that the diffusion coefficients of the diffusion process $((N_t^A, N_t^a), t \geq 0)$ do not explode as $N_t^A + N_t^a$ goes to 0 since $\frac{N_t^A}{\sqrt{N_t^A + N_t^a}} \leq \sqrt{N_t^A + N_t^a}$, $\frac{N_t^a}{\sqrt{N_t^A + N_t^a}} \leq \sqrt{N_t^A + N_t^a}$ and $\frac{N_t^A N_t^a}{N_t^A + N_t^a} \leq N_t^A + N_t^a$. From this

theorem we deduce the convergence of the sequence of processes

$$(N^K, X^K) = \left(\frac{\psi_1(Z^K) + \psi_2(Z^K)}{2}, \frac{\psi_1(Z^K)}{\psi_1(Z^K) + \psi_2(Z^K)} \right)$$

stopped when $N^K \leq \epsilon$ for any $\epsilon > 0$:

Corollary 1 For any $\epsilon > 0$ and $T > 0$, let us define $T_\epsilon^K = \inf\{t \in [0, T] : N_t^K \leq \epsilon\}$. If the sequence of random variables $(N_0^K, X_0^K) \in [\epsilon, +\infty[\times [0, 1]$ converges in law toward a random variable $(N_0, X_0) \in [\epsilon, +\infty[\times [0, 1]$ as K goes to infinity, then the sequence of stopped stochastic processes $\{(N^K, X^K)_{\cdot \wedge T_\epsilon^K}\}_{K \geq 1}$ converges in law in $\mathbb{D}([0, T], [\epsilon, \infty[\times [0, 1])$ as K goes to infinity, toward the stopped diffusion process $(N, X)_{\cdot \wedge T_\epsilon}$ ($T_\epsilon = \inf\{t \in [0, T] : N_t = \epsilon\}$), starting from (N_0, X_0) and satisfying the following diffusion equation:

$$\begin{aligned} dN_t &= \sqrt{2\gamma N_t} dB_t^1 \\ &+ N_t [X_t^2 (\beta_1 - \delta_1 - (\alpha_{11} N_t X_t^2 + \alpha_{21} 2N_t X_t (1 - X_t) + \alpha_{31} N_t (1 - X_t)^2)) \\ &\quad + 2X_t (1 - X_t) (\beta_2 - \delta_2 - (\alpha_{12} N_t X_t^2 + \alpha_{22} 2N_t X_t (1 - X_t) + \alpha_{32} N_t (1 - X_t)^2)) \\ &\quad + (1 - X_t)^2 (\beta_3 - \delta_3 - (\alpha_{13} N_t X_t^2 + \alpha_{23} 2N_t X_t (1 - X_t) + \alpha_{33} N_t (1 - X_t)^2))] dt \\ dX_t &= \sqrt{\frac{\gamma X_t (1 - X_t)}{N_t}} dB_t^2 \\ &+ (1 - X_t) X_t^2 [(\beta_1 - \delta_1) - (\beta_2 - \delta_2) \\ &\quad - N_t ((\alpha_{11} - \alpha_{12}) X_t^2 + (\alpha_{21} - \alpha_{22}) 2X_t (1 - X_t) + (\alpha_{31} - \alpha_{32}) (1 - X_t)^2)] dt \\ &+ X_t (1 - X_t)^2 [(\beta_2 - \delta_2) - (\beta_3 - \delta_3) \\ &\quad - N_t ((\alpha_{12} - \alpha_{13}) X_t^2 + (\alpha_{22} - \alpha_{23}) 2X_t (1 - X_t) + (\alpha_{32} - \alpha_{33}) (1 - X_t)^2)] dt. \end{aligned} \tag{3.4}$$

The population size and the proportion of allele A are therefore driven by two independent Brownian motions. The diffusion equation (3.4) can be simplified in the neutral case:

Remark 1 In the neutral case where $\beta_i = \beta$, $\delta_i = \delta$ and $\alpha_{ij} = \alpha$ for all i, j , the limiting diffusion (N, X) introduced in Equation (3.4) satisfies:

$$\begin{aligned} dN_t &= \sqrt{2\gamma N_t} dB_t^1 + N_t(\beta - \delta - \alpha N_t)dt \\ dX_t &= \sqrt{\frac{\gamma X_t(1-X_t)}{N_t}} dB_t^2. \end{aligned} \tag{3.5}$$

X is then a bounded martingale and this diffusion can be seen as a Wright-Fisher diffusion (see for instance Ethier and Kurtz (1986) p. 411) associated to a population size evolving stochastically with time. This diffusion will later be compared (next remark and Remark 3) to the one studied in Cattiaux and Méléard (2010) for a haploid population.

Remark 2 Depperschmidt et al. (2012) noted (Remark 2.1) that diploid additive selection and haploid selection are equivalent for populations with constant size. In our model, in an additive case for which $\beta_2 - \delta_2 = (\beta_1 - \delta_1) + s$, $\beta_3 - \delta_3 = (\beta_1 - \delta_1) + 2s$ and $\alpha_{ij} = \alpha$ for all i, j , the limiting diffusion (N, X) given in Equation (3.4) satisfies:

$$\begin{aligned} dN_t &= \sqrt{2\gamma N_t} dB_t^1 + N_t(\beta_1 - \delta_1 - \alpha N_t)dt + 2sN_t(1 - X_t)dt \\ dX_t &= \sqrt{\frac{\gamma X_t(1 - X_t)}{N_t}} dB_t^2 - sX_t(1 - X_t)dt. \end{aligned}$$

In this case of diploid additive selection, the drift coefficient of the proportion X_t (equal to $sX_t(1 - X_t)$) has indeed the same form than the drift of a haploid Wright-Fisher diffusion with selection. However, as we will see in the end of this section, due to population size stochasticity, our model of diploid populations with additive selection is different than haploid populations with stochastic population size studied in Cattiaux and Méléard (2010).

We denote by $\mathcal{C}_b^k(E, \mathbb{R})$ the set of functions from E to \mathbb{R} possessing bounded continuous derivatives of order up to k and $\mathcal{C}_c^k(E, \mathbb{R})$ the set of functions of $\mathcal{C}_b^k(E, \mathbb{R})$ with compact support.

Proof (of Theorem 1) Using the Rebolledo-Aldous criterion (Joffe and Métivier (1986)), we prove the tightness of the sequence of processes $(\psi_1(Z^K), \psi_2(Z^K))$ and its convergence toward the unique continuous solution of a martingale problem. The proof is divided in several steps.

STEP 1. Let us denote by L the generator of the diffusion process defined in Equation (3.2). We first prove the uniqueness of a solution $((N_t^A, N_t^a), t \in [0, T])$ to the martingale problem: for any function $f \in \mathcal{C}_b^2((\mathbb{R}_+)^2, \mathbb{R})$,

$$M_t^f = f(N_t^A, N_t^a) - f(N_0^A, N_0^a) - \int_0^t Lf(N_s^A, N_s^a) ds \quad (3.6)$$

is a continuous martingale. From Theorem 6.6.1 of Stroock and Varadhan (1979), for any $\epsilon > 0$, there exists a unique (in law) solution $((N_t^{A,\epsilon}, N_t^{a,\epsilon}), t \in [0, T])$ such that for all $f \in \mathcal{C}_b^2(\mathbb{R}^2)$ the process $(M_t^{f,\epsilon}, t \in [0, T])$ such that for all $t \geq 0$,

$$M_t^{f,\epsilon} = f(N_t^{A,\epsilon}, N_t^{a,\epsilon}) - f(N_0^{A,\epsilon}, N_0^{a,\epsilon}) - \int_0^t Lf(N_s^{A,\epsilon}, N_s^{a,\epsilon}) \mathbf{1}_{\{\epsilon < N_s^{A,\epsilon} + N_s^{a,\epsilon} < 1/\epsilon\}} ds$$

is a continuous martingale. The uniqueness of a solution of (3.6) therefore follows from Theorem 6.2 of Ethier and Kurtz (1986) about localization of martingale problems.

STEP 2. As in the proof of Proposition 2, we obtain easily that there exist two positive constants C_1 and C_2 such that for all $z \in (\mathbb{R}_+)^3$, the generator L^K of Z^K , decomposed in Equation (3.1), satisfies:

$$|L^K \psi_1(z)| + |L^K \psi_2(z)| \leq C_2 [\phi_1(z)^2 + 1 + K \phi_1(z) \mathbf{1}_{\phi_1(z) \geq C_1 K}],$$

and similarly

$$|L^K \psi_1^2(z) - 2\psi_1(z)L^K \psi_1(z) + L^K \psi_2^2(z) - 2\psi_2(z)L^K \psi_2(z)| \leq C_2(\phi_1(z)^2 + 1).$$

Therefore from Proposition 1, under (H1) and (H2), for any sequence of stopping times $\tau_K \leq T$ and for all $\epsilon > 0$:

$$\begin{aligned}
& \sup_{K \geq K_0} \sup_{\sigma \leq \delta} \mathbb{P} \left(\left| \int_{\tau_K}^{\tau_K + \sigma} L^K \psi_1(N_s^K, X_s^K) ds \right| + \left| \int_{\tau_K}^{\tau_K + \sigma} L^K \psi_2(N_s^K, X_s^K) ds \right| > \eta \right) \\
& \leq \sup_{K \geq K_0} \mathbb{P} \left(\delta \sup_{0 \leq s \leq T + \delta} C_2((N_s^K)^2 + 1) > \eta \right) + \sup_{K \geq K_0} \mathbb{P} \left(\sup_{0 \leq s \leq T + \delta} N_s^K \geq C_1 K \right) \quad (3.7) \\
& \leq \epsilon \quad \text{if } K_0 \text{ is large enough and } \delta \text{ is small enough.}
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \sup_{K \geq K_0} \sup_{\sigma \leq \delta} \mathbb{P} \left(\left| \int_{\tau_K}^{\tau_K + \sigma} (L^K \psi_1^2(Z_s^K) - 2\psi_1(Z_s^K)L^K \psi_1(Z_s^K) \right. \right. \\
& \quad \left. \left. + L^K \psi_2^2(Z_s^K) - 2\psi_2(Z_s^K)L^K \psi_2(Z_s^K)) ds \right| > \eta \right) \quad (3.8) \\
& \leq \epsilon \quad \text{if } K_0 \text{ is large enough and } \delta \text{ is small enough.}
\end{aligned}$$

The sequence of processes $(\psi_1(Z^K), \psi_2(Z^K))$ is then tight from the Rebolledo-Aldous criterion (Theorem 2.3.2 of Joffe and Métivier (1986)).

STEP 3. Let us consider a subsequence of $(\psi_1(Z^K), \psi_2(Z^K))$ that converges in law in $\mathbb{D}([0, T], \mathbb{R}^2)$ toward a process (N^A, N^a) . Since for all $K > 0$, $\sup_{t \in [0, T]} \|(N_t^A, N_t^a) - (N_{t-}^A, N_{t-}^a)\| \leq 2/K$ by construction, almost all trajectories of the limiting process (N^A, N^a) belong to $C([0, T], \mathbb{R}^2)$.

STEP 4. Finally we prove that the sequence $\{(\psi_1(Z^K), \psi_2(Z^K))\}_{K \in \{1, 2, \dots\}}$ of stochastic processes converges in law toward the unique continuous solution of the martingale problem given by Equation (3.6). Indeed for every function $f \in \mathcal{C}_c^3(\mathbb{R}^2)$, from Equation (3.1) there exists a constant C_4 such that

$$\begin{aligned}
& |L^K f(\psi_1(z), \psi_2(z)) - Lf(\psi_1(z), \psi_2(z))| \\
& \leq C_4 \left[\frac{\phi_1(z)^2}{K} + |\phi_3(z)|(1 + \phi_1(z)) + \gamma K \phi_1(z) \mathbf{1}_{\phi_1(z) \geq C_2 K} + (\phi_1(z)^2 + 1) \mathbf{1}_{\phi_1(z) \geq C_2 K} \right] \quad (3.9)
\end{aligned}$$

Note here that the fast-scale property shown in Proposition 2, combined with Proposition 1, will ensure that $\sup_{t \leq u \leq t+s} \mathbb{E}(|\phi_3(Z_t^K)| | \phi_1(Z_t^K))$ converges to 0 as K goes to infinity. Then for all $0 \leq t_1 < t_2 < \dots < t_k \leq t < t + s$, for all bounded continuous functions h_1, \dots, h_k on $(\mathbb{R}_+)^2$ and every $f \in \mathcal{C}_c^3(\mathbb{R}^2)$:

$$\begin{aligned}
& \mathbb{E} \left[\left(f(\psi_1(Z_{t+s}^K), \psi_2(Z_{t+s}^K)) - f(\psi_1(Z_t^K), \psi_2(Z_t^K)) - \int_t^{t+s} Lf(\psi_1(Z_u^K), \psi_2(Z_u^K)) du \right) \right. \\
& \quad \left. \times \prod_{i=1}^k h_i(\psi_1(Z_{t_i}^K), \psi_2(Z_{t_i}^K)) \right] = \\
& \mathbb{E} \left[\int_t^{t+s} (L^K f(\psi_1(Z_u^K), \psi_2(Z_u^K)) - Lf(\psi_1(Z_u^K), \psi_2(Z_u^K))) du \times \prod_{i=1}^k h_i(\psi_1(Z_{t_i}^K), \psi_2(Z_{t_i}^K)) \right] \\
& \leq \sup_i \|h_i\|_\infty \mathbb{E} \left[\int_t^{t+s} |L^K f(\psi_1(Z_u^K), \psi_2(Z_u^K)) - Lf(\psi_1(Z_u^K), \psi_2(Z_u^K))| du \right] \\
& \leq \sup_i \|h_i\|_\infty s \sup_{t \leq u \leq t+s} \mathbb{E} [|L^K f(\psi_1(Z_u^K), \psi_2(Z_u^K)) - Lf(\psi_1(Z_u^K), \psi_2(Z_u^K))|] \\
& \xrightarrow{K \rightarrow \infty} 0,
\end{aligned}$$

under (H1) and (H2), from Equation (3.9) and Propositions 1 and 2. The extension of this result to any $f \in \mathcal{C}_b^2((\mathbb{R}_+)^2, \mathbb{R})$ is easy to obtain by approximating uniformly f by a sequence of functions $f_n \in \mathcal{C}_c^3((\mathbb{R}_+)^2, \mathbb{R})$. Then from Theorem 8.10 (p. 234) of Ethier and Kurtz (1986), $(\psi_1(Z^K), \psi_2(Z^K))$ converges in law in $\mathbb{D}([0, T], \mathbb{R}^2)$ toward the unique (in law) solution of the martingale problem given in Equation (3.6), which is equal in law to the diffusion process (N^A, N^a) of Equation (3.2).

□

The proof of Corollary 1 relies on the following analytic lemma:

Lemma 2 *For any $x = (x_t^1, x_t^2)_{0 \leq t \leq T} \in \mathbb{D}([0, T], (\mathbb{R}_+)^2)$ and any $\epsilon > 0$, let us define*

$$\zeta_\epsilon(x) = \inf\{t \in [0, T] : x_t^1 + x_t^2 \leq 2\epsilon\}.$$

Let $x = (x_t^1, x_t^2)_{0 \leq t \leq T} \in \mathcal{C}([0, T], (\mathbb{R}_+)^2)$ such that $x_0^1 + x_0^2 > 2\epsilon$ and $\epsilon' \mapsto \zeta_{\epsilon'}(x)$ is continuous in ϵ . Consider a sequence of functions $(x_n)_{n \in \mathbb{Z}_+}$ such that for any $n \in \mathbb{Z}_+$, $x_n = (x_t^{1,n}, x_t^{2,n})_{0 \leq t \leq T} \in \mathbb{D}([0, T], (\mathbb{R}_+)^2)$ and x_n converges to x for the Skorohod topology. Then the sequence of processes $((x_{t \wedge \zeta_\epsilon(x_n)}^{1,n}, x_{t \wedge \zeta_\epsilon(x_n)}^{2,n}), t \in [0, T])$ converges to $((x_{t \wedge \zeta_\epsilon(x)}^1, x_{t \wedge \zeta_\epsilon(x)}^2), t \in [0, T])$ as n goes to infinity.

Proof We first prove that $\zeta_\epsilon(x_n)$ converges to $\zeta_\epsilon(x)$ as n goes to infinity. For any $\delta > 0$, since $\epsilon' \mapsto \zeta_{\epsilon'}(x)$ is continuous in ϵ' , there exists $n' \in \mathbb{Z}_+^*$ such that $\zeta_{\epsilon-1/n'}(x) - \delta < \zeta_\epsilon(x) < \zeta_{\epsilon+1/n'}(x) + \delta$. Now let us assume that $\zeta_\epsilon(x_n)$ does not converge to $\zeta_\epsilon(x)$ as n goes to infinity. Then there exists

δ such that for all n there exists $k_n > n$ such that $|\zeta_\epsilon(x_{k_n}) - \zeta_\epsilon(x)| > \delta$. Consequently there exists m such that

$$\overline{\lim}_{n \rightarrow +\infty} \sup_{0 \leq t \leq \zeta_{\epsilon^{-1}/m}(x)} |x_n^1(t) + x_n^2(t) - (x^1(t) + x^2(t))| \geq 1/m$$

which is impossible if x is continuous. Now we prove that $(x_n)_{\cdot \wedge \zeta_\epsilon(x_n)}$ converges to $x_{\cdot \wedge \zeta_\epsilon(x)}$ as n goes to infinity. Let us denote by $r(v, w)$ the Euclidean distance between two points v and w of \mathbb{R}^2 . Since x_n converges to x in $\mathbb{D}([0, T], (\mathbb{R}_+)^2)$, there exists a sequence of strictly increasing functions λ_n mapping $[0, \infty)$ onto $[0, \infty)$ such that

$$\gamma(\lambda_n) \xrightarrow{n \rightarrow +\infty} 0 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} r(x_n(t), x(\lambda_n(t))) = 0 \quad (3.10)$$

where $\gamma(\lambda) = \sup_{0 \leq t < s} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right|$ (Ethier and Kurtz (1986), p. 117). Now for all $t \geq 0$,

$$\begin{aligned} r(x_n(t \wedge \zeta_\epsilon(x_n)), x(\lambda_n(t) \wedge \zeta_\epsilon(x))) &\leq r(x_n(t \wedge \zeta_\epsilon(x_n)), x(\lambda_n(t \wedge \zeta_\epsilon(x_n)))) \\ &\quad + r(x(\lambda_n(t \wedge \zeta_\epsilon(x_n))), x(\lambda_n(t) \wedge \zeta_\epsilon(x))), \quad \text{and} \end{aligned}$$

$$\begin{aligned} r(x(\lambda_n(t \wedge \zeta_\epsilon(x_n))), x(\lambda_n(t) \wedge \zeta_\epsilon(x))) &= r(x(\lambda_n(\zeta_\epsilon(x_n))), x(\zeta_\epsilon(x))) \mathbf{1}_{\{t > \zeta_\epsilon(x_n), \lambda_n(t) > \zeta_\epsilon(x)\}} \\ &\quad + r(x(\zeta_\epsilon(x)), x(\lambda_n(t))) \mathbf{1}_{\{t \leq \zeta_\epsilon(x_n), \lambda_n(t) > \zeta_\epsilon(x)\}} \\ &\quad + r(x(\lambda_n(\zeta_\epsilon(x_n))), x(\lambda_n(t))) \mathbf{1}_{\{t > \zeta_\epsilon(x_n), \lambda_n(t) \leq \zeta_\epsilon(x)\}}. \end{aligned}$$

Therefore, using that x is continuous, that $\zeta_\epsilon(x_n) \rightarrow \zeta_\epsilon(x)$ and that $\sup_{0 \leq t \leq T} |\lambda_n(t) - t| \rightarrow 0$ as n goes to infinity, and from Equation (3.10), we obtain that $\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} r(x_n(t \wedge \zeta_\epsilon(x_n)), x(\lambda_n(t) \wedge \zeta_\epsilon(x))) = 0$ which gives the result. \square

Proof (of Corollary 1) The function ζ_ϵ defined in Lemma 2 satisfies $T_\epsilon^K = \zeta_\epsilon(\psi_1(Z^K), \psi_2(Z^K)) = \inf\{t \in [0, T] : N_t^K \leq \epsilon\}$, and $T_\epsilon = \zeta_\epsilon(N^A, N^a) = \inf\{t \in [0, T] : N_t \leq \epsilon\}$. From the Theorem 3.3 of Pinsky (1995), we know that the function $\epsilon' \mapsto \zeta_{\epsilon'}(N^A, N^a)$ is almost surely continuous in ϵ' . Therefore from Lemma 2, the function f such that for all $x \in \mathbb{D}([0, T], (\mathbb{R}_+)^2)$, $f(x) = (x_{t \wedge \zeta_\epsilon(x)}, t \in [0, T])$ is continuous in almost all trajectories of the diffusion process (N^A, N^a) . Therefore from

Corollary 1.9 p.103 of Ethier and Kurtz (1986) and Theorem 1, if the sequence of random variables $(\psi_1(Z_0^K), \psi(Z_0^K)) \in (\mathbb{R}_+)^2$ converges in law toward a random variable (N_0^A, N_0^a) as K goes to infinity, then for all $T > 0$, the sequence of stochastic processes $(\psi_1(Z_{\cdot \wedge T_\epsilon}^K), \psi(Z_{\cdot \wedge T_\epsilon}^K))$ converges in law in $\mathbb{D}([0, T], (\mathbb{R}_+)^2)$ toward $(N_{\cdot \wedge T_\epsilon}^A, N_{\cdot \wedge T_\epsilon}^a)$. Since the function $(n^A, n^a) \mapsto \left(\frac{n^A + n^a}{2}, \frac{n^A}{n^A + n^a}\right)$ is Lipschitz continuous on $\{(n^A, n^a) \in (\mathbb{R}_+)^2 : n^A + n^a \geq 2\epsilon\}$, we get the result. \square

Remark 3 The diffusion process $(N_t, X_t)_{t \geq 0}$ of Remark 1 can be compared to the haploid neutral population (which corresponds to a 2-type stochastic Lotka-Volterra competition model) studied in detail in Cattiaux and Méléard (2010) and defined by:

$$\begin{aligned} dH_t^1 &= \sqrt{2\gamma H_t^1} dB_t^{1,h} + (\beta - \delta - \alpha(H_t^1 + H_t^2))H_t^1 dt \\ dH_t^2 &= \sqrt{2\gamma H_t^2} dB_t^{2,h} + (\beta - \delta - \alpha(H_t^1 + H_t^2))H_t^2 dt \end{aligned} \quad (3.11)$$

where $(B^{1,h}, B^{2,h})$ is a 2-dimensional Brownian motion. Here, H^1 is the number of alleles A while H^2 is the number of alleles a . The quantity $N^h = H^1 + H^2$ is then the total number of individuals while $X^h = H^1/(H^1 + H^2)$ is the proportion of alleles A in the haploid population. We easily see that the total population size satisfies the same diffusion equation in the haploid and diploid populations. We therefore compare the stochastic processes (N, X) and (N^h, X^h) . Now by Itô's formula, the stochastic process (N^h, X^h) satisfies a diffusion equation that can be written using a new 2-dimensional Brownian motion $(\tilde{B}^1, \tilde{B}^2)$ as:

$$\begin{aligned} dN_t^h &= (\beta - \delta - \alpha N_t^h)N_t^h dt + \sqrt{2\gamma N_t^h} d\tilde{B}_t^1 \\ dX_t^h &= \sqrt{\frac{2\gamma X_t^h(1 - X_t^h)}{N_t^h}} d\tilde{B}_t^2 \end{aligned} \quad (3.12)$$

Then the differences between the haploid and the diploid neutral models only reside in a variation of the proportion of allele A which is divided by $\sqrt{2}$ in the diploid population (see Equations (3.5) and (3.12)). However, note from Equations (3.3) and (3.11) that this apparently insignificant difference induces that the respective numbers of alleles A and a are directed by correlated Brownian motions in a diploid population which is not the case in a haploid population. In particular, due to the

stochasticity in population size dynamics and since the diffusion term of the population size is not divided by $\sqrt{2}$ for a diploid population, our neutral diploid population does not have the same law as a neutral haploid population with sampling parameter γ replaced by $\gamma/2$.

4 New change of variable and quasi-stationarity

In this section we study the long-time behavior of the diffusion process (N^A, N^a) introduced in Theorem 1. For any process U , we denote by \mathbb{P}_x^U the law of U starting from a point x , and \mathbb{E}_x^U the associated expectation. First we show that the process $N = N^A + N^a$ defined in Corollary 1 reaches 0 almost surely in finite time:

Proposition 3 *Let $T_0 = \inf\{t \geq 0 : N_t = 0\}$. Under (H1), $\mathbb{P}_x^N(T_0 < +\infty) = 1$ for all $x \in \mathbb{R}_+$, and there exists $\lambda > 0$ such that $\sup_x \mathbb{E}_x(e^{\lambda T_0}) < +\infty$.*

Proof Under (H1), as in the proof of Proposition 1, there exists a positive constant m such that N is stochastically dominated by a diffusion process \bar{N} satisfying $d\bar{N}_t = \sqrt{2\gamma\bar{N}_t}dB_t^1 + \bar{N}_t(\sup_i \beta_i - \inf_i \delta_i - m\bar{N}_t)dt$. Theorem 5.2 of Cattiaux et al. (2009) gives the result for \bar{N} and therefore for N . \square

Due to the stochastic dynamics of population size, the long-time behavior of the diffusion (N^a, N^A) is therefore trivial and we now study the long-time behavior of this diffusion process conditioned on non-extinction, i.e. conditioned on not reaching the absorbing state $(0, 0)$. In particular, we are interested in studying the possibility of a long-time coexistence of the two alleles A and a in the population conditioned on non-extinction.

4.1 New change of variables

The quasi-stationary behavior of multi-dimensional diffusions was studied in Cattiaux et al. (2009) and Cattiaux and Méléard (2010) for Kolmogorov diffusions (i.e. diffusion processes with unit diffusion coefficient and gradient drift). To study the quasi-stationary behavior of the diffusion (N, X) conditioned

on non-extinction, we therefore change variables to obtain a 2-dimensional Kolmogorov diffusion (under conditions on the interactions parameters that will be discussed later). Let us define, as long as $N_t > 0$:

$$\begin{aligned} S_t^1 &= \sqrt{\frac{2N_t}{\gamma}} \cos\left(\frac{\arccos(2X_t - 1)}{\sqrt{2}}\right) \\ S_t^2 &= \sqrt{\frac{2N_t}{\gamma}} \sin\left(\frac{\arccos(2X_t - 1)}{\sqrt{2}}\right). \end{aligned} \quad (4.1)$$

If $N_t = 0$, we obviously set $S_t = (S_t^1, S_t^2) = (0, 0)$. To begin with, simple calculations give the following proposition, illustrated in Figure 1.

Proposition 4 *For all $t \geq 0$, one has $S_t^2 \geq 0$ and $S_t^2 \geq uS_t^1$, with $u = \tan\left(\frac{\pi}{\sqrt{2}}\right) < 0$.*

Proof For all $t \geq 0$, $2X_t - 1 \in [-1, 1]$, so $\frac{\arccos(2X_t - 1)}{\sqrt{2}} \in [0, \pi/\sqrt{2}]$ and $\sin\left(\frac{\arccos(2X_t - 1)}{\sqrt{2}}\right) > 0$. Then $S_t^2 \geq 0$ for all $t \geq 0$. Now if $\frac{\arccos(2X_t - 1)}{\sqrt{2}} \in [0, \pi/2]$, then $S_t^1 \geq 0$ and $S_t^2 \geq 0$, so $S_t^2 \geq uS_t^1$. Finally, if $\frac{\arccos(2X_t - 1)}{\sqrt{2}} \in]\pi/2, \pi/\sqrt{2}]$, then $S_t^1 < 0$, $S_t^2 \geq 0$, and $\frac{S_t^2}{S_t^1} = \tan\left(\frac{\arccos(2X_t - 1)}{\sqrt{2}}\right) \in]-\infty, u]$. Thus $S_t^2 \geq uS_t^1$. \square

Remark 4 Let us define for all $(s_1, s_2) \in \mathbb{R}^2$, the sets $\mathbf{A} = \{s_2 = 0, s_1 > 0\}$, $\mathbf{a} = \{s_2 = us_1, s_2 > 0\}$ and $\mathbf{0} = \{s_1 = s_2 = 0\}$. The events $\{S_t \in \mathbf{A}\}$, $\{S_t \in \mathbf{a}\}$, and $\{S_t \in \mathbf{0}\}$ are respectively equal to the events $\{X_t = 1\}$ (fixation of allele A), $\{X_t = 0\}$ (fixation of allele a) and $\{N_t = 0\}$ (extinction of the population).

We denote by $\mathcal{D} = \mathbf{R} \times \mathbf{R}_+ \cap \{(s_1, s_2) : s_2 \geq us_1\}$ the set of values taken by S_t for $t \geq 0$, $\partial\mathcal{D} = \mathbf{A} \cup \mathbf{a} \cup \mathbf{0}$ its boundary in \mathbb{R}^2 , and $T_{\mathbf{D}}$ the hitting time of \mathbf{D} for any $\mathbf{D} \subset \mathcal{D}$. $\mathbf{0}$, $\mathbf{A} \cup \mathbf{0}$ and $\mathbf{a} \cup \mathbf{0}$ are therefore absorbing sets and from Proposition 3, starting from any point $s \in \mathcal{D}$, S reaches any of these sets almost surely in finite time. Finally, we get the

Proposition 5 *The transformation*

$$\begin{aligned} \psi : \mathbb{R}_+^* \times [0, 1] &\rightarrow \mathcal{D} \setminus \mathbf{0} \\ (n, x) &\mapsto (s_1, s_2) = \left(\sqrt{\frac{2n}{\gamma}} \cos\left(\frac{\arccos(2x - 1)}{\sqrt{2}}\right), \sqrt{\frac{2n}{\gamma}} \sin\left(\frac{\arccos(2x - 1)}{\sqrt{2}}\right) \right) \end{aligned}$$

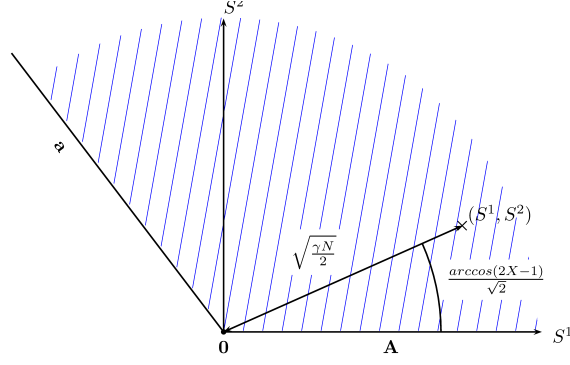


Fig. 1 Set \mathcal{D} of the possible values of S_t , for $t \geq 0$.

introduced in Equation (4.1) is a bijection.

See Figure 1 for an illustration.

Proof For any $(s_1, s_2) \in \mathcal{D} \setminus \mathbf{0}$, we easily get the following inverse transformation:

$$x = \begin{cases} \frac{1 + \cos\left(\sqrt{2} \arctan\left(\frac{s_2}{s_1}\right)\right)}{2} & \text{if } s_1 \geq 0, \\ \frac{1 + \cos\left(\sqrt{2} \arctan\left(\frac{s_2}{s_1} + \pi\right)\right)}{2} & \text{if } s_1 \leq 0, \end{cases} \quad \text{and} \quad n = \frac{((s_1)^2 + (s_2)^2)\gamma}{2},$$

for which we obviously have $n \in \mathbb{R}_+^*$ and $x \in [0, 1]$. □

Now from Itô's formula, S satisfies the following diffusion equation:

$$\begin{aligned} dS_t^1 &= dW_t^1 - q_1(S_t)dt \\ dS_t^2 &= dW_t^2 - q_2(S_t)dt, \end{aligned} \tag{4.2}$$

where, in the neutral case (Equation (2.5)), $q(s)$ is defined for all $s = (s_1, s_2) \in \mathcal{D}$ such that $s_1 \geq 0$ by

$$q(s) = \begin{pmatrix} -\frac{s_2}{(s_1)^2 + (s_2)^2} \frac{1}{\sqrt{2} \tan\left(\sqrt{2} \arctan\left(\frac{s_2}{s_1}\right)\right)} \\ -s_1 \left[\left(\beta - \delta - \frac{\alpha\gamma}{2}((s_1)^2 + (s_2)^2)\right) \frac{1}{2} - \frac{1}{(s_1)^2 + (s_2)^2} \right] \\ \frac{s_1}{(s_1)^2 + (s_2)^2} \frac{1}{\sqrt{2} \tan\left(\sqrt{2} \arctan\left(\frac{s_2}{s_1}\right)\right)} \\ -s_2 \left[\left(\beta - \delta - \frac{\alpha\gamma}{2}((s_1)^2 + (s_2)^2)\right) \frac{1}{2} - \frac{1}{(s_1)^2 + (s_2)^2} \right] \end{pmatrix} \tag{4.3}$$

and when $s_1 \leq 0$ by

$$q(s) = \begin{pmatrix} -\frac{s_2}{(s_1)^2+(s_2)^2} \frac{1}{\sqrt{2} \tan\left(\sqrt{2}\left(\arctan\left(\frac{s_2}{s_1}\right)+\pi\right)\right)} \\ -s_1 \left[\left(\beta - \delta - \frac{\alpha\gamma}{2}((s_1)^2 + (s_2)^2)\right) \frac{1}{2} - \frac{1}{(s_1)^2+(s_2)^2} \right] \\ \frac{s_1}{(s_1)^2+(s_2)^2} \frac{1}{\sqrt{2} \tan\left(\sqrt{2}\left(\arctan\left(\frac{s_2}{s_1}\right)+\pi\right)\right)} \\ -s_2 \left[\left(\beta - \delta - \frac{\alpha\gamma}{2}((s_1)^2 + (s_2)^2)\right) \frac{1}{2} - \frac{1}{(s_1)^2+(s_2)^2} \right] \end{pmatrix}.$$

The formula for q in the general case and if $s_1 \geq 0$ is given in Appendix A.2.

We now give conditions on the demographic parameters so that S is a Kolmogorov diffusion, i.e. satisfies $dS_t = dW_t - \nabla Q(S_t)dt$, for a real-valued function Q of two variables, called the potential of the diffusion process S . We need $q = (q_1, q_2) = \nabla Q$, which requires at least that $\frac{\partial q_2}{\partial s_1}(s) = \frac{\partial q_1}{\partial s_2}(s)$ for all $s \in \mathcal{D}$. We state the following

Proposition 6 (i) $\frac{\partial q_2(s)}{\partial s_1} = \frac{\partial q_1(s)}{\partial s_2}$ for all $s = (s_1, s_2) \in \mathcal{D}$ if and only α is symmetric, i.e.

$$\alpha_{12} = \alpha_{21}, \alpha_{31} = \alpha_{13}, \alpha_{23} = \alpha_{32}.$$

(ii) In this case we have

$$dS_t = dW_t - \nabla Q(S_t)dt, \quad (4.4)$$

with, in the neutral case and for all $s = (s_1, s_2) \in \mathcal{D}$,

$$Q(s) = \begin{cases} \frac{\ln((s_1)^2+(s_2)^2)}{2} + \frac{1}{2} \ln\left(\sin\left(\sqrt{2}\arctan\left(\frac{s_2}{s_1}\right)\right)\right) \\ \quad - \left(\beta - \delta - \frac{\alpha\gamma}{4}((s_1)^2 + (s_2)^2)\right) \frac{(s_1)^2+(s_2)^2}{4} \text{ if } s_1 \geq 0 \\ \frac{\ln((s_1)^2+(s_2)^2)}{2} + \frac{1}{2} \ln\left(\sin\left(\sqrt{2}\left(\arctan\left(\frac{s_2}{s_1}\right) + \pi\right)\right)\right) \\ \quad - \left(\beta - \delta - \frac{\alpha\gamma}{4}((s_1)^2 + (s_2)^2)\right) \frac{(s_1)^2+(s_2)^2}{4} \text{ if } s_1 \leq 0. \end{cases} \quad (4.5)$$

The corresponding potential Q in the non-neutral case is given in Appendix A.1.

Proof For (i), we can decompose the functions q_1 and q_2 as:

$$\begin{aligned}
q_1(s) &= \frac{\gamma}{2n} s_1 + \frac{\gamma}{\sqrt{2n}} s_2 \frac{2x-1}{4\sqrt{x(1-x)}} - \frac{s_1}{2} [x^2 U + 2x(1-x)V + (1-x)^2 W] \\
&\quad - \frac{s_2}{\sqrt{2}} \sqrt{x(1-x)} [x(U-V) + (1-x)(V-W)]
\end{aligned} \tag{4.6}$$

and

$$\begin{aligned}
q_2(s) &= \frac{\gamma}{2n} s_2 - \frac{\gamma}{\sqrt{2n}} s_1 \frac{2x-1}{4\sqrt{x(1-x)}} - \frac{s_2}{2} [x^2 U + 2x(1-x)V + (1-x)^2 W] \\
&\quad + \frac{s_1}{\sqrt{2}} \sqrt{x(1-x)} [x(U-V) + (1-x)(V-W)]
\end{aligned} \tag{4.7}$$

where

$$x = \begin{cases} \frac{1 + \cos(\sqrt{2} \arctan(\frac{s_2}{s_1}))}{2} & \text{if } s_1 \geq 0, \\ \frac{1 + \cos(\sqrt{2} \arctan(\frac{s_2}{s_1} + \pi))}{2} & \text{if } s_1 \leq 0, \end{cases} \quad n = \frac{((s_1)^2 + (s_2)^2)\gamma}{2}, \tag{4.8}$$

$$\begin{aligned}
U &= \beta_1 - \delta_1 - n(\alpha_{11}x^2 + \alpha_{21}2x(1-x) + \alpha_{31}(1-x)^2) \\
V &= \beta_2 - \delta_2 - n(\alpha_{12}x^2 + \alpha_{22}2x(1-x) + \alpha_{32}(1-x)^2), \text{ and} \\
W &= \beta_3 - \delta_3 - n(\alpha_{13}x^2 + \alpha_{23}2x(1-x) + \alpha_{33}(1-x)^2).
\end{aligned}$$

From Equation (4.8), we easily obtain that:

$$\frac{\partial n(s_1, s_2)}{\partial s_1} = \frac{s_1}{s_2} \frac{\partial n(s_1, s_2)}{\partial s_2} \quad \text{and} \quad \frac{\partial x(s_1, s_2)}{\partial s_1} = -\frac{s_2}{s_1} \frac{\partial x(s_1, s_2)}{\partial s_2}.$$

Finally, after some calculations and using that

$$\frac{\partial n}{\partial s_1} = \gamma s_1 \quad \text{and} \quad \frac{\partial x}{\partial s_1} \times \left(\frac{(s_1)^2 + (s_2)^2}{s_2} \right) = \sqrt{2x(1-x)},$$

we obtain that $\frac{\partial q_1(s)}{\partial s_2} = \frac{\partial q_2(s)}{\partial s_1}$ if and only if for all $x \in [0, 1]$,

$$x^2[\alpha_{21} - \alpha_{31} - \alpha_{12} + \alpha_{13} + \alpha_{32} - \alpha_{23}] + x[\alpha_{31} - \alpha_{13} + 2\alpha_{23} - 2\alpha_{32}] + [\alpha_{32} - \alpha_{23}] = 0$$

which happens if and only if α is symmetric. For (ii), the result comes from straightforward calculations that are given in the general case in Appendix A.1. \square

Assuming that α is symmetric, we will therefore be able, in Section 4.3, to describe the quasi-stationary behavior of the diffusion process S . This restriction of our model (which is necessary to use the existing theory on quasi-stationarity of diffusion processes developed in Cattiaux et al. (2009)) still allows the whole model to be asymmetrical (since $\beta_1 - \delta_1$ might be different than $\beta_3 - \delta_3$), and does not exclude some biologically interesting situations such as overdominance.

The following proposition gives some sufficient conditions on the parameters α_{ij} so that the function g introduced in Proposition 1 is positive, i.e. Hypothesis (H1) is satisfied.

Proposition 7 *Let us now assume that $\alpha_{ij} = \alpha_{ji}$ for all $i, j \in \{1, 2, 3\}$. If $\alpha_{ii} > 0$ for all $i \in \{1, 2, 3\}$ and one of the following conditions is satisfied:*

- (i) $\alpha_{ij} > 0$ for all i, j .
- (ii) *There exists $i \in \{1, 2, 3\}$ such that $\alpha_{ik} > 0$ for all k , and $\alpha_{jl}^2 < \alpha_{jj}\alpha_{ll}$ if i, j and l are all distinct.*
- (iii) *There exists $i \in \{1, 2, 3\}$ such that $\alpha_{ii}\alpha_{jl} > \alpha_{ij}\alpha_{il}$, $\alpha_{ij}^2 < \alpha_{ii}\alpha_{jj}$, and $\alpha_{il}^2 < \alpha_{ii}\alpha_{ll}$ where i, j and l are all distinct.*
- (iv) *There exists $i \in \{1, 2, 3\}$ such that $\alpha_{ij}^2 < \alpha_{ii}\alpha_{jj}$, $\alpha_{il}^2 < \alpha_{ii}\alpha_{ll}$, and $(\alpha_{ii}\alpha_{jl} - \alpha_{ij}\alpha_{il})^2 < (\alpha_{ii}\alpha_{ll} - \alpha_{il}^2)(\alpha_{ii}\alpha_{jj} - \alpha_{ij}^2)$ where i, j and l are all distinct.*

then Hypothesis (H1) is satisfied.

Proof Since α is symmetric, we have for all $z = (z_1, z_2, z_3) \in (\mathbb{R}_+)^3$:

$$g(Z) = \alpha_{11}(z_1)^2 + 2\alpha_{12}z_1z_2 + \alpha_{22}(z_2)^2 + 2\alpha_{23}z_2z_3 + 2\alpha_{13}z_1z_3 + \alpha_{33}(z_3)^2.$$

Considering g as a polynomial function of z_1 , we easily obtain that g is positive if (1) : the discriminant $\Delta_1(z_2, z_3) = (2\alpha_{12}z_2 + 2\alpha_{13}z_3)^2 - 4\alpha_{11}(\alpha_{22}(z_2)^2 + \alpha_{33}(z_3)^2 + 2\alpha_{23}z_2z_3)$ is negative or if (2) : $(2\alpha_{12}z_2 + 2\alpha_{13}z_3) > \sqrt{\Delta_1(z_2, z_3)}$. If $\alpha_{12} > 0$, $\alpha_{13} > 0$, and $\alpha_{23} > 0$ (case (i)) or $\alpha_{23}^2 < \alpha_{22}\alpha_{33}$ (case (i) or (ii)), then (2) is true for all $z \in (\mathbb{R}_+)^3$. If $\alpha_{11}\alpha_{22} > \alpha_{12}^2$, $\alpha_{11}\alpha_{33} > \alpha_{13}^2$, and $\alpha_{11}\alpha_{23} > \alpha_{12}\alpha_{13}$ or $(\alpha_{11}\alpha_{23} - \alpha_{12}\alpha_{13})^2 < (\alpha_{11}\alpha_{33} - \alpha_{13}^2)(\alpha_{11}\alpha_{22} - \alpha_{12}^2)$ (case (iii) or (iv)),

then (1) is true for all $z \in (\mathbb{R}_+)^3$, which gives the result, allowing in the end for permutations of indices 1, 2, and 3. \square

Note that these conditions mean that, for Hypothesis (H1) to be true, we need that cooperation is not too strong or is compensated in some way by competition.

4.2 Absorption of the diffusion process S

In this section, we establish more precise results concerning the absorption of the process S in the absorbing sets $\mathbf{0}$, $\mathbf{A} \cup \mathbf{0}$, $\mathbf{a} \cup \mathbf{0}$ and $\mathbf{A} \cup \mathbf{a} \cup \mathbf{0}$.

Theorem 2 (i) For all $s \in \mathcal{D} \setminus \mathbf{0}$, $\mathbb{P}_s^S(T_{\mathbf{A}} \wedge T_{\mathbf{a}} < T_{\mathbf{0}}) = 1$.

(ii) For all $s \in \mathcal{D} \setminus \partial\mathcal{D}$, $\mathbb{P}_s^S(T_{\mathbf{A}} < T_{\mathbf{0}}) > 0$, and $\mathbb{P}_s^S(T_{\mathbf{a}} < T_{\mathbf{0}}) > 0$.

Let us recall that the diffusion process $(S_t, t \geq 0)$ is obtained by a change of variable from the diffusion process $((N_t, X_t), t \geq 0)$ such that in the neutral case:

$$\begin{aligned} dN_t &= \sqrt{2\gamma N_t} dB_t^1 + N_t(\beta - \delta - \alpha N_t)dt \\ dX_t &= \sqrt{\frac{\gamma X_t(1-X_t)}{N_t}} dB_t^2, \end{aligned} \tag{4.9}$$

where (B^1, B^2) is a 2-dimensional Brownian motion. If we denote by T_0 and $T_{\{0,1\}}^X$ the respective absorbing times of $(N_t, t \geq 0)$ and $(X_t, t \geq 0)$ (in 0 and $\{0,1\}$ respectively), the two results of Theorem 2 are then respectively equivalent to (i) $T^X < T_0$ a.s., and (ii) $\mathbb{P}_{(n,x)}^{(N,X)}(X_{T_{\{0,1\}}^X} = 1) > 0$ and $\mathbb{P}_{(n,x)}^{(N,X)}(X_{T_{\{0,1\}}^X} = 1) > 0$ for all $(n, x) \in \mathbb{R}_+^* \times]0, 1[$.

We start with the following

Lemma 3 The logistic diffusion process $(N_s, s \geq 0)$ defined by the first equation of (4.9) satisfies

$$\int_0^{T_0} \frac{1}{N_s} ds = +\infty \quad a.s.$$

Proof (of Lemma 3) We know from Foucart and Hénard (2013) (Proposition 3.2) that this result is true for the Feller diffusion $(Y_t, t \geq 0)$ which obeys $dY_t = \sqrt{2\gamma Y_t} dB_t^1$. Let us set $\tilde{N}_t = \sqrt{2N_t/\gamma}$ and $\tilde{Y}_t = \sqrt{2Y_t/\gamma}$ for all t . We have

$$\begin{aligned} d\tilde{N}_t &= dB_t^1 + \left[\frac{\tilde{N}_t}{2} \left(\beta - \delta - \alpha \frac{\gamma}{2} (\tilde{N}_t)^2 \right) - \frac{1}{2\tilde{N}_t} \right] dt \\ d\tilde{Y}_t &= dB_t^1 - \frac{1}{2\tilde{Y}_t} dt \end{aligned}$$

The drift of the process $(\tilde{N}_t)_{t \geq 0}$ is unbounded near 0. Still, applying Proposition 2.2 in Cattiaux et al. (2009), we obtain a local Girsanov formula. Indeed for all time $t > 0$, for all bounded Borel function f on $\mathcal{C}([0, t], \mathbb{R}_+)$ for all $x \in \mathbb{R}_+$,

$$\mathbb{E}_x^{\tilde{N}}(f(w)\mathbf{1}_{t < T_0(w)}) = \mathbb{E}_x^{\tilde{Y}}(f(w)\mathbf{1}_{t < T_0(w)}e^{A(t)})$$

with $A(t) = G(w_0) - G(w_t) - \frac{1}{2} \int_0^t [(G'(w_u))^2 - G''(w_u)] du$ and $G(x) = \frac{(\beta - \delta)x^2}{4} - \frac{\alpha\gamma x^4}{16}$ for all $x \in \mathbb{R}_+$. Note that $A(t \wedge T_0)$ is well defined, therefore $\mathbb{E}_x^{\tilde{N}}[f(w)] = \mathbb{E}_x^{\tilde{Y}}[f(w)e^{A(t \wedge T_0(w))}]$ for all $t > 0$ so the laws of the processes $(\tilde{N}_{t \wedge T_0}, t \geq 0)$ and $(\tilde{Y}_{t \wedge T_0}, t \geq 0)$ are equivalent, which gives the result. \square

Proof (of Theorem 2) We first consider the neutral case. To prove (i), we introduce the time change $(\tau(t), t \geq 0)$ such that for all $t \geq 0$ (from the previous lemma),

$$\int_0^{\tau(t)} \frac{1}{N_s} ds = t.$$

Let B^2 be the Brownian motion from the 2nd equation of the System of equations 4.9. Since $X_{\tau(t)} = B_{\int_0^{\tau(t)} \gamma \frac{X_s(1-X_s)}{N_s} ds}^2$ for all $t \geq 0$, setting $\hat{X}_t = X_{\tau(t)}$ for all $t \geq 0$ and using the change of variables $s = \tau(u)$, we obtain that $(\hat{X}_t, t \geq 0)$ is a Wright-Fisher diffusion, i.e.

$$d\hat{X}_t = \gamma \hat{X}_t(1 - \hat{X}_t) dB_t^2$$

for all $t \geq 0$. Therefore the diffusion process $(\hat{X}_t, t \geq 0)$ reaches $\{0, 1\}$ almost surely in a finite time denoted by T and the diffusion process $(X_t, t \geq 0)$ reaches $\{0, 1\}$ almost surely at time $T_{\{0,1\}}^X = \tau(T)$. Now if $\tau(T) \geq T_0$, then $T = \int_0^{\tau(T)} \frac{1}{N_s} ds \geq \int_0^{T_0} \frac{1}{N_s} ds$ which is impossible from Lemma 3, since $T < \infty$ a.s.. Therefore $T_{\{0,1\}}^X = \tau(T) < T_0$ which gives (i). Now for (ii), in the neutral case, from Corollary 1, the proportion X of allele A is a bounded martingale, which gives that for every $(n, x) \in \mathbb{R}_+^* \times]0, 1[$,

$$\mathbb{P}_{(n,x)}^{N,X}(T_{\mathbf{A}} > T_{\mathbf{a}}) = \mathbb{E}_{(n,x)}^{N,X}(X_{T_{\mathbf{A}} \wedge T_{\mathbf{a}}}) = x > 0.$$

In the non-neutral case, by the extended Girsanov approach again, for any time $t > 0$, for any bounded Borel function f on $\mathcal{C}([0, t], \mathbb{R}_+)$ and for any $x \in \mathcal{D}$,

$$\mathbb{E}_s^S(f(w)) = \mathbb{E}_s^{\tilde{S}}[f(w)e^{C(t \wedge T_{\mathbf{A}} \wedge T_{\mathbf{a}})}]$$

where \tilde{S} is a neutral process whose potential is denoted by \tilde{Q} , with $C(t) = \hat{Q}(w_0) - \hat{Q}(w_t) - \frac{1}{2} \int_0^t [|\nabla \hat{Q}(w_u)|^2 - \Delta \hat{Q}(w_u)] du$ and $\hat{Q}(s) = Q(s) - \tilde{Q}(s)$ for all $s \in \mathcal{D}$, which gives the results. \square

4.3 Quasi-stationary behavior of S

In Cattiaux and Méléard (2010), the study of quasi-stationary distributions has been developed for diffusion processes of the form (4.4). In particular, existence and uniqueness is given under some conditions on the potential Q . Let us prove that these conditions are satisfied in our case.

Proposition 8 (i) *There exists a finite constant C such that for all $s = (s_1, s_2) \in \mathcal{D}$,*

$$|\nabla Q(s)|^2 - \Delta Q(s) \geq C.$$

(ii) $\inf\{|\nabla Q(s)|^2 - \Delta Q(s), |s| \geq R, s \in \mathcal{D}\} \rightarrow +\infty$ when $R \rightarrow \infty$.

Proof Let us define $F(s) = |\nabla Q(s)|^2 - \Delta Q(s)$ for all $s \in \mathcal{D}$. In the neutral case, we find:

$$\begin{aligned}
F(s) &= \frac{1}{2((s_1)^2 + (s_2)^2) \tan^2 \left(\sqrt{2} \arctan \frac{s_2}{s_1} \right)} \\
&+ ((s_1)^2 + (s_2)^2) \left[\frac{\beta - \delta - \frac{\alpha\gamma}{2}((s_1)^2 + (s_2)^2)^2}{4} + \frac{1}{((s_1)^2 + (s_2)^2)^2} \right] \\
&- ((s_1)^2 + (s_2)^2) \frac{\alpha\gamma}{2} + \frac{1}{(s_1)^2 + (s_2)^2} \frac{1 + \tan^2 \left(\sqrt{2} \arctan \frac{s_2}{s_1} \right)}{\tan^2 \left(\sqrt{2} \arctan \frac{s_2}{s_1} \right)}
\end{aligned}$$

which is not smaller than C . We also have

$$F(s) \geq ((s_1)^2 + (s_2)^2) \left[\frac{(\beta - \delta - \frac{\alpha\gamma}{2}((s_1)^2 + (s_2)^2)^2)}{4} + \frac{1}{(s_1)^2 + (s_2)^2} - \frac{\alpha\gamma}{2} \right]$$

which gives (ii). The proof of the two points in the non-neutral case is given in Appendix A.3. \square

As in Cattiaux and Méléard (2010), the quasi-stationary behavior of S is first studied with respect to the absorbing set $\partial\mathcal{D}$ and then for the absorbing set $\mathbf{0}$ that corresponds to the extinction of the population.

Theorem 3 (i) *There exists a unique distribution ν on $\mathcal{D} \setminus \partial\mathcal{D}$ such that for all $E \subset \mathcal{D} \setminus \partial\mathcal{D}$ and all $t \geq 0$,*

$$\mathbb{P}_\nu^S(S_t \in E | T_{\partial\mathcal{D}} > t) = \nu(E).$$

What is more, this distribution is a Yaglom limit for S , i.e. for all $s \in \mathcal{D} \setminus \partial\mathcal{D}$,

$$\lim_{t \rightarrow \infty} \mathbb{P}_s^S(S_t \in E | T_{\partial\mathcal{D}} > t) = \nu(E).$$

(ii) *There exists a unique probability measure ν_0 on $\mathcal{D} \setminus \mathbf{0}$ such that for all $s \in \mathcal{D} \setminus \partial\mathcal{D}$ and for all $E \subset \mathcal{D} \setminus \mathbf{0}$,*

$$\lim_{t \rightarrow \infty} \mathbb{P}_s^S(S_t \in E | T_{\mathbf{0}} > t) = \nu_0(E). \quad (4.10)$$

Proof The set of assumptions (H) of Cattiaux and Méléard (2010) (p. 816 – 818) is satisfied from Propositions 3 and 8, which gives (i) from Cattiaux and Méléard (2010) (Proposition B.12). (ii) is obtained as in Cattiaux and Méléard (2010) by using Theorem 2 and decomposing:

$$\mathbb{P}_s(S_t \in E | T_{\mathbf{0}} > t) = \frac{\mathbb{P}_s(S_t \in E) \mathbb{P}_s(T_{\partial\mathcal{D}} > t)}{\mathbb{P}_s(T_{\partial\mathcal{D}} > t) \mathbb{P}_s(T_{\mathbf{0}} > t)}. \quad (4.11)$$

□

Note that the quasi-stationary behavior of the diffusion process $((N_t, X_t), t \geq 0)$ conditioned on non extinction is obtained easily since

$$\mathbb{P}_{(n,x)}^{N,X}((N_t, X_t) \in F | N_t > 0) = \mathbb{P}_s^S(S_t \in E | T_{\mathbf{0}} > t),$$

if $s = \psi(n, x)$ and $E = \psi(F)$ where ψ is defined in Proposition 5. Let us remind that we are interested in studying the possibility of a long-time coexistence of the two alleles A and a in the population conditioned on non-extinction. This means that we would like to approximate the quasi-stationary distribution ν_X such that

$$\nu_X(\cdot) := \lim_{t \rightarrow \infty} \mathbb{P}_{(n,x)}^{N,X}(X_t \in \cdot | N_t > 0) \quad (4.12)$$

and we are interested in knowing whether $\nu_X(]0, 1[) = 0$ or not. Indeed, if $\nu_X(]0, 1[) \neq 0$ we can observe a long-time coexistence of the two alleles in the population conditioned on non-extinction whereas if $\nu_X(]0, 1[) = 0$, no such coexistence is possible. Note that $\nu_X(]0, 1[) = \nu_0(\mathcal{D} \setminus \partial\mathcal{D})$. For a haploid population with clonal reproduction, Cattiaux and Méléard (2010) proved that in a pure competition case, i.e. when every individual competes with every other one, no coexistence of alleles is possible. However, in our diploid population, we face a different situation and believe that this result is not true anymore. Indeed, from Equation (4.11),

$$\mathbb{P}_s(S_t \in \mathcal{D} \setminus \partial\mathcal{D}) = \frac{\mathbb{P}_s(T_{\partial\mathcal{D}} > t)}{\mathbb{P}_s(T_{\mathbf{0}} > t)},$$

therefore the possibility of coexistence of the two alleles relies on the fact that the time spent by the population in $\mathcal{D} \setminus \partial\mathcal{D}$ is not negligible compared to the time spent in $\mathcal{D} \setminus \mathbf{0}$. In a diploid population, if the heterozygotes are favored compared to homozygous individuals (this situation is called overdominance), they can make the coexistence period last longer than the remaining lifetime of the population once one of the alleles has disappeared. Similarly, as in Cattiaux and Méléard (2010), cooperation can favor the long-time

coexistence of alleles in the population conditioned on non-extinction. These biological and mathematical intuitions are now supported by numerical results.

5 Numerical results

Numerical simulations of ν_X are obtained following the Fleming-Viot algorithm introduced in Burdzy et al. (1996), which has been extensively studied in Villemonais (2011) and Villemonais (2013). This approach consists in approximating the conditioned distribution $\mathbb{P}_{(n,x)}^{N,X}((N_t, X_t) \in \cdot | T_0 > t)$ by the empirical distribution of an interacting particle system. More precisely, we consider a large number k of particles, that all start from a given $(n, x) \in \mathbb{R}_+^* \times]0, 1[$ and evolve independently from each other according to the law of the diffusion process (N, X) defined by the diffusion equation (3.4), until one of them hits $N = 0$. At that time, the absorbed particle jumps to the position of one of the remaining $k - 1$ particles, chosen uniformly at random among them. Then the particles evolve independently according to the law of the diffusion process (N, X) until one of them reaches $N = 0$, and so on. Theorem 1 of Villemonais (2013) gives the convergence as k goes to infinity of the empirical distribution of the k particles at time t toward the conditioned distribution $\mathbb{P}_{(n,x)}^{N,X}((N_t, X_t) \in \cdot | T_0 > t)$. Here we present three biologically relevant examples. For each case, we set $k = 2000$ and plot the empirical distribution at a large enough time T of the 2000 proportions of allele A given by the respective positions of the 2000 particles, starting from $(n, x) = (10, 1/2)$. First, we consider a neutral competitive case, in which each individual is in competition with every other one, independently from their genotypes. Here, our simulation of the quasi-stationary distribution ν_X of the proportion X is a sum of two Dirac functions in 0 and 1 (Figure 2), i.e. alleles A and a do not coexist in a long time limit.

Second (Figure 3), we show an overdominance case: every individual competes equally with every other ones but heterozygous individuals are favored compared to homozygotes, as their reproduction rate is higher. In this case, our simulation of the quasi-stationary distribution ν_X charges only points of $]0, 1[$, i.e. alleles A and a coexist with probability 1 or close to 1. This behavior is specific to the Mendelian re-

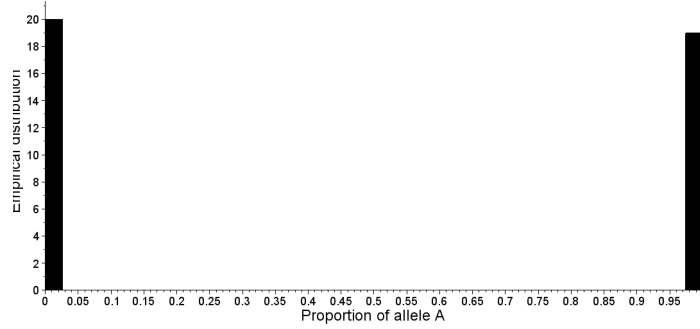


Fig. 2 Approximation of the quasi-stationary distribution ν_X of the proportion X of allele A (Equation (4.12)), in a neutral competitive case. In this figure, $\beta_i = 1$, $\delta_i = 0$, and $\alpha_{ij} = 0.1$ for all i, j , and $T = 40$.

production: in Cattiaux and Méléard (2010), the authors proved that no coexistence of alleles is possible in a haploid population with clonal reproduction, if every individual is in competition with every other one.

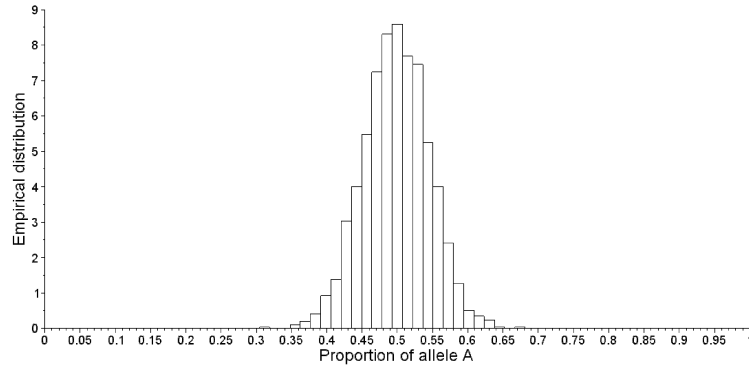


Fig. 3 Approximation of the quasi-stationary distribution ν_X of the proportion X of allele A (Equation (4.12)), in an overdominance case. In this figure, $\beta_i = 1$ for all $i \neq 2$, $\beta_2 = 5$, $\delta_i = 0$ for all i , $\alpha_{ij} = 0.1$ for all (i, j) , and $T = 500$.

Third (Figure 4), we show a case in which individuals only compete with individuals with same genotype; this can happen if different genotypes feed differently and/or have different predators. In this case, we can observe either a coexistence of the two alleles A and a or an elimination of one of the alleles, since the distribution ν_X charges both $\{0\} \cup \{1\}$ and $]0, 1[$.

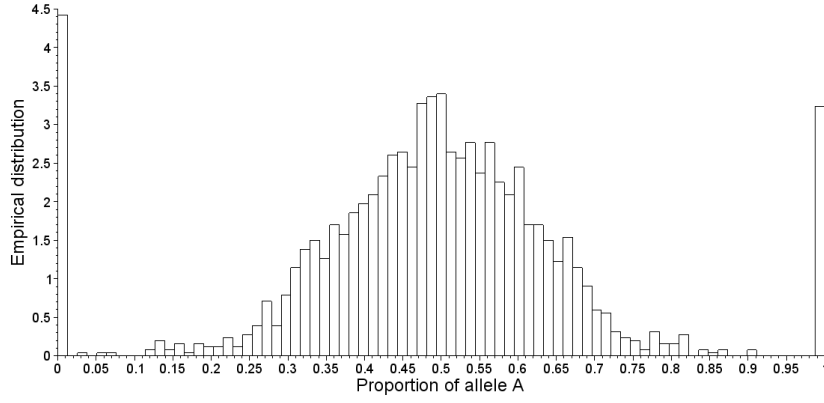


Fig. 4 Approximation of the quasi-stationary distribution ν_X of the proportion X of allele A (Equation (4.12)), in a case where individuals with different genotypes do not compete or cooperate with each other. In this figure, $\beta_i = 1$, $\delta_i = 0$, $\alpha_{ii} = 0.1$ for all i , $\alpha_{ij} = 0$ for all $i \neq j$, and $T = 2500$.

A Calculations in the general case

A.1 Form of the function Q

If α is symmetric, we use Equations (4.6), (4.7) and (4.8) and search a function Q such that $\frac{\partial Q(s)}{\partial s_1} = q_1(S)$ and $\frac{\partial Q(s)}{\partial s_2} = q_2(S)$.

After calculating the partial derivatives of functions of the form:

$$(s_1, s_2) \mapsto \begin{cases} ((s_1)^2 + (s_2)^2)^k \cos^l \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} \right) \right) & \text{if } s_1 \geq 0 \\ ((s_1)^2 + (s_2)^2)^k \cos^l \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} + \pi \right) \right) & \text{if } s_1 \leq 0 \end{cases}$$

for $k \in \{1, 2\}$ and $l \in \{1, 2, 3, 4\}$, we find that

$$Q(s) = \begin{cases} \frac{\ln((s_1)^2+(s_2)^2)}{2} + \frac{1}{2} \ln \left(\sin \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} \right) \right) \right) \\ - \frac{(s_1)^2+(s_2)^2}{4} \left[\frac{\beta_1 - \delta_1 + 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{4} - \frac{(s_1)^2+(s_2)^2}{4} \gamma \frac{\alpha_{11} + 4\alpha_{12} + 2\alpha_{13} + 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{16} \right] \\ - ((s_1)^2 + (s_2)^2) \left[h(s) \frac{\beta_1 - \delta_1 - (\beta_3 - \delta_3)}{8} + h(s)^2 \frac{\beta_1 - \delta_1 - 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{16} \right] \\ + \frac{((s_1)^2+(s_2)^2)^2}{16} \gamma h(s) \left[\frac{\alpha_{11} + 2\alpha_{12} - 2\alpha_{23} - \alpha_{33}}{4} + h(s) \frac{3\alpha_{11} - 2\alpha_{13} - 4\alpha_{22} + 3\alpha_{33}}{8} \right. \\ \left. + h(s)^2 \frac{\alpha_{11} - 2\alpha_{12} + 2\alpha_{23} - \alpha_{33}}{4} + h(s)^3 \frac{\alpha_{11} - 4\alpha_{12} + 2\alpha_{13} - 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{16} \right] \quad \text{if } s_1 \geq 0 \\ \\ \frac{\ln((s_1)^2+(s_2)^2)}{2} + \frac{1}{2} \ln \left(\sin \left(\sqrt{2} \left(\arctan \left(\frac{s_2}{s_1} \right) + \pi \right) \right) \right) \\ - \frac{(s_1)^2+(s_2)^2}{4} \left[\frac{\beta_1 - \delta_1 + 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{4} - \frac{(s_1)^2+(s_2)^2}{4} \gamma \frac{\alpha_{11} + 4\alpha_{12} + 2\alpha_{13} + 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{16} \right] \\ - ((s_1)^2 + (s_2)^2) \left[h(s) \frac{\beta_1 - \delta_1 - (\beta_3 - \delta_3)}{8} + h(s)^2 \frac{\beta_1 - \delta_1 - 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{16} \right] \\ + \frac{((s_1)^2+(s_2)^2)^2}{16} \gamma h(s) \left[\frac{\alpha_{11} + 2\alpha_{12} - 2\alpha_{23} - \alpha_{33}}{4} + h(s) \frac{3\alpha_{11} - 2\alpha_{13} - 4\alpha_{22} + 3\alpha_{33}}{8} \right. \\ \left. + h(s)^2 \frac{\alpha_{11} - 2\alpha_{12} + 2\alpha_{23} - \alpha_{33}}{4} + h(s)^3 \frac{\alpha_{11} - 4\alpha_{12} + 2\alpha_{13} - 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{16} \right] \quad \text{if } s_1 \leq 0 \end{cases} \quad (\text{A.1})$$

where

$$h(s) = \begin{cases} \cos \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} \right) \right) & \text{when } s_1 \geq 0 \\ \cos \left(\sqrt{2} \left(\arctan \left(\frac{s_2}{s_1} \right) + \pi \right) \right) & \text{when } s_1 \leq 0. \end{cases}$$

A.2 Form of the function q

Therefore if $s_1 \geq 0$:

$$q_1(s) = \frac{s_1}{(s_1)^2 + (s_2)^2} - \frac{s_2}{(s_1)^2 + (s_2)^2} \frac{1}{\sqrt{2} \tan \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} \right) \right)} \\ - s_1 \left[\frac{\beta_1 - \delta_1 + 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{8} - \frac{(s_1)^2 + (s_2)^2}{4} \gamma \frac{\alpha_{11} + 4\alpha_{12} + 2\alpha_{13} + 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{16} \right] \\ - 2s_1 \left[h(s) \frac{\beta_1 - \delta_1 - (\beta_3 - \delta_3)}{8} + h(s)^2 \frac{\beta_1 - \delta_1 - 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{16} \right] \\ + s_1 \frac{((s_1)^2 + (s_2)^2)}{4} \gamma h(s) \left[\frac{\alpha_{11} + 2\alpha_{12} - 2\alpha_{23} - \alpha_{33}}{4} + h(s) \frac{3\alpha_{11} - 2\alpha_{13} - 4\alpha_{22} + 3\alpha_{33}}{8} \right. \\ \left. + h(s)^2 \frac{\alpha_{11} - 2\alpha_{12} + 2\alpha_{23} - \alpha_{33}}{4} + h(s)^3 \frac{\alpha_{11} - 4\alpha_{12} + 2\alpha_{13} - 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{16} \right] \\ - \sqrt{2} s_2 \sin \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} \right) \right) \left[\frac{\beta_1 - \delta_1 - (\beta_3 - \delta_3)}{8} + h(s) \frac{\beta_1 - \delta_1 - 2(\beta_2 - \delta_2) + \beta_3 - \delta_3}{8} \right] \\ + \frac{(s_1)^2 + (s_2)^2}{16} \gamma \sqrt{2} s_2 \sin \left(\sqrt{2} \arctan \left(\frac{s_2}{s_1} \right) \right) \left[\frac{\alpha_{11} + 2\alpha_{12} - 2\alpha_{23} - \alpha_{33}}{4} + h(s) \frac{3\alpha_{11} - 2\alpha_{13} - 4\alpha_{22} + 3\alpha_{33}}{4} \right. \\ \left. + h(s)^2 \frac{3(\alpha_{11} - 2\alpha_{12} + \alpha_{23} - \alpha_{33})}{4} + h(s)^3 \frac{\alpha_{11} - 4\alpha_{12} + 2\alpha_{13} - 4\alpha_{23} + 4\alpha_{22} + \alpha_{33}}{4} \right]. \quad (\text{A.2})$$

We have similar formulas for q_2 and when $s_1 \leq 0$.

A.3 Proof of Proposition 8

Recall that

$$F(s) = |\nabla Q(s)|^2 - \Delta Q(s) = (q_1(s))^2 + (q_2(s))^2 - \frac{\partial q_1}{\partial s_1}(s) - \frac{\partial q_2}{\partial s_2}(s).$$

Besides, note that under (H1), $\frac{\alpha_{11}+4\alpha_{12}+2\alpha_{13}+4\alpha_{23}+4\alpha_{22}+\alpha_{33}}{16} > 0$. Therefore using Equations (4.6) and (4.7) we easily obtain that there exists a positive constant C_1 such that $(q_1(s))^2 + (q_2(s))^2 \geq C_1((s_1)^2 + (s_2)^2)^3$. Finally, from Equation (A.2), we obtain after some calculations that there exists a positive constant C_2 such that $\frac{\partial q_1}{\partial s_1}(s) + \frac{\partial q_2}{\partial s_2}(s) \leq C_2((s_1)^2 + (s_2)^2)^2$. Therefore Proposition 8 is true if $s_1 \geq 0$. If $s_1 \leq 0$, the result is true as well by symmetry.

Acknowledgements I fully thank my Phd director Sylvie Méléard for suggesting me this research subject, and for her continual guidance during my work. I would also like to thank Denis Villemonais for his help for the simulation results. This article benefited from the support of the ANR MANEGE (ANR-09-BLAN-0215) and from the Chair "Modélisation Mathématique et Biodiversité" of Veolia Environnement - École Polytechnique - Museum National d'Histoire Naturelle - Fondation X.

References

- Ball, K., Kurtz, T. G., Popovic, L., and Rempala, G. (2006). Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, 16(4):1925–1961.
- Berglund, N. and Gentz, B. (2006). *Noise-induced phenomena in slow-fast dynamical systems*. Probability and its Applications (New York). Springer-Verlag London Ltd., London.
- Burdzy, K., Holyst, R., Ingeman, D., and March, P. (1996). Configurational transition in a fleming-viot-type model and probabilistic interpretation of laplacian eigenfunctions. *J. Phys. A*, 29(11):2633–2642.
- Cattiaux, P., Collet, P., Lambert, A., Martínez, S., Méléard, S., and San Martín, J. (2009). Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5):1926–1969.
- Cattiaux, P. and Méléard, S. (2010). Competitive or weak cooperative stochastic Lotka-Volterra systems conditioned on non-extinction. *J. Math. Biol.*, 60(6):797–829.
- Champagnat, N. (2006). A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Process. Appl.*, 116(8):1127–1160.
- Champagnat, N., Ferrière, R., and Méléard, S. (2006). Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theor. Popul. Biol.*, 69:297–321.
- Champagnat, N. and Méléard, S. (2007). Invasion and adaptive evolution for individual-based spatially structured populations. *J. Math. Biol.*, 55:147–188.

- Collet, P., Méléard, S., and Metz, J. A. (2013). A rigorous model study of the adaptive dynamics of Mendelian diploids. *J. Math. Biol.*, 67(3):569–607.
- Coron, C. (2014). Stochastic modeling of density-dependent diploid populations and extinction vortex. *Adv. in Appl. Probab.*, 46:446–477.
- Crow, J. F. and Kimura, M. (1970). *An introduction to population genetics theory*. Harper & Row Publishers, New York.
- Depperschmidt, A., Greven, A., and Pfaffelhuber, P. (2012). Tree-valued Fleming-Viot dynamics with mutation and selection. *Ann. Appl. Probab.*, 22(6):2560–2615.
- Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.
- Ethier, S. N. and Nagylaki, T. (1980). Diffusion approximations of markov chains with two time scales and applications to population genetics. *Adv. in App. Probab.*, 12(1):14–49.
- Ethier, S. N. and Nagylaki, T. (1988). Diffusion approximations of markov chains with two time scales and applications to population genetics, ii. *Adv. in App. Probab.*, 20(3):525–545.
- Foucart, C. and Hénard, O. (2013). Stable continuous-state branching processes with immigration and beta-fleming-viot processes with immigration. *Electron. J. Probab.*, 18:no. 23, 1–21.
- Fournier, N. and Méléard, S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.*, 14(4):1880–1919.
- Hofbauer, J., Schuster, P., and Sigmund, K. (1982). Game dynamics in mendelian populations. *Biol. Cybernetics*, 43(1):51–57.
- Hofbauer, J. and Sigmund, K. (1998). *Evolutionary games and population genetics*. Cambridge University Press, Cambridge, UK.
- Hoppensteadt, F. (1975). *Mathematical theories of populations: demographics, genetics and epidemics*. Regional conference series in applied mathematics. Society for Industrial and Applied Mathematics.
- Joffe, A. and Métivier, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.*, 18(1):20–65.
- Katzenberger, G. S. (1991). Solutions of a stochastic differential equation forced onto a manifold by a large drift. *Ann. Probab.*, 19(4):1587–1628.
- Kurtz, T. G. (1992). Averaging for martingale problems and stochastic approximation. In *Applied stochastic analysis (New Brunswick, NJ, 1991)*, volume 177 of *Lecture Notes in Control and Inform. Sci.*, pages 186–209. Springer, Berlin.
- Méléard, S. and Tran, V. C. (2012). Slow and fast scales for superprocess limits of age-structured populations. *Stochastic Process. Appl.*, 122(1):250–276.
- Méléard, S. and Villemonais, D. (2012). Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410.
- Nagylaki, T. (1992). *Introduction to theoretical population genetics*. Springer-Verlag, Berlin.
- Nagylaki, T. and Crow, J. (1974). Continuous selective models. *Theor. Pop. Biol.*, 5:257–283.
- Norton, H. (1928). Natural selection and mendelian variation. *Proc. Lond. Math. Soc.*, 2(1):1–45.

-
- Papanicolaou, G. C., Stroock, D., and Varadhan, S. R. S. (1977). Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference (Duke Univ., Durham, N.C., 1976), Paper No. 6*, pages ii+120 pp. Duke Univ. Math. Ser., Vol. III. Duke Univ., Durham, N.C.
- Pinsky, R. G. (1995). *Positive harmonic functions and diffusion*, volume 45 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge.
- Stroock, D. W. and Varadhan, S. R. S. (1979). *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin.
- Villemonais, D. (2011). Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.*, 16(61):1663–1692.
- Villemonais, D. (2013). General approximation method for the distribution of markov processes conditioned not to be killed. *ESAIM: Probability and Statistics*, eFirst.