



HAL
open science

An Eco-Evolutionary approach of Adaptation and Recombination in a large population of varying size

Charline Smadi

► **To cite this version:**

Charline Smadi. An Eco-Evolutionary approach of Adaptation and Recombination in a large population of varying size. 2014. hal-00948097v1

HAL Id: hal-00948097

<https://hal.science/hal-00948097v1>

Preprint submitted on 17 Feb 2014 (v1), last revised 24 Dec 2014 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AN ECO-EVOLUTIONARY APPROACH OF ADAPTATION AND RECOMBINATION IN A LARGE POPULATION OF VARYING SIZE

CHARLINE SMADI

ABSTRACT. We identify the genetic signature left by a selective sweep in the vicinity of an allele fixed by positive selection. We consider a sexually reproducing haploid population with non constant population size described by a birth and death process with density dependent competition between individuals. Parameter K scales the population size, and we study the limit behaviour for large K . We focus on two linked loci: one under selection and a neutral one. We distinguish soft selective sweep occurring after an environmental change, from hard selective sweep occurring after a new mutation. In the first case, we express the variations of neutral allele proportions as a function of the ecological parameters, recombination probability and solutions of a deterministic two-dimensional Lotka-Volterra system with competition. In the case of strong selective sweep, we find two recombination regimes. In the strong one, the recombination probability per reproductive event is large with respect to $1/\log K$ and the selective sweep does not modify the neutral proportions. In the weak regime, the recombination probability per reproductive event is of order $1/\log K$, and the natural selection on the mutant drives an increase in frequency of the neutral allele carried by the first mutant. We express this increase as an explicit function of the ecological parameters, recombination probability and K .

1. INTRODUCTION

There are at least two different ways of adaptation for a population: selection can either act on a new mutation (hard selective sweep), either on preexisting alleles that become advantageous after an environmental change (soft selective sweep). New mutations are sources of diversity, and hard selective sweep was until recently the only considered way of adaptation. Soft selective sweep allows a faster adaptation to novel environments, and its importance is growing in empirical and theoretical studies (Prezeworski, Coop and Wall [24], Barrett and Schluter [2]). These distinct selective sweeps entail different genetic signatures in the vicinity of the newly fixed allele, and the multiplication of genetic data available allows to detect these signatures in current populations as described by Peter and al in [23]. To do this in an effective way, it is necessary to identify accurately the signatures left by these two modes of adaptation.

In this work, we consider a sexual haploid population with non constant population size, modeled by a birth and death process with density dependent competition. The ability to survive and reproduce of each individual depends on its own genotype and on the population state. More precisely, each individual is characterized by some ecological parameters: birth rate, intrinsic death rate and competition kernel describing the competition with other individuals depending on their genotype. The differential reproductive success of individuals generated by their interactions entail progressive variations in the number of individuals carrying a given genotype. This process, called natural selection, is a key mechanism of evolution. Such eco-evolutionary approach has been introduced by Metz and coauthors in [22] and made rigorous in the seminal

UNIVERSITÉ PARIS-EST, CERMICS (ENPC), F-77455 MARNE LA VALLÉE, FRANCE AND CMAP UMR 7641, ÉCOLE POLYTECHNIQUE CNRS, ROUTE DE SACLAY, 91128 PALAISEAU CEDEX, FRANCE

E-mail address: charline.smadi@polytechnique.edu.

Date: February 17, 2014.

2010 Mathematics Subject Classification. 60J75; 60F99; 60K35; 92D15.

Key words and phrases. eco-evolution; birth and death process; coalescent; selective sweep; genetic hitch-hiking.

paper of Fournier and Méléard [17]. Then it has been developed by Champagnat [5], Champagnat and coauthors [6, 7, 8, 9], and Méléard and Tran [21] for the haploid asexual case and by Collet, Méléard and Metz [10] and Coron [11, 12] for the diploid sexual case. Finally, the recent work of Billiard and coauthors [4] studies the dynamics of a two loci model in an haploid asexual population. Following these works, we introduce a parameter K , linked to the concept of carrying capacity which scales the population size, and study the limit behavior for large K . But unlike them, we focus on two loci in a sexual haploid population and take into account recombinations: one locus is under selection and has two possible alleles A and a and the second one is neutral with allele b_1 or b_2 . When two individuals give birth, either a recombination occurs with probability r_K and the newborn inherits one allele from each parent, or he is the clone of one of his parents.

We first focus on soft selective sweep occurring after a change in the environment (new pathogen, environmental catastrophe, occupation of a new ecological niche,...). We assume that before the change the alleles A and a were neutral and represented both a positive fraction of the population, and that in the new environment the allele a becomes favorable and goes to fixation. We can divide the selective sweep in two periods: a first one where both the alleles A and a represent a non negligible proportion of the population and the population process is well approximated by the solution of a deterministic dynamical system, and a second one where A -individuals are near extinction, the deterministic approximation fails and the fluctuations of the A -population size become predominant. We give the asymptotic value of the final neutral allele proportions as a function of the ecological parameters, recombination probability r_K and solutions of a two-dimensional competitive Lotka-Volterra system.

We then focus on hard selective sweeps. We assume that a mutant a appears in a monomorphic A -population at ecological equilibrium. As stated by Champagnat in [5], the selective sweep is divided in three periods: during the first one, the resident population size stays near its equilibrium value, and the mutant population size grows until it reaches a non-negligible fraction of the total population size. The two other periods are the ones described for the soft selective sweep. Moreover, the time needed for the mutant a to invade the population is of order $\log K$. We prove that the distribution of neutral alleles at the end of the sweep has different shapes according to the order of the recombination probability per reproductive event r_K with respect to $1/\log K$. More precisely, we find two recombination regimes: a strong one where $r_K \log K$ is large, and a weak one where r_K is of order $1/\log K$. In both recombination regimes, we give the asymptotic value of the final neutral allele proportions as a function of the ecological parameters and recombination probability r_K . In the strong recombination regime, the frequent exchanges of neutral alleles between the A and a -populations yield an homogeneous neutral allele repartition in the two populations, and the sweep does not modify the neutral allele repartition. In the weak recombination regime, the frequency of the neutral allele carried by the first mutant increases because it is linked to the positively selected allele. This phenomenon, called genetic hitch-hiking by Maynard Smith and Haigh [26], has been studied by many authors. Maynard Smith and Haigh [26] and Stephan and coauthors [27] use deterministic models for the change in the frequency of the selected allele. Kaplan and coauthors [19] and Barton [3] present more precise models taking into account the randomness of the first and third periods of the mutant invasion. Finally, Durrett and Schweinsberg [14, 25], Etheridge and coauthors [15], and Leocard [20] describe the population process by a structured coalescent and finely study genealogies of neutral alleles during the sweep. Unlike our model, in all these works, the population size is constant and the individuals' "selective value" does not depend on the population state, but only on the individuals' genotype.

The structure of the paper is the following. In Section 2 we describe the model, review some results of Champagnat in [5] about the two-dimensional population process when we do not consider the neutral locus, and present the main results. Section 3 is devoted to a semimartingale

decomposition of neutral proportions, key tool in the different proofs. In Section 4 we prove our result for the soft sweep. Our proof relies on a comparison of the population process with a four dimensional dynamical system. In Section 5 we describe a coupling of the population process with two birth and death processes widely use in Sections 6 and 7. Section 6 is devoted to the proof for the strong regime of hard sweep. The proof for the weak recombination regime is presented in Section 7 and requires a fine study of the genealogies in a structured coalescent process during the first phase of the selective sweep. We use here some ideas developed in [25]. Finally in Appendix we state technical results.

Throughout the rest of the paper we use c to denote a positive finite constant. Its value can change from line to line but it is always independent of K and ε . The set $\mathbb{N} := \{1, 2, \dots\}$ will denote the set of positive integers.

2. MODEL AND MAIN RESULTS

We introduce the sets $\mathcal{A} = \{A, a\}$, $\mathcal{B} = \{b_1, b_2\}$, and $\mathcal{E} = \{A, a\} \times \{b_1, b_2\}$ to describe the genetic background of individuals. The state of the population will be given by the four dimensional Markov process $N^{(z,K)} = (N_{\alpha\beta}^{(z,K)}(t), (\alpha, \beta) \in \mathcal{E})_{t \geq 0}$ where $N_{\alpha\beta}^{(z,K)}(t)$ denotes the number of individuals with alleles (α, β) at time t when the carrying capacity is $K \in \mathbb{N}$ and the initial state is $\lfloor zK \rfloor$ with $z = (z_{\alpha\beta}, (\alpha, \beta) \in \mathcal{E}) \in \mathbb{R}_+^{\mathcal{E}}$. We recall that b_1 and b_2 are neutral, thus ecological parameters only depend on the allele, A or a , carried by the individuals at their first locus. There are the following:

- For $\alpha \in \mathcal{A}$, f_α and D_α denote the birth rate and the intrinsic death rate of an individual carrying allele α .
- For $(\alpha_1, \alpha_2) \in \mathcal{A}^2$, C_{α_1, α_2} represents the competitive pressure felt by an individual carrying allele α_1 from an individual carrying allele α_2 .
- $K \in \mathbb{N}$ is a parameter rescaling the competition between individuals. It can be interpreted as a scale of resources or area available, and is related to the concept of carrying capacity, which is the maximum population size that the environment can sustain indefinitely. In the sequel K will be large.
- r_K is the recombination probability per reproductive event. When two individuals with respective genotypes (α_1, β_1) and (α_2, β_2) in \mathcal{E} give birth, the newborn individual, either is a clone of one parent and carries alleles (α_1, β_1) or (α_2, β_2) with probability $(1 - r_K)/2$, or has a mixed genotype (α_1, β_2) or (α_2, β_1) with probability $r_K/2$.

We will use, for every $n = (n_{\alpha\beta}, (\alpha, \beta) \in \mathcal{E}) \in \mathbb{Z}_+^{\mathcal{E}}$, and $(\alpha, \beta) \in \mathcal{E}$, the notations

$$(2.1) \quad n_\alpha = n_{\alpha b_1} + n_{\alpha b_2}, \quad n_\beta = n_{A\beta} + n_{a\beta}, \quad \text{and} \quad |n| = n_A + n_a = n_{b_1} + n_{b_2}.$$

Let us now give the transition rates of $N^{(z,K)}$ when $N^{(z,K)}(t) = n \in \mathbb{Z}_+^{\mathcal{E}}$. An individual can die either from a natural death or from competition, whose strength depends on the carrying capacity K . Thus death rate of individuals $(\alpha, \beta) \in \mathcal{E}$ is given by:

$$(2.2) \quad d_{\alpha\beta}^K(n) = [D_\alpha + C_{\alpha, A} n_A / K + C_{\alpha, a} n_a / K] n_{\alpha\beta}.$$

An individual carrying allele $\alpha \in \mathcal{A}$ produces gametes with rate f_α , thus the relative frequencies of gametes available for reproduction are $p_{\alpha\beta}(n) = f_\alpha n_{\alpha\beta} / (f_A n_A + f_a n_a)$, $(\alpha, \beta) \in \mathcal{E}$. When an individual gives birth, he chooses his mate uniformly among the gametes available. Hence the probability of giving birth to an individual of a given genotype depends on the parents (the couple (ab_2, ab_1) is not able to generate an individual Ab_1). We detail the computation of $b_{Ab_1}^K(n)$:

$$\begin{aligned} b_{Ab_1}^K(n) &= f_A n_{Ab_1} [p_{Ab_1} + p_{Ab_2}/2 + p_{ab_1}/2 + (1 - r_K) p_{ab_2}/2] + f_A n_{Ab_2} [p_{Ab_1}/2 + r_K p_{ab_1}/2] \\ &\quad + f_a n_{ab_1} [p_{Ab_1}/2 + r_K p_{Ab_2}/2] + f_a n_{ab_2} (1 - r_K) p_{Ab_1}/2 \\ &= f_A n_{Ab_1} + r_K f_A f_a (n_{ab_1} n_{Ab_2} - n_{Ab_1} n_{ab_2}) / (f_A n_A + f_a n_a). \end{aligned}$$

If we denote by $\bar{\alpha}$ (resp. $\bar{\beta}$) the complement of α in \mathcal{A} (resp. β in \mathcal{B}), we obtain in the same way:

$$(2.3) \quad b_{\alpha\beta}^K(n) = f_\alpha n_{\alpha\beta} + r_K f_a f_A \frac{n_{\bar{\alpha}\bar{\beta}} n_{\alpha\bar{\beta}} - n_{\alpha\beta} n_{\bar{\alpha}\bar{\beta}}}{f_A n_A + f_a n_a}, \quad (\alpha, \beta) \in \mathcal{E}.$$

The definitions of death and birth rates in (2.2) and (2.3) ensure that the number of jumps is finite on every finite interval, and the population process is well defined.

When we focus on the dynamics of traits under selection A and a , we get the process $(N_A^{(z,K)}, N_a^{(z,K)})$. It has been studied by Champagnat in [5] and its death and birth rates, which are direct consequences of (2.2) and (2.3), satisfy:

$$(2.4) \quad d_\alpha^K(n) = \sum_{\beta \in \mathcal{B}} d_{\alpha\beta}^K(n) = \left[D_\alpha + C_{\alpha,A} \frac{n_A}{K} + C_{\alpha,a} \frac{n_a}{K} \right] n_\alpha, \quad b_\alpha^K(n) = \sum_{\beta \in \mathcal{B}} b_{\alpha\beta}^K(n) = f_\alpha n_\alpha, \quad \alpha \in \mathcal{A}.$$

Champagnat has proved that under some conditions the rescaled population process $(N_A^{(z,K)}/K, N_a^{(z,K)}/K)$ is well approximated by the following dynamical system,

$$(2.5) \quad \dot{n}_\alpha^{(z)} = (f_\alpha - D_\alpha - C_{\alpha,A} n_A^{(z)} - C_{\alpha,a} n_a^{(z)}) n_\alpha^{(z)}, \quad n_\alpha^{(z)}(0) = z_\alpha, \quad \alpha \in \mathcal{A}.$$

More precisely Theorem 3 (b) in [5] for every compact subset $B \in (\mathbb{R}_+^{A \times \mathcal{B}})^* \times (\mathbb{R}_+^{a \times \mathcal{B}})^*$ and finite real number T , we have for any $\delta > 0$,

$$(2.6) \quad \lim_{K \rightarrow \infty} \sup_{z \in B} \mathbb{P} \left(\sup_{0 \leq t \leq T, \alpha \in \mathcal{A}} |N_\alpha^{(z,K)}(t)/K - n_\alpha^{(z)}(t)| \geq \delta \right) = 0.$$

Moreover, if we assume

$$(2.7) \quad f_A > D_A, \quad f_a > D_a, \quad \text{and} \quad f_A - D_A > (f_A - D_A) \cdot \sup \left\{ C_{a,A}/C_{A,A}, C_{a,a}/C_{A,a} \right\},$$

then the dynamical system (2.5) has a unique attracting equilibrium $(0, \bar{n}_a)$ for initial condition z satisfying $z_a > 0$, and an unstable steady state $(\bar{n}_A, 0)$ where

$$(2.8) \quad \bar{n}_\alpha = \frac{f_\alpha - D_\alpha}{C_{\alpha,\alpha}} > 0, \quad \alpha \in \mathcal{A}.$$

Hence, Assumption (2.7) avoids the coexistence of alleles A and a , and \bar{n}_α is the equilibrium density of a monomorphic α -population per unit of carrying capacity. This implies that when K is large, the size of a monomorphic α -population stays near $\bar{n}_\alpha K$ for a long time (Theorem 3 (c) in [5]). Moreover, if we introduce the invasion fitness $S_{\alpha\bar{\alpha}}$ of a mutant α in a population $\bar{\alpha}$,

$$(2.9) \quad S_{\alpha\bar{\alpha}} = f_\alpha - D_\alpha - C_{\alpha,\bar{\alpha}} \bar{n}_{\bar{\alpha}}, \quad \alpha \in \mathcal{A},$$

it corresponds to the per capita growth rate of a mutant α when it appears in a population $\bar{\alpha}$ at its equilibrium density $\bar{n}_{\bar{\alpha}}$. Assumption (2.7) is equivalent to

Assumption 1. *Ecological parameters satisfy*

$$\bar{n}_A > 0, \quad \bar{n}_a > 0, \quad \text{and} \quad S_{Aa} < 0 < S_{aA}.$$

Under Assumption 1, with positive probability, the A -population becomes extinct and the a -population size reaches a vicinity of its equilibrium value $\bar{n}_a K$.

Let us now present the main results of this paper. We introduce the extinction time of the A -population, and the fixation event of the a -population. For $(z, K) \in \mathbb{R}_+^{\mathcal{E}} \times \mathbb{N}$:

$$(2.10) \quad T_{\text{ext}}^{(z,K)} := \inf \left\{ t \geq 0, N_A^{(z,K)}(t) = 0 \right\}, \quad \text{and} \quad \text{Fix}^{(z,K)} := \left\{ T_{\text{ext}}^{(z,K)} < \infty, N_a^{(z,K)}(T_{\text{ext}}^{(z,K)}) > 0 \right\}.$$

We are interested in the neutral allele proportions. We thus define for $t \geq 0$,

$$(2.11) \quad P_{\alpha,\beta}^{(z,K)}(t) = \frac{N_{\alpha\beta}^{(z,K)}(t)}{N_\alpha^{(z,K)}(t)}, \quad (\alpha, \beta) \in \mathcal{E}, K \in \mathbb{N}, z \in \mathbb{R}_+^{\mathcal{E}},$$

the proportion of alleles β in the α -population at time t , with the convention $0/0 = 0$.

We first focus on soft selective sweep. We assume that the alleles A and a were neutral and coexisted in a population with large carrying capacity K . At time 0, an environmental change makes the allele a favorable (in the sense that $S_{Aa} < 0 < S_{aA}$). Before stating the result, let us introduce the function F , defined for every $(z, r, t) \in (\mathbb{R}_+^{\mathcal{E}})^* \times [0, 1] \times \mathbb{R}_+$ by

$$(2.12) \quad F(z, r, t) = \int_0^t \frac{r f_A f_a n_A^{(z)}(s)}{f_A n_A^{(z)}(s) + f_a n_a^{(z)}(s)} \exp\left(-r f_A f_a \int_0^s \frac{n_A^{(z)}(u) + n_a^{(z)}(u)}{f_A n_A^{(z)}(u) + f_a n_a^{(z)}(u)} du\right) ds,$$

where $(n_A^{(z)}, n_a^{(z)})$ is the solution of the dynamical system (2.5). We notice that $F : t \in \mathbb{R}^+ \mapsto F(z, r, t)$ is non-negative and non-decreasing. Moreover, if we introduce the non-decreasing function $h : (z, r, t) \in (\mathbb{R}_+^{\mathcal{E}})^* \times [0, 1] \times \mathbb{R}_+ \mapsto r f_A f_a \int_0^t n_A^{(z)}(s) / (f_A n_A^{(z)}(s) + f_a n_a^{(z)}(s)) ds$, then

$$0 \leq F(z, r, t) \leq \int_0^t \partial_s h(z, r, s) e^{-h(z, r, s)} ds = e^{-h(z, r, 0)} - e^{-h(z, r, t)}.$$

Thus $F(z, r, t)$ has a limit in $[0, 1]$ when t goes to infinity and we can define

$$(2.13) \quad F(z, r) := \lim_{t \rightarrow \infty} F(z, r, t) \in [0, 1].$$

In the case of soft selective sweep, the selected allele gets to fixation with high probability. More precisely, Champagnat proved the following asymptotic result in [5]: for $z \in \mathbb{R}_+^{A \times \mathcal{B}} \times (\mathbb{R}_+^{a \times \mathcal{B}})^*$ and under Assumption 1,

$$(2.14) \quad \lim_{K \rightarrow \infty} \mathbb{P}(\text{Fix}^{(z, K)}) = 1.$$

We consider the soft selective sweep with recombination probability r_K satisfying:

Assumption 2.

$$\lim_{K \rightarrow \infty} r_K = r \in [0, 1].$$

Then we have the following result whose proof is deferred in Section 4:

Theorem 1. *Let z be in $\mathbb{R}_+^{A \times \mathcal{B}} \times (\mathbb{R}_+^{a \times \mathcal{B}})^*$ and Assumption 1 and 2 hold. Then on the fixation event $\text{Fix}^{(z, K)}$, the proportion $P_{a, b_1}^{(z, K)}$ of alleles b_1 at the A -population extinction time $T_{\text{ext}}^{(z, K)}$ converges in probability:*

$$\mathbb{P}\left(\mathbb{1}_{\text{Fix}^{(z, K)}} \left| P_{a, b_1}^{(z, K)}(T_{\text{ext}}^{(z, K)}) - \left[\frac{z_{Ab_1}}{z_A} F(z, r) + \frac{z_{ab_1}}{z_a} (1 - F(z, r)) \right] \right| > \varepsilon\right) \xrightarrow{K \rightarrow \infty} 0, \quad \forall \varepsilon > 0.$$

The neutral proportion at the end of the soft sweep is thus a weighted mean of initial proportions in populations A and a . In particular, soft sweep is responsible for a diminution of the number of neutral alleles with very low or very high proportions in the population, as remarked in [24]. We notice that the variation of neutral proportions only depends on initial state and on the solution of a two-dimensional Lotka-Volterra competitive system.

Now we focus on hard selective sweep: a mutant a appears in a large population and gets to fixation. We assume that the mutant appears in an A -population at ecological equilibrium, and carries the neutral allele b_1 . In other words, recalling Definition (2.8), we assume:

Assumption 3. *There exists $z_{Ab_1} \in]0, \bar{n}_A[$ such that $N^{(z^{(K)}, K)}(0) = \lfloor z^{(K)} K \rfloor$ with*

$$z^{(K)} = (z_{Ab_1}, \bar{n}_A - z_{Ab_1}, K^{-1}, 0).$$

In this case, the selected allele gets to fixation with positive probability. More precisely, Champagnat proved the following asymptotic result in [5]: under Assumptions 1 and 3,

$$(2.15) \quad \lim_{K \rightarrow \infty} \mathbb{P}\left(\text{Fix}^{(z^{(K)}, K)}\right) = \frac{S_{aA}}{f_a}.$$

In the case of strong selective sweep we will distinguish two different recombination regimes:

Assumption 4. *Strong recombination*

$$\lim_{K \rightarrow \infty} r_K \log K = \infty.$$

Assumption 5. *Weak recombination*

$$\limsup_{K \rightarrow \infty} r_K \log K < \infty.$$

We have the following results whose proofs are deferred in Sections 6 and 7:

Theorem 2. *Suppose that Assumption 1 and 3 hold. Then on the fixation event $\text{Fix}^{(z,K)}$ and under Assumption 4 or 5, the proportion $P_{a,b_1}^{(z,K)}$ of alleles b_1 at the A -population extinction time $T_{\text{ext}}^{(z,K)}$ converges in probability. More precisely, if Assumption 4 holds, then,*

$$\mathbb{P}\left(\mathbb{1}_{\text{Fix}^{(z,K)}} \left| P_{a,b_1}^{(z(K),K)}(T_{\text{ext}}^{(z(K),K)}) - \frac{z_{Ab_1}}{z_A} \right| > \varepsilon\right) \xrightarrow{K \rightarrow \infty} 0, \quad \forall \varepsilon > 0,$$

and if Assumption 5 holds, then,

$$\mathbb{P}\left(\mathbb{1}_{\text{Fix}^{(z,K)}} \left| P_{a,b_1}^{(z(K),K)}(T_{\text{ext}}^{(z(K),K)}) - \left[\frac{z_{Ab_1}}{z_A} + \frac{z_{Ab_2}}{z_A} \exp\left(-\frac{f_a r_K \log K}{S_{aA}}\right) \right] \right| > \varepsilon\right) \xrightarrow{K \rightarrow \infty} 0, \quad \forall \varepsilon > 0.$$

As stated in [5], the selective sweep has a characteristic time of order $\log K$. Thus, when $r_K \log K$ is large, a lot of recombinations occur during the sweep, and the neutral alleles are constantly exchanged by the populations A and a . As a consequence in the strong recombination case, the selective sweep does not modify the proportion of neutral alleles. On the contrary, when the recombination probability r_K is of order $1/\log K$ the number of recombinations undergone by a given lineage does not go to infinity, and the frequency of the neutral allele carried by the first mutant a increases. This phenomenon is called genetic hitch-hiking [26]: the selective sweep leads to a diminution of diversity around the novel fixed allele.

Remark 1. *The limits in the two regimes are consistent in the sense that*

$$\lim_{r_K \log K \rightarrow \infty} \left\{ \frac{z_{Ab_1}}{z_A} + \frac{z_{Ab_2}}{z_A} \exp\left(-\frac{f_a r_K \log K}{S_{aA}}\right) \right\} = \frac{z_{Ab_1}}{z_A}.$$

Moreover, let us notice that we can easily extend the results of Theorems 1 and 2 to a finite number of possible alleles b_1, b_2, \dots, b_i on the neutral locus.

3. A SEMI-MARTINGALE DECOMPOSITION

The expression of birth rate in (2.3) shows that the effect of recombination depends on the recombination probability r_K but also on the population state via the term $n_{\bar{\alpha}\bar{\beta}} n_{\alpha\beta} - n_{\alpha\beta} n_{\bar{\alpha}\bar{\beta}}$. This quantity is linked with the linkage disequilibrium of the population, which is the occurrence of some allele combinations more or less often than would be expected from a random formation of haplotypes (see [13] Section 3.3 for an introduction to this notion). Proposition 1 states a semi-martingale representation of the neutral allele proportions.

Proposition 1. *Let (α, z, K) be in $\mathcal{A} \times (\mathbb{R}_+^{\mathcal{G}})^* \times \mathbb{N}$. The process $(P_{\alpha,b_1}^{(z,K)}(t), t \geq 0)$ defined in (2.11) is a semi-martingale and we have the following decomposition:*

$$(3.1) \quad P_{\alpha,b_1}^{(z,K)}(t) = P_{\alpha,b_1}^{(z,K)}(0) + M_{\alpha}^{(z,K)}(t) + r_K f_A f_a \int_0^t \frac{(N_{\bar{\alpha}b_1}^{(z,K)} N_{\alpha b_2}^{(z,K)} - N_{\alpha b_1}^{(z,K)} N_{\bar{\alpha}b_2}^{(z,K)})(s)}{(N_{\alpha}^{(z,K)}(s) + 1)(f_A N_A^{(z,K)}(s) + f_a N_a^{(z,K)}(s))} ds,$$

where the process $(M_{\alpha}^{(z,K)}(t), t \geq 0)$ is a martingale bounded on every interval $[0, t]$ whose quadratic variation is given by (3.7).

Remark 2. The process $N_{ab_2}N_{Ab_1} - N_{ab_1}N_{Ab_2}$ will play a major role in the dynamics of neutral proportions. Indeed it is a measure of the neutral proportion disequilibrium between the A and a-populations as it satisfies:

$$(3.2) \quad N_A N_a (P_{A,b_1} - P_{a,b_1}) = N_{ab_2} N_{Ab_1} - N_{ab_1} N_{Ab_2}.$$

Proof of Proposition 1. In the vein of Fournier and Méléard [17] we represent the population process in terms of Poisson measure. Let $Q(ds, d\theta)$ be a Poisson random measure on \mathbb{R}_+^2 with intensity $dsd\theta$, and $(e_{\alpha\beta}, (\alpha, \beta) \in \mathcal{E})$ the canonical basis of $\mathbb{R}^{\mathcal{E}}$. According to (2.4) a jump occurs at rate $\sum_{(\alpha,\beta) \in \mathcal{E}} (b_{\alpha\beta}^K(N) + d_{\alpha\beta}^K(N)) = f_a N_a + d_a^K(N) + f_A N_A + d_A^K(N)$. We decompose on possible jumps that may occur: births and deaths for a-individuals and births and deaths for A-individuals. Itô's formula with jumps (see [18] p. 66) yields for every function h measurable and bounded on $\mathbb{R}_+^{\mathcal{E}}$:

$$(3.3) \quad \begin{aligned} h(N(t)) &= h(N(0)) + \int_0^t \int_{\mathbb{R}_+} \left\{ \sum_{\alpha \in \mathcal{A}} \left(h(N(s^-) + e_{\alpha b_1}) \mathbb{1}_{0 < \theta - \mathbb{1}_{\alpha=A}(f_a N_a(s^-) + d_a^K(N(s^-)) \leq b_{\alpha b_1}^K(N(s^-))} \right. \right. \\ &\quad + h(N(s^-) + e_{\alpha b_2}) \mathbb{1}_{b_{\alpha b_1}^K(N(s^-)) < \theta - \mathbb{1}_{\alpha=A}(f_a N_a(s^-) + d_a^K(N(s^-)) \leq f_a N_a(s^-)} \\ &\quad + h(N(s^-) - e_{\alpha b_1}) \mathbb{1}_{0 < \theta - f_a N_a(s^-) - \mathbb{1}_{\alpha=A}(f_a N_a(s^-) + d_a^K(N(s^-)) \leq d_{\alpha b_1}^K(N(s^-))} \\ &\quad + h(N(s^-) - e_{\alpha b_2}) \mathbb{1}_{d_{\alpha b_1}^K(N(s^-)) < \theta - f_a N_a(s^-) - \mathbb{1}_{\alpha=A}(f_a N_a(s^-) + d_a^K(N(s^-)) \leq d_a^K(N(s^-))} \\ &\quad \left. \left. - h(N(s^-)) \mathbb{1}_{\theta \leq f_a N_a(s^-) + d_a^K(N(s^-)) + f_A N_A(s^-) + d_A^K(N(s^-))} \right\} Q(ds, d\theta). \end{aligned}$$

Let us introduce the functions μ_K^α defined for $\alpha \in \mathcal{A}$ and (s, θ) in $\mathbb{R}_+ \times \mathbb{R}_+$ by,

$$(3.4) \quad \begin{aligned} \mu_K^\alpha(N, s, \theta) &= \frac{N_{\alpha b_2}(s)}{(N_\alpha(s) + 1)N_\alpha(s)} \mathbb{1}_{0 < \theta - \mathbb{1}_{\alpha=A}(f_a N_a(s) + d_a^K(N(s)) \leq b_{\alpha b_1}^K(N(s))} \\ &\quad - \frac{N_{\alpha b_1}(s)}{(N_\alpha(s) + 1)N_\alpha(s)} \mathbb{1}_{b_{\alpha b_1}^K(N(s)) < \theta - \mathbb{1}_{\alpha=A}(f_a N_a(s) + d_a^K(N(s)) \leq f_a N_a(s)} \\ &\quad - \frac{N_{\alpha b_2}(s)}{(N_\alpha(s) - 1)N_\alpha(s)} \mathbb{1}_{0 < \theta - f_a N_a(s) - \mathbb{1}_{\alpha=A}(f_a N_a(s) + d_a^K(N(s)) \leq d_{\alpha b_1}^K(N(s))} \\ &\quad + \frac{N_{\alpha b_1}(s)}{(N_\alpha(s) - 1)N_\alpha(s)} \mathbb{1}_{d_{\alpha b_1}^K(N(s)) < \theta - f_a N_a(s) - \mathbb{1}_{\alpha=A}(f_a N_a(s) + d_a^K(N(s)) \leq d_a^K(N(s))}, \end{aligned}$$

with the convention $0/0 = 0$. We can represent the neutral allele proportions P_{α, b_1} in terms of Poisson measure as,

$$(3.5) \quad P_{\alpha, b_1}(t) = P_{\alpha, b_1}(0) + \int_0^t \int_{\mathbb{R}_+} \mu_K^\alpha(N, s^-, \theta) Q(ds, d\theta), \quad t \geq 0.$$

A direct calculation gives

$$\int_0^\infty \mu_K^\alpha(N, s, \theta) d\theta = r_K f_A f_a \frac{N_{\bar{\alpha} b_1}(s) N_{\alpha b_2}(s) - N_{\alpha b_1}(s) N_{\bar{\alpha} b_2}(s)}{(N_\alpha(s) + 1)(f_A N_A(s) + f_a N_a(s))}.$$

Thus if we introduce the compensated Poisson measure $\tilde{Q}(ds, d\theta) := Q(ds, d\theta) - dsd\theta$, we get that

$$\begin{aligned} M_\alpha(t) &:= \int_0^t \int_{\mathbb{R}_+} \mu_K^\alpha(N, s^-, \theta) \tilde{Q}(ds, d\theta) \\ &= P_{\alpha, b_1}(t) - P_{\alpha, b_1}(0) - r_K f_A f_a \int_0^t \frac{N_{\bar{\alpha} b_1}(s) N_{\alpha b_2}(s) - N_{\alpha b_1}(s) N_{\bar{\alpha} b_2}(s)}{(N_\alpha(s) + 1)(f_A N_A(s) + f_a N_a(s))} ds \end{aligned}$$

is a local martingale. By construction the process P_{α, b_1} has values in $[0, 1]$ and as $r_K \leq 1$,

$$(3.6) \quad \sup_{s \leq t} \left| r_K f_A f_a \int_0^s \frac{N_{\bar{\alpha} b_1}(s) N_{\alpha b_2}(s) - N_{\alpha b_1}(s) N_{\bar{\alpha} b_2}(s)}{(N_\alpha(s) + 1)(f_A N_A(s) + f_a N_a(s))} ds \right| \leq r_K f_\alpha t \leq f_\alpha t, \quad t \geq 0.$$

Thus M_α is a square integrable pure jump martingale bounded on every finite interval with quadratic variation

$$\begin{aligned}
\langle M_\alpha \rangle_t &= \int_0^t \int_{\mathbb{R}^+} \left(\mu_K^\alpha(N, s, \theta) \right)^2 ds d\theta \\
&= \int_0^t \left\{ P_{\alpha, b_1}(s)(1 - P_{\alpha, b_1}(s)) \left[(D_\alpha + \frac{C_{\alpha, \alpha}}{K} N_\alpha(s) + \frac{C_{\alpha, \bar{\alpha}}}{K} N_{\bar{\alpha}}(s)) \frac{\mathbb{1}_{N_\alpha(s) \geq 2} N_\alpha(s)}{(N_\alpha(s) - 1)^2} \right. \right. \\
(3.7) \quad &\quad \left. \left. + \frac{f_\alpha N_\alpha(s)}{(N_\alpha(s) + 1)^2} \right] + \frac{r_K f_A f_a (N_{\bar{\alpha} b_1} N_{\alpha b_2} - N_{\alpha b_1} N_{\bar{\alpha} b_2})(s) (1 - 2P_{\alpha, b_1}(s))}{(N_\alpha(s) + 1)^2 (f_A N_A(s) + f_a N_a(s))} \right\} ds.
\end{aligned}$$

This ends up the proof of Proposition 1 \square

Remark 3. Let us mention two properties of the functions μ_K^α , defined in (3.4), which will be useful in the sequel. Firstly by definition we have for all (s, θ) in $\mathbb{R}_+ \times \mathbb{R}_+$,

$$(3.8) \quad \mu_K^A(N, s, \theta) \mu_K^a(N, s, \theta) = 0.$$

Secondly, the convention $0/0 = 0$ yields the following equality for $\alpha \in \mathcal{A}$ and all (s, θ) in $\mathbb{R}_+ \times \mathbb{R}_+$,

$$(3.9) \quad \mu_K^\alpha(N, s, \theta) \mathbb{1}_{N_\alpha(s) \geq 1} = \mu_K^\alpha(N, s, \theta).$$

For sake of simplicity we will use more often the second notation, but in Section 6 the first notation will also be useful.

Lemma 3.1 states properties of the quadratic variation widely used in the forthcoming proofs. To this aim we introduce a compact interval containing the equilibrium size of the A -population,

$$(3.10) \quad I_\varepsilon^K := \left[K \left(\bar{n}_A - 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right), K \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right) \right],$$

and the stopping times T_ε^K and \tilde{T}_ε^K , which denote respectively the hitting time of $[\varepsilon K]$ by the mutant population and the exit time of I_ε^K by the resident population,

$$(3.11) \quad T_\varepsilon^K := \inf\{t \geq 0, N_A^K(t) = \lfloor \varepsilon K \rfloor\}, \quad \tilde{T}_\varepsilon^K := \inf\{t \geq 0, N_A^K(t) \notin I_\varepsilon^K\}.$$

Lemma 3.1. For $v < \infty$, there exists a finite $C(v)$ such that $(N_A^{(z,K)}(t), N_a^{(z,K)}(t)) \in [0, vK]^2$ implies

$$(3.12) \quad \frac{d}{dt} \langle M_\alpha^{(z,K)} \rangle_t = \int_{\mathbb{R}^+} \left(\mu_K^\alpha(N^{(z,K)}, t, \theta) \right)^2 d\theta \leq C(v) \frac{\mathbb{1}_{N_\alpha(t) \geq 1}}{N_\alpha(t)}, \quad \alpha \in \mathcal{A}.$$

Under Assumptions 1 and 3, there exist $k_0 \in \mathbb{N}$, $\varepsilon_0 > 0$ and a pure jump martingale \tilde{M} such that for $\varepsilon \leq \varepsilon_0$ and $t \geq 0$,

$$(3.13) \quad e^{\frac{S_{AA}}{2(k_0+1)} t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \int_{\mathbb{R}^+} \left[\mu_K^a(N^{(z^{(k_0)}, K)}, t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K, \theta) \right]^2 d\theta \leq (k_0 + 1) C \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right) \tilde{M}_{t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K},$$

and

$$(3.14) \quad \mathbb{E} \left[\tilde{M}_{t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \right] \leq \frac{1}{k_0 + 1}.$$

Proof. Equation (3.12) is a direct consequence of (3.7). To prove (3.13) and (3.14), let us first remark that according to Assumption 1, there exists $k_0 \in \mathbb{N}$ such that for ε small enough and $k \in \mathbb{Z}_+$,

$$\frac{f_a(k_0 + k - 1) - (D_a + C_{a,A} \bar{n}_A + \varepsilon(C_{a,a} + 2C_{A,a} C_{a,A} / C_{A,A}))(k_0 + k + 1)}{k_0 + k - 1} \geq \frac{S_{AA}}{2}.$$

This implies in particular that for every $t < T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K$,

$$(3.15) \quad \frac{f_a N_a(t)(N_a(t) + k_0 - 1) - d_a(N(t))(N_a(t) + k_0 + 1)}{(N_a(t) + k_0 - 1)(N_a(t) + k_0 + 1)} \geq \frac{S_{AA} N_a(t)}{2(N_a(t) + k_0 + 1)} \geq \frac{S_{AA} \mathbb{1}_{N_a(t) \geq 1}}{2(k_0 + 1)},$$

where the death rate d_a has been defined in (2.4). For sake of simplicity let us introduce the process X defined as follows:

$$X(t) = \frac{\mathbb{1}_{N_a(t) \geq 1}}{N_a(t) + k_0} \exp\left(\frac{S_{aA}t}{2(k_0 + 1)}\right), \quad \forall t \geq 0.$$

Applying Itô's formula with jumps we get for every $t \geq 0$:

$$(3.16) \quad X(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K) = \tilde{M}(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K) + \int_0^{t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \left(\frac{S_{aA}}{2(k_0 + 1)} - \frac{f_a N_a(s)(N_a(s) + k_0 - 1) - d_a^K(N(s))(N_a(s) + k_0 + 1)}{(N_a(s) + k_0 - 1)(N_a(s) + k_0 + 1)} \right) X(s) ds,$$

where the martingale \tilde{M} has the following expression:

$$(3.17) \quad \tilde{M}(t) = \frac{1}{k_0 + 1} + \int_0^t \int_{\mathbb{R}_+} \tilde{Q}(ds, d\theta) \mathbb{1}_{N_a(s^-) \geq 1} \exp\left(\frac{S_{aA}s}{2(k_0 + 1)}\right) \left[\frac{\mathbb{1}_{\theta \leq f_a N_a(s^-)}}{N_a(s^-) + k_0 + 1} + \frac{\mathbb{1}_{f_a N_a(s^-) < \theta \leq f_a N_a(s^-) + d_a(N(s^-))}}{N_a(s^-) + k_0 - 1} - \frac{\mathbb{1}_{\theta \leq f_a N_a(s^-) + d_a(N(s^-))}}{N_a(s^-) + k_0} \right].$$

Thanks to (3.15) the integral in (3.16) is nonpositive. Moreover, according to (3.12), for $t \leq T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K$,

$$\int_{\mathbb{R}_+} \left(\mu_K^a(N^{(z^{(K)}, K)}, t, \theta) \right)^2 d\theta \leq C \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right) \frac{\mathbb{1}_{N_a(t) \geq 1}}{N_a(t)} \leq (k_0 + 1) C \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right) X(t) \exp\left(-\frac{S_{aA}t}{2(k_0 + 1)}\right)$$

which ends the proof. \square

4. PROOF OF THEOREM 1

In this section we suppose that Assumptions 1 and 2 hold. For $\varepsilon \leq C_{a,a}/C_{a,A} \wedge 2|S_{AA}|/C_{A,a}$ and z in $(\mathbb{R}_+^{A \times \mathcal{B}})^* \times (\mathbb{R}_+^{a \times \mathcal{B}})^*$ we introduce a deterministic time $t_\varepsilon(z)$ after which the solution $(n_A^{(z)}, n_a^{(z)})$ of the dynamical system (2.5) is close to the stable equilibrium $(0, \bar{n}_a)$:

$$(4.1) \quad t_\varepsilon(z) := \inf\{s \geq 0, \forall t \geq s, (n_A^{(z)}(t), n_a^{(z)}(t)) \in [0, \varepsilon^2/2] \times [\bar{n}_a - \varepsilon/2, \infty)\}.$$

Once $(n_A^{(z)}, n_a^{(z)})$ has reached the set $[0, \varepsilon^2/2] \times [\bar{n}_a - \varepsilon/2, \infty)$ it no more escapes from it. Moreover, according to Assumption 1 on stable equilibrium, $t_\varepsilon(z)$ is finite.

First we compare the population process with a four dimensional dynamical system on the finite time interval $[0, t_\varepsilon(z)]$. Then we study this dynamical system and get an approximation of the proportion of neutral alleles b_1 at time $t_\varepsilon(z)$. Finally, Lemma 4.2 states that during the A-population extinction period, this proportion stays nearly constant.

4.1. Comparison with a four dimensional dynamical system. We denote by $n^{(z)} = (n_{\alpha\beta}^{(z)}, (\alpha, \beta) \in \mathcal{E})$ the solution of

$$(4.2) \quad \dot{n}_{\alpha\beta}^{(z)} = \left(f_\alpha - (D_\alpha + C_{\alpha,A} n_A^{(z)} + C_{\alpha,a} n_a^{(z)}) \right) n_{\alpha\beta}^{(z)} + r f_A f_a \frac{n_{\bar{\alpha}\bar{\beta}}^{(z)} n_{\alpha\bar{\beta}}^{(z)} - n_{\alpha\beta}^{(z)} n_{\bar{\alpha}\bar{\beta}}^{(z)}}{f_A n_A^{(z)} + f_a n_a^{(z)}}, \quad (\alpha, \beta) \in \mathcal{E},$$

with initial condition $n^{(z)}(0) = z \in \mathbb{R}_+^{\mathcal{E}}$. Then we have the following comparison result:

Lemma 4.1. *Let z be in $\mathbb{R}_+^{\mathcal{E}}$. Then*

$$(4.3) \quad \lim_{K \rightarrow \infty} \sup_{s \leq t_\varepsilon(z)} \|N^{(z,K)}(s)/K - n^{(z)}(s)\| = 0 \quad a.s.$$

where $\|\cdot\|$ denotes the L^1 -Norm on $\mathbb{R}^{\mathcal{E}}$.

Proof. The proof relies on a slight modification of Theorem 2.1 p. 456 in Ethier and Kurtz [16]. The rescaled birth and death rates

$$\tilde{b}_{\alpha\beta}^K(n) = \frac{1}{K} b_{\alpha\beta}^K(Kn) = f_\alpha n_{\alpha\beta} + r_K f_A f_A \frac{n_{\bar{\alpha}\beta} n_{\alpha\bar{\beta}} - n_{\alpha\beta} n_{\bar{\alpha}\bar{\beta}}}{f_A n_A + f_a n_a}, \quad (\alpha, \beta) \in \mathcal{E}, n \in N^\mathcal{E},$$

and

$$\tilde{d}_{\alpha\beta}(n) = \frac{1}{K} d_{\alpha\beta}^K(Kn) = [D_\alpha + C_{\alpha,A} n_A + C_{\alpha,a} n_a] n_{\alpha\beta}, \quad (\alpha, \beta) \in \mathcal{E}, n \in N^\mathcal{E},$$

are Lipschitz and bounded on every compact subset of $\mathbb{N}^\mathcal{E}$. The only difference with [16] is that $\tilde{b}_{\alpha\beta}^K$ depends on K . We have

$$\frac{N^{(z,K)}(t)}{K} = \frac{\lfloor zK \rfloor}{K} + \text{Mart}^{(z,K)}(t) + \int_0^t \sum_{(\alpha,\beta) \in \mathcal{E}} e_{\alpha\beta} \left(\tilde{b}_{\alpha\beta}^K \left(\frac{N^{(z,K)}(s)}{K} \right) - \tilde{d}_{\alpha\beta} \left(\frac{N^{(z,K)}(s)}{K} \right) \right) ds,$$

where $\text{Mart}^{(z,K)}$ is a martingale, and we recall that $(e_{\alpha\beta}, (\alpha, \beta) \in \mathcal{E})$ is the canonical basis of $\mathbb{R}_+^\mathcal{E}$. If we denote by $\tilde{b}_{\alpha\beta}^\infty$ the function

$$\tilde{b}_{\alpha\beta}^\infty(n) = f_\alpha n_{\alpha\beta} + r f_A f_A \frac{n_{\bar{\alpha}\beta} n_{\alpha\bar{\beta}} - n_{\alpha\beta} n_{\bar{\alpha}\bar{\beta}}}{f_A n_A + f_a n_a}, \quad (\alpha, \beta) \in \mathcal{E}, n \in N^\mathcal{E},$$

we get

$$n^{(z)}(t) = z + \int_0^t \sum_{(\alpha,\beta) \in \mathcal{E}} e_{\alpha\beta} \left(\tilde{b}_{\alpha\beta}^\infty \left(n^{(z)}(s) \right) - \tilde{d}_{\alpha\beta} \left(n^{(z)}(s) \right) \right) ds.$$

Hence we have for every $t \leq t_\varepsilon(z)$,

$$\begin{aligned} \left| \frac{N^{(z,K)}(t)}{K} - n^{(z)}(t) \right| &\leq \left| \frac{\lfloor zK \rfloor}{K} - z \right| + \left| \text{Mart}^{(z,K)}(t) \right| + \int_0^t \sum_{(\alpha,\beta) \in \mathcal{E}} \left| \tilde{b}_{\alpha\beta}^\infty \left(n^{(z)}(s) \right) - \tilde{b}_{\alpha\beta}^K \left(n^{(z)}(s) \right) \right| ds \\ &\quad + \int_0^t \sum_{(\alpha,\beta) \in \mathcal{E}} \left| \left(\tilde{b}_{\alpha\beta}^K - \tilde{d}_{\alpha\beta} \right) \left(\frac{N^{(z,K)}(s)}{K} \right) - \left(\tilde{b}_{\alpha\beta}^K - \tilde{d}_{\alpha\beta} \right) \left(n^{(z)}(s) \right) \right| ds, \end{aligned}$$

and there exists a finite constant M such that

$$\left| \frac{N^{(z,K)}(t)}{K} - n^{(z)}(t) \right| \leq \frac{1}{K} + \left| \text{Mart}^{(z,K)}(t) \right| + |r - r_K| M t_\varepsilon(z) + M \int_0^t \left| \frac{N^{(z,K)}(s)}{K} - n^{(z)}(s) \right| ds.$$

But following Ethier and Kurtz, we get

$$\lim_{K \rightarrow \infty} \sup_{s \leq t_\varepsilon(z)} |\text{Mart}^{(z,K)}| = 0, \quad \text{a.s.},$$

and we end the proof by using Gronwall's Lemma. \square

Once we know that the rescaled population process is close to the solution of the dynamical system (4.2), we can study this latter.

Lemma 4.2. *Let z be in $\mathbb{R}^\mathcal{E}$ such that $z_a > 0$. Then $n_a^{(z)}(t)$ and $n_{ab_1}^{(z)}(t)$ have a finite limit when t goes to infinity, and there exists a positive constant ε_0 such that for every $\varepsilon \leq \varepsilon_0$,*

$$\left| \frac{n_{ab_1}^{(z)}(\infty)}{n_a^{(z)}(\infty)} - \frac{n_{ab_1}^{(z)}(t_\varepsilon(z))}{n_a^{(z)}(t_\varepsilon(z))} \right| \leq \frac{2f_a \varepsilon^2}{\bar{n}_A |S_{AA}|}.$$

Proof. We already know that $n_a^{(z)}(t)$ goes to \bar{n}_a at infinity. If we define the functions

$$p_{\alpha,b_1}^{(z)} = n_{\alpha b_1}^{(z)} / n_\alpha^{(z)}, \quad \alpha \in \mathcal{A}, \quad \text{and} \quad g^{(z)} = p_{A,b_1}^{(z)} - p_{a,b_1}^{(z)},$$

we easily check that $\phi : (n_{Ab_1}^{(z)}, n_{Ab_2}^{(z)}, n_{ab_1}^{(z)}, n_{ab_2}^{(z)}) \mapsto (n_A^{(z)}, n_a^{(z)}, g^{(z)}, p_{a,b_1}^{(z)})$ defines a change of variables from $(\mathbb{R}_+^*)^{\mathcal{S}}$ to $\mathbb{R}_+^{2*} \times]-1, 1[\times]0, 1[$, and (4.2) is equivalent to:

$$(4.4) \quad \begin{cases} \dot{n}_\alpha^{(z)} = (f_\alpha - (D_\alpha + C_{\alpha,A}n_A^{(z)} + C_{\alpha,a}n_a^{(z)}))n_\alpha^{(z)}, & \alpha \in \mathcal{A} \\ \dot{g}^{(z)} = -g^{(z)} \left(r f_A f_a (n_A^{(z)} + n_a^{(z)}) / (f_A n_A^{(z)} + f_a n_a^{(z)}) \right) \\ \dot{p}_{a,b_1}^{(z)} = g^{(z)} \left(r f_A f_a n_A^{(z)} / (f_A n_A^{(z)} + f_a n_a^{(z)}) \right). \end{cases}$$

Moreover, a direct integration yields

$$(4.5) \quad p_{a,b_1}^{(z)}(t) = p_{a,b_1}^{(z)}(0) - (p_{a,b_1}^{(z)}(0) - p_{A,b_1}^{(z)}(0))F(z, r, t),$$

where the function F has been defined in (2.12). According to (2.13), $F(z, r, \cdot) : t \mapsto F(z, r, t)$ has a finite limit when t goes to infinity. Hence $p_{a,b_1}^{(z)}$ also admits a limit at infinity. Let $\varepsilon \leq |S_{Aa}|/C_{A,a} \wedge C_{a,a}/C_{A,a} \wedge \bar{n}_a/2$, and $t_\varepsilon(z)$ defined in (4.1). Then for $t_\varepsilon(z) \leq t$,

$$\dot{n}_A^{(z)}(t) \leq (f_A - D_A - C_{Aa}(\bar{n}_a - \varepsilon/2))n_A^{(z)}(t) \leq S_{Aa}n_A^{(z)}(t)/2 < 0.$$

Recalling that $r \leq 1$ and $|g(t)| \leq 1$ for all $t \geq 0$ we get:

$$(4.6) \quad \left| p_{a,b_1}^{(z)}(\infty) - p_{a,b_1}^{(z)}(t_\varepsilon(z)) \right| \leq \int_{t_\varepsilon(z)}^\infty \frac{f_A f_a n_A^{(z)}(s) ds}{f_A n_A^{(z)}(s) + f_a n_a^{(z)}(s)} \leq \frac{f_A \varepsilon^2}{\bar{n}_a} \int_0^\infty e^{S_{Aa}s/2} ds \leq \frac{2f_A \varepsilon^2}{\bar{n}_a |S_{Aa}|},$$

which ends up the proof. \square

4.2. A-population extinction. The deterministic approximation (4.2) fails when the A -population size becomes too small. We will compare N_A with birth and death processes to study the last period of mutant invasion. We show that during this period, the number of A individuals is so small that it has no influence on the neutral proportion in the a -population, which stays nearly constant. Before stating the result, we recall that $T_{\text{ext}}^{(z,K)}$ is defined in (2.10), and we introduce the compact set Θ :

$$(4.7) \quad \Theta := \{z \in \mathbb{R}_+^{A \times \mathcal{B}} \times (\mathbb{R}_+^{a \times \mathcal{B}})^*, z_A \leq \varepsilon^2 \text{ and } |z_a - \bar{n}_a| \leq \varepsilon\}.$$

Lemma 4.3. *Let z be in Θ . Then under Assumption 1, there exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$,*

$$\limsup_{K \rightarrow \infty} \mathbb{P} \left(\sup_{t \leq T_{\text{ext}}^{(z,K)}} \left| P_{a,b_1}^{(z,K)}(t) - P_{a,b_1}^{(z,K)}(0) \right| > \varepsilon \right) \leq c\varepsilon.$$

Proof. Let $M'' = 3 + (f_a + C_{a,A})/C_{a,a}$ and let us introduce the stopping time:

$$(4.8) \quad S_\varepsilon^K(z) := \inf\{t \geq 0, N_A^{(z,K)}(t) > \varepsilon K \text{ or } |N_a^{(z,K)}(t) - \bar{n}_a K| > M'' \varepsilon K\}.$$

Let z be in Θ and Z^1 be a birth and death process with birth rate f_A , death rate $D_A + (\bar{n}_a - M'' \varepsilon)C_{A,a}$, and initial state $\lceil \varepsilon^2 K \rceil$. Then on $[0, S_\varepsilon^K(z)]$, N_A and Z^1 have the same birth rate, and Z^1 has a smaller death rate than N_A . Thus according to Theorem 2 in [5], we can construct the processes N and Z^1 on the same probability space such that:

$$(4.9) \quad N_A(t) \leq Z_t^1, \quad \forall t < S_\varepsilon^K(z).$$

Moreover, if we denote by T_0^1 the extinction time of Z^1 , $T_0^1 := \inf\{t \geq 0, Z_t^1 = 0\}$, and recall that

$$(4.10) \quad f_A - D_A - (\bar{n}_a - M'' \varepsilon)C_{A,a} = S_{Aa} + M'' C_{A,a} \varepsilon < S_{Aa}/2 < 0, \quad \forall \varepsilon < |S_{Aa}|/(2M'' C_{A,a}),$$

we get according to (A.10) that for $z \leq \varepsilon^2$ and $L(\varepsilon, K) = 2 \log K / |S_{Aa} + M'' \varepsilon C_{A,a}|$,

$$\mathbb{P}_{\lceil zK \rceil} \left(T_0^1 \leq L(\varepsilon, K) \right) \geq \exp \left(\lceil \varepsilon^2 K \rceil \left[\log(K^2 - 1) - \log(K^2 - f_A(D_A + (\bar{n}_a - M'' \varepsilon)C_{A,a})^{-1}) \right] \right).$$

Thus:

$$(4.11) \quad \lim_{K \rightarrow \infty} \mathbb{P}_{\lceil zK \rceil} \left(T_0^1 < L(\varepsilon, K) \right) = 1.$$

Moreover, Equation (A.4) ensures the existence of a finite c such that for ε small enough,

$$(4.12) \quad \mathbb{P}\left(L(\varepsilon, K) < S_\varepsilon^K(z)\right) \geq 1 - c\varepsilon.$$

Equations (4.11) and (4.12) imply

$$(4.13) \quad \liminf_{K \rightarrow \infty} \mathbb{P}\left(T_0^1 < L(\varepsilon, K) < S_\varepsilon^K(z)\right) \geq 1 - c\varepsilon.$$

According to Coupling (4.9) we have the inclusion $\{T_0^1 < L(\varepsilon, K) < S_\varepsilon^K(z)\} \subset \{T_{\text{ext}}^K < L(\varepsilon, K) < S_\varepsilon^K(z)\}$. Adding (4.13) we finally get:

$$(4.14) \quad \liminf_{K \rightarrow \infty} \mathbb{P}(T_{\text{ext}}^K < L(\varepsilon, K) < S_\varepsilon^K(z)) \geq 1 - c\varepsilon.$$

Recall the martingale decomposition of P_{a,b_1} in (3.1). To bound the difference $|P_{a,b_1}(t) - P_{a,b_1}(0)|$ we bound independently the martingale $M_a(t)$ and the integral $|P_{a,b_1}(t) - P_{a,b_1}(0) - M_a(t)|$. On one hand Doob's Maximal Inequality and Equation (3.12) imply:

$$(4.15) \quad \mathbb{P}\left(\sup_{t \leq L(\varepsilon, K) \wedge S_\varepsilon^K} |M_a(t)| > \frac{\varepsilon}{2}\right) \leq \frac{4}{\varepsilon^2} \mathbb{E}\left[\langle M_a \rangle_{L(\varepsilon, K) \wedge S_\varepsilon^K}\right] \leq \frac{4C(\bar{n}_a + M''\varepsilon)L(\varepsilon, K)}{\varepsilon^2 K(\bar{n}_a - M''\varepsilon)}.$$

On the other hand the inequality $|N_{Ab_1}N_{Ab_2} - N_{Ab_1}N_{Ab_2}| \leq N_A N_a$ yields for $t \geq 0$

$$\left| \int_0^{t \wedge S_\varepsilon^K} \frac{r_K f_A f_a (N_{Ab_1} N_{Ab_2} - N_{Ab_1} N_{Ab_2})(s)}{(N_a(s) + 1)(f_A N_a(s) + f_a N_a(s))} ds \right| \leq \int_0^{t \wedge S_\varepsilon^K} \frac{f_A N_a(s) ds}{(\bar{n}_a - \varepsilon M'')K}.$$

Hence decomposition (3.1), Markov inequality, and Equations (4.9), (A.8) and (4.10) yield

$$(4.16) \quad \mathbb{P}\left(|(P_{a,b_1} - M_a)(t \wedge S_\varepsilon^K) - P_{a,b_1}(0)| > \frac{\varepsilon}{2}\right) \leq \frac{2f_A \varepsilon^2}{\varepsilon(\bar{n}_a - \varepsilon M'')} \int_0^t e^{S_{Aa}s/2} ds \leq \frac{4f_A \varepsilon}{(\bar{n}_a - \varepsilon M'')|S_{Aa}|}.$$

Taking the limit of (4.15) when K goes to infinity and adding (4.16) end the proof. \square

4.3. End of the proof of Theorem 1. We have:

$$\begin{aligned} & \left| P_{a,b_1}^{(z,K)}(T_{\text{ext}}^{(z,K)}) - p_{a,b_1}^{(z)}(\infty) \right| \leq \\ & \left| P_{a,b_1}^{(z,K)}(T_{\text{ext}}^{(z,K)}) - P_{a,b_1}^{(z,K)}(t_\varepsilon(z)) \right| + \left| P_{a,b_1}^{(z,K)}(t_\varepsilon(z)) - p_{a,b_1}^{(z)}(t_\varepsilon(z)) \right| + \left| p_{a,b_1}^{(z)}(t_\varepsilon(z)) - p_{a,b_1}^{(z)}(\infty) \right|. \end{aligned}$$

To bound the two last terms we use respectively Lemmas 4.1 and 4.2. For the first term of right hand side, (2.6) ensures that with high probability, $N^{(z,K)}(t_\varepsilon(z)) \in \Theta$ and $t_\varepsilon(z) < T_{\text{ext}}^{(z,K)}$. Lemma 4.3, Equation (2.14) and Markov's Inequality allow us to conclude that for ε small enough

$$\limsup_{K \rightarrow \infty} \mathbb{P}(\mathbb{1}_{\text{Fix}^{(z,K)}} |P_{a,b_1}^{(z,K)}(T_{\text{ext}}^{(z,K)}) - p_{a,b_1}^{(z)}(\infty)| > 3\varepsilon) \leq c\varepsilon,$$

for a finite c . It is equivalent to the convergence in probability, which concludes the proof.

5. A COUPLING WITH TWO BIRTH AND DEATH PROCESSES

In Sections 6 and 7 we suppose that Assumptions 1 and 3 hold and we denote by N^K the process $N^{(z^{(K)}, K)}$. As it will appear in the proof of Theorem 2 the first period of mutant invasion, which ends at time T_ε^K when the mutant population size hits $\lfloor \varepsilon K \rfloor$, is the most important for the neutral proportions dynamics. Indeed, the neutral proportion in the a -population has already reached its final value at time T_ε^K . Let us describe a coupling of the process N_a^K with two birth and death processes which will be a key argument to control the growing of the population a during the first period of the mutant invasion. To this aim we recall the definition of the stopping time \tilde{T}_ε^K in (3.11) and we define for $\varepsilon < S_{aA}/(2C_{a,A}C_{A,a}/C_{A,A} + C_{a,a})$,

$$(5.1) \quad s_-(\varepsilon) := \frac{S_{aA}}{f_a} - \varepsilon \frac{2C_{a,A}C_{A,a} + C_{a,a}C_{A,A}}{f_a C_{A,A}}, \quad \text{and} \quad s_+(\varepsilon) := \frac{S_{aA}}{f_a} + 2\varepsilon \frac{C_{a,A}C_{A,a}}{f_a C_{A,A}}.$$

Definitions (2.4) and (2.9) ensure that for $t < T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K$,

$$(5.2) \quad f_a(1 - s_+(\varepsilon)) \leq \frac{d_a^K(N^K(t))}{N_a^K(t)} = f_a - S_{aA} + \frac{C_{a,A}}{K}(N_A^K(t) - \bar{n}_A K) + \frac{C_{a,a}}{K}N_a^K(t) \leq f_a(1 - s_-(\varepsilon)),$$

and following Theorem 2 in [5], we can construct on the same probability space the processes Z_ε^- , N^K and Z_ε^+ such that almost surely:

$$(5.3) \quad Z_\varepsilon^-(t) \leq N_a^K(t) \leq Z_\varepsilon^+(t), \quad t < \tilde{T}_\varepsilon^K \wedge T_\varepsilon^K,$$

where for $* \in \{-, +\}$, Z_ε^* is a birth and death process with initial state 1, individual birth rate f_a and individual death rate $f_a(1 - s_*(\varepsilon))$.

6. PROOF OF THEOREM 2 IN THE STRONG RECOMBINATION REGIME

We distinguish the three periods of the selective sweep: (i) rare mutants and resident population size near its equilibrium value, (ii) quasi-deterministic period governed by the dynamical system (2.5), and (iii) A -population extinction. First we prove that at time T_ε^K proportions of b_1 alleles in the populations A and a are close to z_{Ab_1}/z_A . Once the neutral proportions are the same in the two populations, they do not evolve anymore until the end of the sweep.

Lemma 6.1. *There exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$:*

$$\limsup_{K \rightarrow \infty} \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \left| P_{A,b_1}(T_\varepsilon^K) - \frac{z_{Ab_1}}{z_A} \right| \right] \leq c\varepsilon.$$

Proof. First we bound the difference between the neutral proportions in the two populations, $|P_{a,b_1}(t) - P_{A,b_1}(t)|$, then we bound $|P_{A,b_1}(t) - z_{Ab_1}/z_A|$. For sake of simplicity we introduce:

$$(6.1) \quad G(t) := P_{A,b_1}(t) - P_{a,b_1}(t) = \frac{N_{ab_2}(t)N_{Ab_1}(t) - N_{ab_1}(t)N_{Ab_2}(t)}{N_A(t)N_a(t)}, \quad \forall t \geq 0,$$

$$Y(t) = \mathbb{1}_{\{N_A(t) \geq 1, N_a(t) \geq 1\}} G^2(t) e^{r_K(f_A \wedge f_a)t}, \quad \forall t \geq 0.$$

Recalling (3.8) and applying Itô's formula with jumps we get

$$\begin{aligned} Y(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K) &= Y(0) + \hat{M}_{t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} + r_K \int_0^t \mathbb{1}_{s < T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \left((f_A \wedge f_a) - H(s) \right) Y(s) ds \\ &\quad + \int_0^t \mathbb{1}_{s < T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} e^{r_K(f_A \wedge f_a)s} ds \int_{\mathbb{R}_+} \left[\left(\mu_A^K(N, s, \theta) \right)^2 + \left(\mu_a^K(N, s, \theta) \right)^2 \right] \mathbb{1}_{\{N_A(s) \geq 1, N_a(s) \geq 1\}} d\theta, \end{aligned}$$

where \hat{M} is a martingale with zero mean, and H is defined by

$$(6.2) \quad H(s) = \frac{2f_a f_A N_A(s) N_a(s)}{f_A N_A(s) + f_a N_a(s)} \left[\frac{1}{N_A(s) + 1} + \frac{1}{N_a(s) + 1} \right].$$

In particular we can check that for all $s \geq 0$ we have $Y(s)H(s) \geq (f_A \wedge f_a)Y(s)$. Moreover, Equations (3.12) and (3.13) imply the inequalities

$$\begin{aligned} \int_{\mathbb{R}_+} \left(\mu_A^K(N, s \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K, \theta) \right)^2 d\theta &\leq \frac{C(\bar{n}_A + 2\varepsilon C_{A,a}/C_{A,A})}{(\bar{n}_A - 2\varepsilon C_{A,a}/C_{A,A})K}, \\ \int_{\mathbb{R}_+} \left(\mu_a^K(N, s \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K, \theta) \right)^2 d\theta &\leq (k_0 + 1)C(\bar{n}_A + 2\varepsilon C_{A,a}/C_{A,A}) e^{-\frac{S_{aA}}{2(k_0+1)}(s \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K)} \tilde{M}_{s \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K}. \end{aligned}$$

Hence using Cauchy-Schwarz Inequality and (3.14) we obtain

$$\begin{aligned} \mathbb{E}[Y(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K)] &\leq 1 + \frac{C(\bar{n}_A + 2\varepsilon C_{A,a}/C_{A,A})}{r_K(f_A \wedge f_a)(\bar{n}_A - 2\varepsilon C_{A,a}/C_{A,A})K} e^{r_K(f_A \wedge f_a)t} \\ &\quad + \int_0^t (k_0 + 1)C \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right) \mathbb{E} \left[\tilde{M}_{s \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \right] e^{(r_K(f_A \wedge f_a) - \frac{S_{aA}}{2(k_0+1)})s} ds \\ (6.3) \quad &\leq c \left(1 + \frac{1}{Kr_K} e^{r_K(f_A \wedge f_a)t} + e^{(r_K(f_A \wedge f_a) - \frac{S_{aA}}{2(k_0+1)})t} \right), \end{aligned}$$

where c is a finite constant which can be chosen independently of ε and K if ε is small enough and K large enough. Combining Equation (3.1), Cauchy-Schwarz Inequality, and Equations (3.12) and (6.3) we get for every $t \geq 0$,

$$\begin{aligned} & \mathbb{E} \left[\left| P_{A,b_1}(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K) - \frac{\lfloor z_{Ab_1} K \rfloor}{\lfloor z_A K \rfloor} \right| \right] \\ & \leq \mathbb{E} \left[|M_A(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K)| \right] + \frac{r_K f_a \varepsilon}{\bar{n}_A - 2\varepsilon C_{A,a}/C_{A,A}} \int_0^t \mathbb{E} \left[\mathbb{1}_{s < T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} |G(s)| \mathbb{1}_{\{N_A(s) \geq 1, N_a(s) \geq 1\}} \right] ds \\ & \leq \mathbb{E}^{1/2} \left[\langle M_A \rangle_{t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \right] + cr_K \varepsilon \int_0^t \mathbb{E}^{1/2} \left[Y(s \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K) \right] e^{-r_K(f_A \wedge f_a)s/2} ds \\ & \leq c \left(\sqrt{t/K} + \varepsilon \int_0^t \left(r_K e^{-r_K(f_A \wedge f_a)s} + \frac{1}{K} + e^{-S_{aA}s/2(k_0+1)} \right)^{1/2} ds \right), \end{aligned}$$

where c is a finite constant. A simple integration then yields

$$(6.4) \quad \mathbb{E} \left[\left| P_{A,b_1}(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K) - \frac{\lfloor z_{Ab_1} K \rfloor}{\lfloor z_A K \rfloor} \right| \right] \leq c \left(\sqrt{t/K} + \varepsilon \left(\frac{1}{f_A \wedge f_a} + \frac{t}{\sqrt{K}} + \frac{2(k_0+1)}{S_{aA}} \right) \right).$$

But according to Coupling (5.3) and limit (A.11) we have the asymptotic behavior

$$(6.5) \quad \lim_{K \rightarrow \infty} \mathbb{P}(T_\varepsilon^K > 2S_{aA} \log K | T_\varepsilon^K \leq \tilde{T}_\varepsilon^K) = 0,$$

Hence applying (6.4) at time $t = 2S_{aA} \log K$ and using (6.5), we end up the proof. \square

Remark 4. Using the same arguments (Cauchy-Schwarz Inequality, (6.3) and (6.5)) we also get

$$(6.6) \quad \lim_{K \rightarrow \infty} \mathbb{E}[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} |G(T_\varepsilon^K)|] = 0.$$

The following Lemma states that during the second period, the neutral proportion stays constant in the a -population.

Lemma 6.2. *There exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$:*

$$\limsup_{K \rightarrow \infty} \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \left| P_{a,b_1}(T_\varepsilon^K + t_\varepsilon(\frac{N(T_\varepsilon^K)}{K})) - \frac{z_{Ab_1}}{z_A} \right| \right] \leq c\varepsilon.$$

Proof. Let us introduce, for $z \in \mathbb{R}_+^\varepsilon$ and $\varepsilon > 0$

$$(6.7) \quad \Gamma := \left\{ z \in \mathbb{R}_+^\varepsilon, z_A \in \left[\bar{n}_A - 2\varepsilon \frac{C_{A,a}}{C_{A,A}}, \bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}} \right], z_a \in \left[\frac{\varepsilon}{2}, 3\frac{\varepsilon}{2} \right] \right\}, \quad t_\varepsilon := \sup\{t_\varepsilon(z), z \in \Gamma\},$$

where $t_\varepsilon(z)$ has been defined in (4.1). According to Assumption 1, $t_\varepsilon < \infty$, and

$$I(\Gamma, \varepsilon) := \inf_{z \in \Gamma} \inf_{t \leq t_\varepsilon} \{n_A^{(z)}(t), n_a^{(z)}(t)\} > 0,$$

and we can introduce the stopping time

$$L_\varepsilon^K(z) = \inf\{t \geq 0, (N_A^{(z,K)}(t), N_a^{(z,K)}(t)) \notin [I(\Gamma, \varepsilon)K/2, (\bar{n}_A + \bar{n}_a)K]^2\}.$$

Finally, we denote by $(\mathcal{F}_t, t \geq 0)$ the canonical filtration of the process N . Notice that on the event $\{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\}$, $N(T_\varepsilon^K)/K \in \Gamma$. Thus $t_\varepsilon(\frac{N(T_\varepsilon^K)}{K}) \leq t_\varepsilon$. The semi-martingale decomposition stated in (3.1) and the definition of G in (6.1) then twice the Strong Markov property and the Cauchy-Schwarz

Inequality yield:

$$\begin{aligned}
 (6.8) \quad & \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \left| P_{a,b_1} \left(T_\varepsilon^K + t_\varepsilon \left(\frac{N(T_\varepsilon^K)}{K} \right) \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right) \right) - P_{a,b_1} (T_\varepsilon^K) \right| \right] \\
 & \leq \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \mathbb{E} \left[\left| M_a \left(T_\varepsilon^K + t_\varepsilon \left(\frac{N(T_\varepsilon^K)}{K} \right) \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right) \right) - M_a (T_\varepsilon^K) \right| + f_a \int_{T_\varepsilon^K}^{T_\varepsilon^K + t_\varepsilon \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right)} |G(s)| ds \middle| \mathcal{F}_{T_\varepsilon^K} \right] \right] \\
 & \leq \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \left\{ \mathbb{E}^{1/2} \left[\langle M_a \rangle_{T_\varepsilon^K + t_\varepsilon \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right)} - \langle M_a \rangle_{T_\varepsilon^K} \middle| \mathcal{F}_{T_\varepsilon^K} \right] + f_a \sqrt{t_\varepsilon} \mathbb{E}^{1/2} \left[\int_{T_\varepsilon^K}^{T_\varepsilon^K + t_\varepsilon \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right)} G^2(s) ds \middle| \mathcal{F}_{T_\varepsilon^K} \right] \right\} \right].
 \end{aligned}$$

To bound the first term of the right hand side we use Strong Markov property, Equation (3.12) and the definition of L_ε^K in (6). We get

$$(6.9) \quad \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \mathbb{E}^{1/2} \left[\langle M_a \rangle_{T_\varepsilon^K + t_\varepsilon \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right)} - \langle M_a \rangle_{T_\varepsilon^K} \middle| \mathcal{F}_{T_\varepsilon^K} \right] \right] \leq \frac{2t_\varepsilon C(\bar{n}_A + \bar{n}_a)}{I(\Gamma, \varepsilon)K}.$$

For the second term, Strong Markov property and Itô's formula with jumps yield

$$\begin{aligned}
 & \mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \mathbb{E}^{1/2} \left[\int_{T_\varepsilon^K}^{T_\varepsilon^K + t_\varepsilon \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right)} G^2(s) ds \middle| \mathcal{F}_{T_\varepsilon^K} \right] \\
 & \leq \sqrt{t_\varepsilon} \sup_{z \in \Gamma, t \leq t_\varepsilon} \mathbb{E} \left[|G^2(t \wedge L_\varepsilon^K(z)) - G^2(0)| \middle| N(0) = \lfloor zK \rfloor \right] + \mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} |G(T_\varepsilon^K)| \\
 & \leq 2\sqrt{t_\varepsilon} \sum_{\alpha \in \mathcal{A}} \int_0^{t_\varepsilon} \mathbb{E} \left[\mathbb{1}_{s < L_\varepsilon^K(z)} \frac{d}{ds} \langle M_\alpha \rangle_s \right] ds + \mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} |G(T_\varepsilon^K)|
 \end{aligned}$$

Equations (3.12) and (6.6) finally lead to

$$(6.10) \quad \mathbb{E} \left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \mathbb{E}^{1/2} \left[\int_{T_\varepsilon^K}^{T_\varepsilon^K + t_\varepsilon \wedge L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right)} G^2(s) ds \middle| \mathcal{F}_{T_\varepsilon^K} \right] \right] \leq \sqrt{t_\varepsilon} \left(\frac{8t_\varepsilon C(\bar{n}_A + \bar{n}_a)}{I(\Gamma, \varepsilon)K} \right)^{1/2} + o_K(1)$$

where $o_K(1)$ denotes a function of K going to 0 at infinity. Moreover (2.6) ensures that

$$\mathbb{P} \left(T_\varepsilon^K \leq \tilde{T}_\varepsilon^K, L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right) \leq t_\varepsilon \left(\frac{N(T_\varepsilon^K)}{K} \right) \right) \leq \mathbb{P} \left(\frac{N(T_\varepsilon^K)}{K} \in \Theta, L_\varepsilon^K \left(\frac{N(T_\varepsilon^K)}{K} \right) \leq t_\varepsilon \left(\frac{N(T_\varepsilon^K)}{K} \right) \right) \xrightarrow{K \rightarrow \infty} 0,$$

where Θ has been defined in (4.7). Adding Equations (6.8), (6.9), (6.10) and Lemma 6.1, we finally end the proof of Lemma 6.2. \square

Proof of Theorem 2 in the strong recombination regime. Let us focus on the A-population extinction period. We have thanks to Strong Markov property:

$$\begin{aligned}
 (6.11) \quad & \mathbb{P} \left(\mathbb{1}_{N(T_\varepsilon^K + t_\varepsilon(N(T_\varepsilon^K)/K)) \in \Theta} \left| P_{a,b_1}(T_{\text{ext}}) - P_{a,b_1} \left(T_\varepsilon^K + t_\varepsilon \left(\frac{N(T_\varepsilon^K)}{K} \right) \right) \right| > \sqrt{\varepsilon} \right) \\
 & \leq \sup_{z \in \Theta} \mathbb{P} \left(|P_{a,b_1}(T_{\text{ext}}) - P_{a,b_1}(0)| > \sqrt{\varepsilon} \middle| N(0) = \lfloor zK \rfloor \right).
 \end{aligned}$$

But Equation (2.6) yields $\mathbb{P}(N(T_\varepsilon^K + t_\varepsilon(N(T_\varepsilon^K)/K))/K \in \Theta | N(T_\varepsilon^K)/K \in \Gamma) \xrightarrow{K \rightarrow \infty} 1$, and $\{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\} \subset \{N(T_\varepsilon^K)/K \in \Gamma\}$. Adding Equation (A.7) and Lemma 6.2, Triangle inequality allows us to conclude that for ε small enough

$$\limsup_{K \rightarrow \infty} \mathbb{P} \left(\sup \left\{ \left| P_{a,b_1}^K(T_{\text{ext}}) - \frac{zAb_1}{z_A} \right| \right\} > \sqrt{\varepsilon} \middle| \text{Fix}^K \right) \leq c\varepsilon,$$

which is equivalent to the claim of Theorem 2 in the strong recombination regime. \square

7. PROOF OF THEOREM 2 IN THE WEAK RECOMBINATION REGIME

In this section we suppose that Assumptions 3 and 5 hold. We are interested in proving a convergence on the fixation event Fix^K , defined in (2.10). Inequality (A.7) allows us to restrict our attention to the conditional probability measure:

$$(7.1) \quad \hat{\mathbb{P}}(\cdot) = \mathbb{P}(\cdot | T_\varepsilon^K \leq \tilde{T}_\varepsilon^K),$$

where T_ε^K and \tilde{T}_ε^K have been defined in (3.11). $\hat{\mathbb{E}}$ and $\hat{\text{Var}}$ denote the expectation and variance associated with this probability measure. When we do not mention the initial state, we consider that $(N_A(0), N_a(0)) = (\lfloor \bar{n}_A K \rfloor, 1)$. Else, $\mathbb{P}_{(n_A, n_a)}$ (resp. $\hat{\mathbb{P}}_{(n_A, n_a)}$) denotes the probability \mathbb{P} (resp. $\hat{\mathbb{P}}$) when $(N_A(0), N_a(0)) = (n_A, n_a) \in \mathbb{Z}_+^2$.

Here again we distinguish the three periods of the selective sweep, but we will see that the neutral proportion in the a population has already reached its final value at the end of the first period (time T_ε^K). Thus we focus on the first period and are interested in the proportions of neutral alleles b_1 and b_2 in the $[\varepsilon K]$ individuals of the a population at time T_ε^K . The idea is to consider the neutral alleles at time T_ε^K and follow their ancestral lines back until the beginning of the sweep, to know whether they are descended from the first mutant or not. Two kinds of events can happen to a neutral lineage: coalescences and recombinations (see Section 7.1); we will show that we can neglect the coalescences and the occurrence of several recombinations for a lineage during the first period. Therefore, our approximation of the genealogy will be the following: two neutral lineages are independent, and each of them undergoes one recombination with an A -individual during the first period with probability $1 - \exp(-r_K f_a \log K / S_{aA})$. If it has undergone a recombination with an A -individual, it can be an allele b_1 or b_2 . Otherwise it is descended from the first mutant and is an allele b_1 . To get this approximation we follow the line presented by Schweinsberg and Durrett in [25]. In this paper, the authors describe the population dynamics by a variation of Moran model with two loci and recombinations. In their model, the population size is constant and each individual has a constant selective advantage, 0 or s . In our model, the size fluctuates over time and the individual ability to survive depends on the population state. Moreover, the mutant population size is no more Markovian and it is necessary to consider both A and a -population sizes to get a Markov process.

7.1. Coalescence and m-recombination times. Let us introduce the jump times of the stopped Markov process $(N^K(t), t \leq T_\varepsilon^K)$, $0 =: \tau_0^K < \tau_1^K < \dots < \tau_{J^K}^K := T_\varepsilon^K$, where J^K denotes the jump number of N^K between 0 and T_ε^K , and the time of the m -th jump is:

$$\tau_m^K = \inf \left\{ t > \tau_{m-1}^K, N^K(t) \neq N^K(\tau_{m-1}^K) \right\}, \quad 1 \leq m \leq J^K.$$

Let us sample two individuals uniformly at random at time T_ε^K and denote by β_p and β_q their neutral alleles. We want to follow their genealogy and know at each time between 0 and T_ε^K the types (A or a) of the individuals carrying β_p and β_q .

We say that β_p and β_q coalesce at time τ_m^K if they are carried by two different individuals at time τ_m^K and by the same individual at time τ_{m-1}^K . In other words the individual carrying the allele β_p (or β_q) at time τ_m^K is a newborn and has inherited his neutral allele from the individual carrying allele β_q (or β_p) at time τ_{m-1}^K . The jump number at the coalescence time is denoted by

$$TC^K(\beta_p, \beta_q) := \begin{cases} \sup\{m \leq J^K, \beta_p \text{ and } \beta_q \text{ coalesce at time } \tau_m^K\}, & \text{if } \beta_p \text{ and } \beta_q \text{ coalesce} \\ -\infty, & \text{otherwise.} \end{cases}$$

We say that β_p m -recombines at time τ_m^K if the individual carrying the allele β_p at time τ_m^K is a newborn, carries the allele $\alpha \in \mathcal{A}$, and has inherited his allele β_p from an individual carrying allele $\bar{\alpha}$. In other words, a m -recombination is a recombination which modifies the selected allele

connected to the neutral allele in the newborn individual. The jump numbers of the first and second (backward in time) m-recombinations are denoted by:

$$TR_1^K(\beta_p) := \begin{cases} \sup\{m \leq J^K, \beta_p \text{ m-recombines at time } \tau_m^K\}, & \text{if there is at least one m-recombination} \\ -\infty, & \text{otherwise,} \end{cases}$$

$$TR_2^K(\beta_p) := \begin{cases} \sup\{m < TR_1^K(\beta_p), \beta_p \text{ m-recombines at time } \tau_m^K\}, & \text{if there are at least two} \\ -\infty, & \text{m-recombinations} \\ & \text{otherwise.} \end{cases}$$

Let us now focus on the probability for a coalescence to occur conditionally on the state of the process (N_A, N_a) at two successive jump times. We denote by $p_{\alpha_1 \alpha_2}^{c_K}(n)$ the probability that the genealogies of two random neutral alleles associated respectively with alleles α_1 and $\alpha_2 \in \mathcal{A}$ at time τ_m^K coalesce at this time conditionally on $(N_A^K(\tau_{m-1}^K), N_a^K(\tau_{m-1}^K)) = n \in \mathbb{N}^2$ and on the birth of an individual carrying allele $\alpha_1 \in \mathcal{A}$ at time τ_m^K .

Lemma 7.1. *For every $n = (n_A, n_a) \in \mathbb{N}^2$ and $\alpha \in \mathcal{A}$, we have:*

$$(7.2) \quad p_{\alpha\alpha}^{c_K}(n) = \frac{2}{n_\alpha(n_\alpha + 1)} \left(1 - \frac{r_K f_{\bar{a}} n_{\bar{a}}}{f_A n_A + f_a n_a} \right) \quad \text{and} \quad p_{\alpha\bar{a}}^{c_K}(n) = \frac{r_K f_{\bar{a}}}{(n_\alpha + 1)(f_A n_A + f_a n_a)}.$$

Proof. We only state the expression of $p_{\alpha\alpha}^{c_K}(n)$, as the calculations are similar for $p_{\alpha\bar{a}}^{c_K}(n)$. If there is a m-recombination, we cannot have the coalescence of two neutral alleles associated with allele α at time τ_m^K . With probability $1 - r_K f_{\bar{a}} n_{\bar{a}} / (f_A n_A + f_a n_a)$ there is no m-recombination and the parent giving its neutral allele carries the allele α . When there is no m-recombination, two individuals among those, who carry allele α also carry a neutral allele which was in the same individual at time τ_{m-1}^K . We have a probability $2/n_\alpha(n_\alpha + 1)$ to pick this couple of individuals among the $(n_\alpha + 1)$ α -individuals. \square

Remark 5. *A m-recombination for a neutral allele associated with an α allele is a coalescence with an \bar{a} individual. Thus if we denote by $p_\alpha^{r_K}(n)$ the probability that an α -individual, chosen at random at time τ_m^K , is the newborn and underwent a m-recombination at his birth, conditionally on $(N_A^K(\tau_{m-1}^K), N_a^K(\tau_{m-1}^K)) = n \in \mathbb{N}^2$ and on the birth of an individual α at time τ_m^K we get*

$$(7.3) \quad p_\alpha^{r_K}(n) = n_{\bar{a}} p_{\alpha\bar{a}}^{c_K}(n) = \frac{n_{\bar{a}} r_K f_{\bar{a}}}{(n_\alpha + 1)(f_A n_A + f_a n_a)}.$$

Moreover, if we recall the definition of I_ε^K in (3.10), we notice that there exists a finite constant c such that for $k < \lfloor \varepsilon K \rfloor$,

$$(7.4) \quad (1 - c\varepsilon) \frac{r_K}{k+1} \leq \inf_{n_A \in I_\varepsilon^K} p_a^{r_K}(n_A, k) \leq \sup_{n_A \in I_\varepsilon^K} p_a^{r_K}(n_A, k) \leq \frac{r_K}{k+1}.$$

7.2. Jumps of mutant population during the first period. We want to count the number of coalescences and m-recombinations occurring in the lineages of the two randomly chosen neutral alleles β_p and β_q . By definition, these events can only occur at a birth time. Thus we need to study the upcrossing number of the process N_a^K before T_ε^K (Lemma 7.2). It allows us to prove that the probability that a lineage is affected by two m-recombinations or that two lineages coalesce, and then (backward in time) are affected by a m-recombination are negligible (Lemma 7.3). Then we obtain an approximation of the probability that a lineage is affected by a m-recombination with the A-population (Lemma 7.4), and finally we check that two lineages are approximately independent (Equation (7.21)). The last step consists in controlling the neutral proportion in the population A (Lemma 7.5). Indeed it will give us the probability that a neutral allele which has undergone a m-recombination with the A population is a b_1 or a b_2 .

Let us denote by ζ_k^K the jump number of last visit to k before the hitting of $\lfloor \varepsilon K \rfloor$,

$$(7.5) \quad \zeta_k^K := \sup\{m \leq J^K, N_a^K(\tau_m^K) = k\}, \quad 1 \leq k \leq \lfloor \varepsilon K \rfloor.$$

This allows us to introduce for $0 < j \leq k < \lfloor \varepsilon K \rfloor$ the number of upcrossings from k to $k+1$ for the process N_a^K before and after the last visit to j :

$$(7.6) \quad U_{j,k}^{(K,1)} := \#\{m \in \{0, \dots, \zeta_j^K - 1\}, (N_a^K(\tau_m), N_a^K(\tau_{m+1}^K)) = (k, k+1)\},$$

$$(7.7) \quad U_{j,k}^{(K,2)} := \#\{m \in \{\zeta_j^K, \dots, J^K - 1\}, (N_a^K(\tau_m), N_a^K(\tau_{m+1}^K)) = (k, k+1)\}.$$

We also introduce the number of jumps of the A -population size when there are k a -individuals and the total number of upcrossings from k to $k+1$ before T_ε^K :

$$(7.8) \quad H_k^K := \#\{m < J^K, N_a^K(\tau_m) = N_a^K(\tau_{m+1}^K) = k\},$$

$$(7.9) \quad U_k^K := U_{j,k}^{(K,1)} + U_{j,k}^{(K,2)} = \#\{m < J^K, (N_a^K(\tau_m), N_a^K(\tau_{m+1}^K)) = (k, k+1)\}.$$

The next Lemma states moment properties of these jump numbers. Recall Definition (5.1). Then

Lemma 7.2. *There exist two positive and finite constants ε_0 and c such that for $\varepsilon \leq \varepsilon_0$, K large enough and $1 \leq j \leq k < \lfloor \varepsilon K \rfloor$, $\lambda_\varepsilon := (1 - s_-(\varepsilon))^3(1 - s_+(\varepsilon))^{-2} \in (0, 1)$,*

$$(7.10) \quad \hat{\mathbb{E}}[(U_j^K)^2] \leq \frac{2}{s_-^2(\varepsilon)}, \quad \hat{\mathbb{E}}[H_j^K] \leq \frac{24\bar{n}_A K}{s_-^4(\varepsilon)j}, \quad \hat{\mathbb{E}}[(U_{j,k}^{(K,1)})^2] \leq \frac{4\lambda_\varepsilon^{k-j}}{s_-^7(\varepsilon)(1 - s_+(\varepsilon))},$$

$$(7.11) \quad \left| \widehat{\text{Cov}}(U_{j,k}^{(K,2)}, U_j^K) \right| \leq c(\varepsilon + (1 - s_-(\varepsilon))^{k-j}),$$

and

$$(7.12) \quad r_K \left| \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{\hat{\mathbb{E}}[U_k^K]}{k+1} - \frac{f_a \log K}{S_{AA}} \right| \leq c\varepsilon.$$

This Lemma is widely used in Sections 7.3 and 7.4. Indeed, we shall decompose on the possible states of the population when a birth occurs, and apply Equations (7.2) and (7.3) to express the probability of coalescences and m -recombinations at each birth event. The proof of Lemma 7.2 is quite technical and is postponed to Appendix B.

7.3. Negligible events. The next Lemma bounds the probability that two m -recombinations occur in a neutral lineage and the probability that a couple of neutral lineages coalesce and then m -recombine when we consider the genealogy backward in time.

Lemma 7.3. *There exist two positive finite constants c and ε_0 such that for $K \in \mathbb{N}$ and $\varepsilon \leq \varepsilon_0$,*

$$\hat{\mathbb{P}}(TR_2^K(\beta_p) \neq -\infty) \leq \frac{c}{\log K}, \quad \text{and} \quad \hat{\mathbb{P}}(0 \leq TR_1^K(\beta_p) \leq TC^K(\beta_p, \beta_q)) \leq \frac{c}{\log K}.$$

Proof. By definition, the neutral allele β_p is associated with an allele a at time T_ε^K . If there are at least two m -recombinations it implies that there exists a time between 0 and T_ε^K in the genealogy of β_p at which β_p has undergone a m -recombination when it was associated with an allele A . We shall work conditionally on the stopped process $((N_A(\tau_m^K), N_a(\tau_m^K)), m \leq J^K)$ and decompose according to the a -population size when this m -recombination occurs. We get the inclusion:

$$\{TR_2^K(\beta_p) \neq -\infty\} \subset \bigcup_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \bigcup_{m=1}^{J^K} \left\{ TR_2^K(\beta_p) = m, N_a(\tau_{m-1}^K) = N_a(\tau_m^K) = k \right\}.$$

We recall the definition of I_ε^K in (3.10). Thanks to Equations (7.3) and (7.10), we get:

$$\hat{\mathbb{P}}(TR_2^K(\beta_p) \neq -\infty) \leq \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \sup_{n_A \in I_\varepsilon^K} p_A^{r_K}(n_A, k) \hat{\mathbb{E}}[H_k^K] \leq \frac{24r_K \bar{n}_A f_a \varepsilon}{s_-^4(\varepsilon) f_A (\bar{n}_A - 2\varepsilon C_{A,a} / C_{A,A})^2}.$$

The Assumption 5 on weak recombination completes the proof of the first inequality in Lemma 7.3. The proof of the second one is divided in two steps, presented after introducing the notations

$$(\alpha\beta_p)_m := \{\text{the neutral allele } \beta_p \text{ is associated with the allele } \alpha \text{ at time } \tau_m^K\}, \quad \alpha \in \mathcal{A}, \quad m \leq J^K,$$

$$(\alpha_1\beta_p, \alpha_2\beta_q)_m := (\alpha_1\beta_p)_m \cap (\alpha_2\beta_q)_m, \quad (\alpha_1, \alpha_2) \in \mathcal{A}^2, \quad m \leq J^K.$$

First step: We show that the probability that β_p is associated with an allele A at the coalescence time is negligible. We first recall the inclusion,

$$\{TC^K(\beta_p, \beta_q) \neq -\infty, (A\beta_p)_{TC^K(\beta_p, \beta_q)}\} \subset \bigcup_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \bigcup_{m=1}^{J^K} \{TC^K(\beta_p, \beta_q) = m, N_a(\tau_{m-1}^K) = k, (A\beta_p)_m\},$$

and decompose on the possible selected alleles associated with β_q and on the type of the newborn at coalescence time. Using Lemma 7.1, Equation (7.10) and the inequality $r_K \leq 1$, we get

$$(7.13) \quad \hat{\mathbb{P}}(TC^K(\beta_p, \beta_q) \neq -\infty, (A\beta_p)_{TC^K(\beta_p, \beta_q)}) \\ \leq \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \left[\sup_{n_A \in I_\varepsilon^K} p_{AA}^{c_K}(n_A, k) + \sup_{n_A \in I_\varepsilon^K} p_{Aa}^{c_K}(n_A, k) \right] \hat{\mathbb{E}}[H_k^K] + \sup_{n_A \in I_\varepsilon^K} p_{aA}^{c_K}(n_A, k) \hat{\mathbb{E}}[U_k^K] \leq \frac{c}{K} \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{1}{k},$$

for a finite c , which is of order $\log K/K$.

Second step: Then, we focus on the case where β_p and β_q are associated with an allele a at coalescence time. The inclusion

$$\{N_a(\tau_{TC^K(\beta_p, \beta_q)-1}^K) = k, (a\beta_p, a\beta_q)_{TC^K(\beta_p, \beta_q)}\} \subset \bigcup_{m=1}^{J^K} \{TC^K(\beta_p, \beta_q) = m, N_a(\tau_{m-1}^K) = k, (a\beta_p, a\beta_q)_m\},$$

and Equations (7.2) and (7.10) yield for every $k \in \{1, \dots, \lfloor \varepsilon K \rfloor - 1\}$:

$$\hat{\mathbb{P}}(N_a(\tau_{TC^K(\beta_p, \beta_q)-1}^K) = k, (a\beta_p, a\beta_q)_{TC^K(\beta_p, \beta_q)}) \leq \sup_{n_A \in I_\varepsilon^K} p_{aa}^{c_K}(n_A, k) \hat{\mathbb{E}}[U_k^K] \leq \frac{4}{s_-^2(\varepsilon)k(k+1)}.$$

If β_p and β_q coalesce then undergo their first m -recombination when we look backward in time, and if the a -population has the size k at the coalescence time, it implies that the m -recombination occurs before $\tau_{\zeta_k^K}$ when we look forward in time. For $k, l < \lfloor \varepsilon K \rfloor$,

$$\hat{\mathbb{P}}\left(N_a\left(\tau_{TR_1^K(\beta_p)}^K\right) = l, 0 \leq TR_1^K(\beta_p) \leq TC^K(\beta_p, \beta_q) \mid N_a\left(\tau_{TC^K(\beta_p, \beta_q)-1}^K\right) = k, (a\beta_p, a\beta_q)_{TC^K(\beta_p, \beta_q)}\right) \\ \leq \sup_{n_A \in I_\varepsilon^K} p_a^{r_K}(n_A, l) \left(\mathbb{1}_{k>l} \hat{\mathbb{E}}[U_l^K] + \mathbb{1}_{k \leq l} \hat{\mathbb{E}}[U_{k,l}^{(K,1)}] \right) \leq \frac{2r_K}{(l+1)s_-^2(\varepsilon)} \left(\mathbb{1}_{k>l} + \frac{2\mathbb{1}_{k \leq l} \lambda_\varepsilon^{l-k}}{s_-^2(\varepsilon)(1-s_+(\varepsilon))} \right),$$

where the last inequality is a consequence of (7.4) and (7.10). The two last equations finally yield the existence of a finite c such that for every $K \in \mathbb{N}$:

$$\hat{\mathbb{P}}(0 \leq TR_1^K(\beta_p) \leq TC^K(\beta_p, \beta_q), (a\beta_p, a\beta_q)_{TC^K(\beta_p, \beta_q)}) \leq cr_K \sum_{k,l=1}^{\lfloor \varepsilon K \rfloor} \frac{\mathbb{1}_{k>l} + \mathbb{1}_{k \leq l} \lambda_\varepsilon^{l-k}}{k(k+1)(l+1)} \leq cr_K,$$

which ends up the proof of Lemma 7.3 with Assumption 5. \square

7.4. Probability to be descended from the first mutant. We want to estimate the probability for the neutral lineage of β_p to undergo no m -recombination.

Lemma 7.4. *There exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$:*

$$\limsup_{K \rightarrow \infty} \left| \hat{\mathbb{P}}(TR_1^K(\beta_p) = -\infty) - \exp\left(-\frac{far_K \log K}{S_{aA}}\right) \right| \leq c\varepsilon.$$

Proof. We introduce ρ_m^K , the conditional probability that the neutral lineage of β_p m-recombines at time τ_m^K , given $(N_A(\tau_n^K), N_a(\tau_n^K), n \leq J^K)$ and given that it has not m-recombined during the time interval $]\tau_m^K, T_\varepsilon^K]$. The last condition implies that β_p is associated with an allele a at time τ_m^K .

$$\rho_m^K := \mathbb{1}_{\{N_A^K(\tau_m^K) - N_a^K(\tau_{m-1}^K) = 1\}} p_a^{r_K}(N_A^K(\tau_{m-1}^K), N_a^K(\tau_{m-1}^K)).$$

We also introduce η^K , the sum of these conditional probabilities for $1 \leq m \leq J^K$:

$$\eta^K := \sum_{m=1}^{J^K} \rho_m^K.$$

We want to give a rigorous meaning to the sequence of equivalents:

$$\hat{\mathbb{P}}\left(TR_1^K(\beta_p) = -\infty \mid (N_A(\tau_m^K), N_a(\tau_m^K))_{m \leq J^K}\right) = \prod_{m=1}^{J^K} (1 - \rho_m^K) \sim \prod_{m=1}^{J^K} \exp(-\rho_m^K) \sim \exp(-\mathbb{E}[\eta^K]),$$

when K goes to infinity. Jensen and Triangle inequalities and the Mean Value Theorem imply

$$(7.14) \quad \hat{\mathbb{E}}\left|\hat{\mathbb{P}}\left(TR_1^K(\beta_p) = -\infty \mid (N_A(\tau_m^K), N_a(\tau_m^K))_{m \leq J^K}\right) - e^{-\frac{far_K \log K}{s_{aA}}}\right| \leq \\ \hat{\mathbb{E}}\left|\hat{\mathbb{P}}\left(TR_1^K(\beta_p) = -\infty \mid (N_A(\tau_m^K), N_a(\tau_m^K))_{m \leq J^K}\right) - e^{-\eta^K}\right| + \left|e^{-\hat{\mathbb{E}}\eta^K} - e^{-\frac{far_K \log K}{s_{aA}}}\right| + \hat{\mathbb{E}}\left|\eta^K - \hat{\mathbb{E}}\eta^K\right|.$$

We aim to bound the right hand side of (7.14). The bounding of the first term follows the method developed in Lemma 3.6 in [25]. We refer to this proof, and get

$$(7.15) \quad \hat{\mathbb{E}}\left|\hat{\mathbb{P}}\left(TR_1^K(\beta_p) = -\infty \mid (N_A(\tau_m^K), N_a(\tau_m^K))_{m \leq J^K}\right) - e^{-\eta^K}\right| \leq \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \sup_{n_A \in I_\varepsilon^K} \left(p_a^{r_K}(n_A, k)\right)^2 \hat{\mathbb{E}}[U_k^K] \leq \frac{\pi^2 r_K^2}{6s_-(\varepsilon)},$$

where I_ε^K has been defined in (3.10) and the last inequality follows from (7.4) and (7.10). To bound the second term, we need to estimate $\hat{\mathbb{E}}[\eta^K]$. Inequality (7.4) implies

$$(7.16) \quad (1 - c\varepsilon)r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{U_k^K}{k+1} \leq \eta^K \leq r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{U_k^K}{k+1}.$$

Adding (7.12) we get that for ε small enough,

$$(7.17) \quad \limsup_{K \rightarrow \infty} \left| \exp(-\hat{\mathbb{E}}[\eta^K]) - \exp\left(-\frac{r_K fa \log K}{s_{aA}}\right) \right| \leq c\varepsilon.$$

The bounding of the last term of (7.14) requires a fine study of dependences between upcrossing numbers before and after the last visit to a given integer by the mutant population size. In particular, we widely use Equation (7.11). We observe that $\hat{\mathbb{E}}[|\eta^K - \hat{\mathbb{E}}\eta^K|] \leq (\hat{\text{Var}} \eta^K)^{1/2}$, but the variance of η^K is quite complicated to study and according to Assumption 5 and Equations (7.16) and (7.10),

$$(7.18) \quad \left| \hat{\text{Var}} \eta^K - \hat{\text{Var}} \left(r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{U_k^K}{k+1} \right) \right| \leq c\varepsilon \hat{\mathbb{E}} \left[\left(r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{U_k^K}{k+1} \right)^2 \right] \\ \leq r_K^2 \sum_{k,l=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{\hat{\mathbb{E}}[(U_k^K)^2] + \hat{\mathbb{E}}[(U_l^K)^2]}{(k+1)(l+1)} \leq c\varepsilon,$$

for a finite c . Let $k \leq l \leq \lfloor \varepsilon K \rfloor$, and recall that by definition, $U_l^K = U_{k,l}^{(K,1)} + U_{k,l}^{(K,2)}$:

$$(7.19) \quad |\hat{\text{Cov}}(U_k^K, U_l^K)| \leq \left(\hat{\mathbb{E}}[(U_k^K)^2] \hat{\mathbb{E}}[(U_{k,l}^{(K,1)})^2] \right)^{1/2} + |\hat{\text{Cov}}(U_k^K, U_{k,l}^{(K,2)})|.$$

Applying Inequalities (7.10) and (7.11) and noticing that $(1 - s_-(\varepsilon)) < \lambda_\varepsilon^{1/2} < 1$ (see proof of Lemma 7.2 in Appendix B) lead to

$$|\hat{\text{Cov}}(U_k^K, U_l^K)| \leq c(\lambda_\varepsilon^{(l-k)/2} + \varepsilon + (1 - s_-(\varepsilon))^{l-k}) \leq c(\lambda_\varepsilon^{(l-k)/2} + \varepsilon)$$

for a finite c . We finally get:

$$(7.20) \quad \begin{aligned} \widehat{\text{Var}}\left(r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{U_k^K}{k+1}\right) &\leq 2r_K^2 \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{1}{k+1} \sum_{l=k}^{\lfloor \varepsilon K \rfloor - 1} \frac{\widehat{\text{Cov}}(U_k^K, U_l^K)}{l+1} \\ &\leq cr_K^2 \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{1}{k+1} \sum_{l=k}^{\lfloor \varepsilon K \rfloor - 1} \frac{\lambda_\varepsilon^{(l-k)/2} + \varepsilon}{l+1} \leq cr_K^2 \varepsilon \log^2 K, \end{aligned}$$

where we used (7.19) for the second inequality. Applying Jensen's Inequality to the left hand side of (7.14) and adding Equations (7.15), (7.17), (7.18) and (7.20) complete the proof of Lemma 7.4. \square

We finally focus on the dependence between genealogies of β_p and β_q . Following [25] pp. 1622 to 1624 in the case $J = 1$, we can prove that for d in $\{0, 1, 2\}$,

$$(7.21) \quad \limsup_{K \rightarrow \infty} \left| \widehat{\mathbb{P}}(\mathbb{1}_{TR_1^K(\beta_p) \geq 0} + \mathbb{1}_{TR_1^K(\beta_q) \geq 0} = d) - \binom{2}{d} \left(1 - e^{-\frac{far_K \log K}{S_{aA}}}\right)^d e^{-\frac{(2-d)far_K \log K}{S_{aA}}} \right| \leq c\varepsilon.$$

In particular we use here the weak dependence between two neutral lineages stated in Lemma 7.3 and the probability to descend from the first mutant for β_p and β_q obtained in Lemma 7.4.

7.5. Neutral proportion at time T_ε^K . According to Lemma 7.3 and Equation (7.21), it is enough to distinguish two cases for the randomly chosen neutral allele β_p : either its lineage has undergone one m-recombination, or no m-recombination. In the second case, β_p is a b_1 . In the first one, the probability that β_p is a b_1 depends on the neutral proportion in the A population at the coalescence time. We now state that this proportion stays nearly constant during the first period.

Lemma 7.5. *There exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$,*

$$(7.22) \quad \limsup_{K \rightarrow \infty} \mathbb{P}\left(\sup_{t \leq T_\varepsilon^K} \left| P_{A, b_1}^K(t) - \frac{z_{Ab_1}}{z_A} \right| > \sqrt{\varepsilon}, T_\varepsilon^K < \infty\right) \leq c\varepsilon.$$

The proof of Lemma 7.5 is quite technical and is postponed to Appendix B. This allows us to focus on the neutral proportions in the a -population at time T_ε^K .

Lemma 7.6. *There exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$,*

$$\limsup_{K \rightarrow \infty} \widehat{\mathbb{P}}\left(\left| P_{a, b_2}^K(T_\varepsilon^K) - (1 - \exp(-far_K \log K / S_{aA})) z_{Ab_2} / z_A \right| > \varepsilon^{1/3}\right) \leq c\varepsilon^{1/6}.$$

Proof. We denote by $(\beta_i, i \leq \lfloor \varepsilon K \rfloor)$ the neutral alleles carried by the a -individuals at time T_ε^K and introduce the notation

$$A_2^K(i) := \{\beta_i \text{ has undergone exactly one m-recombination and is an allele } b_2\}.$$

If β_i is a b_2 , either its genealogy has undergone one m-recombination with an individual Ab_2 , either it has undergone more than two m-recombinations. Thus

$$0 \leq N_{ab_2}^K(T_\varepsilon^K) - \sum_{i=1}^{\lfloor \varepsilon K \rfloor} \mathbb{1}_{A_2^K(i)} \leq \sum_{i=1}^{\lfloor \varepsilon K \rfloor} \mathbb{1}_{\{TR_2^K(\beta_i) \neq -\infty\}}.$$

Moreover, the probability of $A_2^K(i)$ depends on the neutral proportions in the A -population when β_i m-recombines. For $i \leq \lfloor \varepsilon K \rfloor$,

$$(7.23) \quad \left| \widehat{\mathbb{P}}\left(A_2^K(i) \mid TR_1^K(\beta_i) \geq 0, TR_2^K(\beta_i) = -\infty, \sup_{t \leq T_\varepsilon^K} \left| P_{A, b_1}^K(t) - \frac{z_{Ab_1}}{z_A} \right| \leq \sqrt{\varepsilon}\right) - \left(1 - \frac{z_{Ab_1}}{z_A}\right) \right| \leq \sqrt{\varepsilon}.$$

Lemma 7.5 and Equation (A.6) ensure that $\limsup_{K \rightarrow \infty} \widehat{\mathbb{P}}\left(\sup_{t \leq T_\varepsilon^K} |P_{A, b_1}^K(t) - z_{Ab_1} / z_A| > \sqrt{\varepsilon}\right) \leq c\varepsilon$, and Lemmas 7.3 and 7.4 ensure that $|\widehat{\mathbb{P}}(TR_1^K(\beta_i) \geq 0, TR_2^K(\beta_i) = -\infty) - (1 - \exp(-far_K \log K / S_{aA}))| \leq c\varepsilon$. It yields:

$$\left| \widehat{\mathbb{P}}\left(TR_1^K(\beta_i) \geq 0, TR_2^K(\beta_i) = -\infty, \sup_{t \leq T_\varepsilon^K} \left| P_{A, b_1}^K(t) - \frac{z_{Ab_1}}{z_A} \right| \leq \sqrt{\varepsilon}\right) - \left(1 - \exp\left(-\frac{far_K \log K}{S_{aA}}\right)\right) \right| \leq c\sqrt{\varepsilon}$$

for a finite c . Adding (7.23) we get:

$$(7.24) \quad \limsup_{K \rightarrow \infty} \left| \hat{\mathbb{E}}[P_{a,b_2}^K(T_\varepsilon^K)] - (1 - \exp(-f_a r_K \log K / S_{aA})) (1 - z_{Ab_1} / z_A) \right| \leq c\sqrt{\varepsilon}.$$

In the same way, using the weak dependence between lineages stated in (7.21), we prove that $\limsup_{K \rightarrow \infty} |\hat{\mathbb{E}}[(P_{a,b_2}^K)^2](T_\varepsilon^K) - (1 - e^{-f_a r_K \log K / S_{aA}})^2 (1 - z_{Ab_1} / z_A)^2| \leq c\sqrt{\varepsilon}$. This implies, adding (7.24) that $\limsup_{K \rightarrow \infty} \hat{\text{Var}}(P_{a,b_2}^K(T_\varepsilon^K)) \leq c\sqrt{\varepsilon}$. We end the proof by using Chebyshev's inequality. \square

7.6. Second and third periods. Now we prove that during the second period the neutral proportion in the a population remains nearly constant. This is due to the short duration of this period, which does not go to infinity with the carrying capacity K .

Lemma 7.7. *There exist two positive finite constants c and ε_0 such that for $\varepsilon \leq \varepsilon_0$,*

$$(7.25) \quad \limsup_{K \rightarrow \infty} \hat{\mathbb{P}}\left(\left|P_{a,b_1}^K(T_{\text{ext}}^K) - P_{a,b_1}^K(T_\varepsilon^K)\right| > \varepsilon^{1/3}\right) \leq c\varepsilon^{1/3}.$$

Proof. Let us introduce the stopping time U_ε^K :

$$U_\varepsilon^K := \inf\left\{t \geq 0, \sup_{t \leq t_\varepsilon, \alpha \in \mathcal{A}} |N_\alpha(T_\varepsilon^K + t) / K - n_\alpha^{(N(T_\varepsilon^K))}(t)| > \varepsilon^3\right\}.$$

Recall the definition of t_ε in (6.7) and that $(\mathcal{F}_t, t \geq 0)$ denotes the canonical filtration of N . Strong Markov property, Doob's Maximal Inequality and Equation (3.12) yield:

$$\begin{aligned} & \mathbb{P}\left(T_\varepsilon^K \leq \tilde{T}_\varepsilon^K, \sup_{t \leq t_\varepsilon} |M_a^K(T_\varepsilon^K + t \wedge U_\varepsilon^K) - M_a^K(T_\varepsilon^K)| > \sqrt{\varepsilon}\right) \\ &= \mathbb{E}\left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \mathbb{P}\left(\sup_{t \leq t_\varepsilon} |M_a^K(T_\varepsilon^K + t \wedge U_\varepsilon^K) - M_a^K(T_\varepsilon^K)| > \sqrt{\varepsilon} \mid \mathcal{F}_{T_\varepsilon^K}\right)\right] \\ &\leq \frac{1}{\varepsilon} \mathbb{E}\left[\mathbb{1}_{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K} \left(\langle M_a \rangle_{T_\varepsilon^K + t \wedge U_\varepsilon^K} - \langle M_a \rangle_{T_\varepsilon^K}\right)\right] \leq \frac{c(\varepsilon)t_\varepsilon}{\varepsilon K}, \end{aligned}$$

where $c(\varepsilon)$ is finite. But according to Equation (2.6) with $\delta = \varepsilon^3$, $\limsup_{K \rightarrow \infty} \mathbb{P}(U_\varepsilon^K < t_\varepsilon \mid T_\varepsilon^K \leq \tilde{T}_\varepsilon^K) = 0$. Moreover, Equations (3.1) and (3.6) imply for every $t \geq 0$

$$\sup_{t \leq t_\varepsilon} |P_{a,b_1}^K(T_\varepsilon^K + t) - P_{a,b_1}^K(T_\varepsilon^K)| \leq \sup_{t \leq t_\varepsilon} |M_a^K(T_\varepsilon^K + t) - M_a^K(T_\varepsilon^K)| + r_K t_\varepsilon f_a.$$

As r_K goes to 0 under Assumption 5, we finally get:

$$(7.26) \quad \limsup_{K \rightarrow \infty} \mathbb{P}\left(\sup_{t \leq t_\varepsilon} |P_{a,b_1}^K(T_\varepsilon^K + t) - P_{a,b_1}^K(T_\varepsilon^K)| > \sqrt{\varepsilon}, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\right) = 0.$$

Adding Lemma 4.3 ends the proof of Lemma 7.7 \square

7.7. End of the proof of Theorem 2 in the weak recombination regime. Thanks to Lemmas 7.6 and 7.7 we get that for ε small enough,

$$\limsup_{K \rightarrow \infty} \hat{\mathbb{P}}\left(\left|P_{a,b_2}^K(T_{\text{ext}}^K) - (1 - \exp(-f_a r_K \log K / S_{aA})) z_{Ab_2} / z_A\right| > 2\varepsilon^{1/3}\right) \leq c\varepsilon^{1/6}.$$

Moreover, (A.7) ensures that $\liminf_{K \rightarrow \infty} \mathbb{P}(T_\varepsilon^K \leq \tilde{T}_\varepsilon^K \mid \text{Fix}^K) \geq 1 - c\varepsilon$, which implies

$$\limsup_{K \rightarrow \infty} \mathbb{P}\left(\mathbb{1}_{\text{Fix}^K} \left|P_{a,b_2}^K(T_{\text{ext}}^K) - (1 - \exp(-f_a r_K \log K / S_{aA})) z_{Ab_2} / z_A\right| > 2\varepsilon^{1/3}\right) \leq c\varepsilon^{1/6}.$$

This is equivalent to the convergence in probability and ends the proof of Theorem 2.

APPENDIX A. TECHNICAL RESULTS

We first present some results stated in [5]. We recall Definitions (2.9), (2.10), (4.7), (4.8), (3.11) and (6.7) and that the notation \cdot^K refers to processes that satisfy Assumption 3. Proposition 2 is a direct consequence of Equations (42), (71)-(72) and (74) in [5]:

Proposition 2. *There exist two positive finite constants $M_1 < \infty$ and ε_0 such that for every $\varepsilon \leq \varepsilon_0$*

$$(A.1) \quad \lim_{K \rightarrow \infty} \mathbb{P} \left(\left| N_a^K(T_{ext}^K) - K\bar{n}_a \right| > \varepsilon K \mid \text{Fix}^K \right) = 0, \quad \text{and} \quad \limsup_{K \rightarrow \infty} \left| \mathbb{P}(T_\varepsilon^K < \infty) - \frac{S_{aA}}{f_a} \right| \leq M_1 \varepsilon.$$

Moreover there exists $M_2 > 0$ such that for every $\varepsilon \leq \varepsilon_0$, the probability of the event

$$(A.2) \quad F_\varepsilon^K = \{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K, N_A^K(T_\varepsilon^K + t_\varepsilon) < \varepsilon^2 K/2, |N_a^K(T_\varepsilon^K + t_\varepsilon) - \bar{n}_a K| < \varepsilon K/2\}$$

satisfies

$$(A.3) \quad \liminf_{K \rightarrow \infty} \mathbb{P}(F_\varepsilon^K) \geq S_{aA}/f_a - M_2 \varepsilon,$$

and if $z \in \Theta$, then there exist $V > 0$ and $c < \infty$ such that:

$$(A.4) \quad \liminf_{K \rightarrow \infty} \mathbb{P}(S_\varepsilon^K(z) > e^{VK}) \geq 1 - c\varepsilon.$$

Remark 6. *From Equation (A.1), we deduce that for $\varepsilon < \bar{n}_a/2$*

$$(A.5) \quad \limsup_{K \rightarrow \infty} \mathbb{P}(T_\varepsilon^K = \infty \mid \text{Fix}^K) \leq \lim_{K \rightarrow \infty} \mathbb{P} \left(|N_a^K(T_{ext}^K) - K\bar{n}_a| > \varepsilon K \mid \text{Fix}^K \right) = 0,$$

and adding Equation (A.3) we also obtain the existence of a finite c such that for ε small enough,

$$(A.6) \quad \mathbb{P}(T_\varepsilon^K < \infty, T_\varepsilon^K > \tilde{T}_\varepsilon^K) \leq c\varepsilon.$$

Let us notice that \tilde{T}_ε^K is almost surely finite. Hence $\{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K, T_\varepsilon^K < \infty\} = \{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\}$ a.s. and adding (A.5) and (A.6), we get the existence of a finite c such that

$$(A.7) \quad \liminf_{K \rightarrow \infty} \mathbb{P}(T_\varepsilon^K \leq \tilde{T}_\varepsilon^K \mid \text{Fix}^K) \geq 1 - c\varepsilon.$$

We also recall some results on birth and death processes which proofs can be found in Lemma 3.1 in [25] and in [1] p 109 and 112.

Proposition 3. *Let $Z = (Z_t)_{t \geq 0}$ be a birth a death process with birth and death rates b and d . For $i \in \mathbb{Z}^+$ we define $T_i = \inf\{t \geq 0, Z_t = i\}$ and \mathbb{P}_i (resp. \mathbb{E}_i) the law (resp. expectation) of Z when $Z_0 = i$. Then*

- For $i \in \mathbb{N}$ and $t \geq 0$,

$$(A.8) \quad \mathbb{E}_i[Z_t] = i e^{(b-d)t}.$$

- For $(i, j, k) \in \mathbb{Z}_+^3$ such that $j \in (i, k)$,

$$(A.9) \quad \mathbb{P}_j(T_k < T_i) = \frac{1 - (d/b)^{j-i}}{1 - (d/b)^{k-i}}.$$

- If $0 < d < b$, for every $i \in \mathbb{Z}_+$ and $t \geq 0$,

$$(A.10) \quad \mathbb{P}_i(T_0 \leq t) = \left(\frac{d(1 - e^{(d-b)t})}{b - d e^{(d-b)t}} \right)^i.$$

- If $0 < d < b$, on the non-extinction event of Z , which has a probability $1 - d/b$, the following convergence holds:

$$(A.11) \quad T_N / \log N \xrightarrow{N \rightarrow \infty} (1 - d/b)^{-1}, \text{ a.s.}$$

Let us now state two technical results. The first one is necessary to get Lemma 7.5. The proof is left to the reader. The second one helps us to control upcrossing numbers of the process N_a^K before reaching the size $\lfloor \varepsilon K \rfloor$. We introduce for $0 < s_1, s_2 < 1$ and $0 \leq j, k < \lfloor \varepsilon K \rfloor$, the quantities

$$(A.12) \quad q_{j,k}^{(s_1, s_2)} := \frac{\mathbb{P}_{k+1}^{(s_1)}(\tau_{\varepsilon K} < \tau_k)}{\mathbb{P}_{k+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_j)} = \frac{s_1}{1 - (1 - s_1)^{\lfloor \varepsilon K \rfloor - k}} \frac{1 - (1 - s_2)^{\lfloor \varepsilon K \rfloor - j}}{1 - (1 - s_2)^{k+1-j}}, \quad 0 < s_1, s_2 < 1,$$

whose expressions are direct consequences of (A.9), and get

Lemma A.1. • *Let $\mathcal{D}, \mathcal{C}, \mathcal{F}, \mathcal{G}$ and \mathcal{H} be measurable events such that $\mathcal{D} \subset \mathcal{C} \subset \mathcal{F}$. Then*

$$\mathbb{P}(\mathcal{H} \cap (\mathcal{C} \cup \mathcal{G})) - \mathbb{P}(\mathcal{H} \cap \mathcal{D}) \leq \mathbb{P}(\mathcal{F}) - \mathbb{P}(\mathcal{D}) + \mathbb{P}(\mathcal{G} \setminus \mathcal{C}).$$

• *For $a \in]0, 1/2[$, $(s_1, s_2) \in [a, 1 - a]^2$, $0 \leq j \leq k < l < \lfloor \varepsilon K \rfloor$,*

$$(A.13) \quad q_{0,k}^{(s_1 \wedge s_2, s_1 \vee s_2)} \geq s_1 \wedge s_2 \quad \text{and} \quad \left| \frac{1}{q_{k,l}^{(s_1, s_2)}} - \frac{1}{q_{j,l}^{(s_2, s_1)}} \right| \leq \frac{2(1 + 1/s_2)}{a^2 |\log(1 - a)| e} |s_2 - s_1| + \frac{(1 - s_2)^{l+1-k}}{s_2^3}.$$

Proof. The first part of (A.13) is a direct consequence of Definition (A.12). Let a be in $]0, 1/2[$ and consider functions $f_{\alpha, \beta}: x \mapsto (1 - x^\alpha)/(1 - x^\beta)$, $(\alpha, \beta) \in \mathbb{N}^2$, $x \in [a, 1 - a]$. Then for $x \in [a, 1 - a]$,

$$(A.14) \quad \|f'_{\alpha, \beta}\|_\infty \leq 2(ea^2 |\ln(1 - a)|)^{-1}.$$

From Equation (A.9), we get for $0 < s < 1$ and $0 \leq j \leq k < \lfloor \varepsilon K \rfloor$,

$$(A.15) \quad \left| \frac{\mathbb{P}_{l+1}^{(s)}(\tau_{\lfloor \varepsilon K \rfloor} < \tau_k)}{\mathbb{P}_{l+1}^{(s)}(\tau_{\lfloor \varepsilon K \rfloor} < \tau_j)} - \frac{\mathbb{P}_{l+1}^{(s)}(\tau_{\lfloor \varepsilon K \rfloor} < \tau_l)}{\mathbb{P}_{l+1}^{(s)}(\tau_{\lfloor \varepsilon K \rfloor} < \tau_l)} \right| = \frac{(1 - (1 - s)^{k-j})((1 - s)^{l+1-k} - (1 - s)^{\lfloor \varepsilon K \rfloor - k})}{(1 - (1 - s)^{\lfloor \varepsilon K \rfloor - k})(1 - (1 - s)^{\lfloor \varepsilon K \rfloor - j})} \leq (1 - s)^{l+1-k} s^{-2}.$$

Triangle Inequality leads to:

$$\begin{aligned} \left| \frac{1}{q_{k,l}^{(s_1, s_2)}} - \frac{1}{q_{j,l}^{(s_2, s_1)}} \right| &= \left| \frac{\mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_k)}{\mathbb{P}_{l+1}^{(s_1)}(\tau_{\varepsilon K} < \tau_l)} - \frac{\mathbb{P}_{l+1}^{(s_1)}(\tau_{\varepsilon K} < \tau_j)}{\mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_l)} \right| \\ &\leq \left| \frac{1}{\mathbb{P}_{l+1}^{(s_1)}(\tau_{\varepsilon K} < \tau_l)} - \frac{1}{\mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_l)} \right| \mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_k) \\ &\quad + \frac{1}{\mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_l)} \left| \mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_k) - \mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_j) \right| \\ &\quad + \frac{1}{\mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_l)} \left| \mathbb{P}_{l+1}^{(s_1)}(\tau_{\varepsilon K} < \tau_j) - \mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_j) \right|. \end{aligned}$$

Noticing that $\mathbb{P}_{l+1}^{(s_2)}(\tau_{\varepsilon K} < \tau_l) \geq \mathbb{P}_{l+1}^{(s_2)}(\tau_\infty < \tau_l) = \mathbb{P}_1^{(s_2)}(\tau_\infty < \tau_0) = s_2$, and using (A.15) and the Mean Value Theorem with (A.14), we get the second part of (A.13). \square

APPENDIX B. PROOFS OF LEMMAS 7.2 AND 7.5

Proof of Lemma 7.2. In the whole proof, the integer n_A denotes the state of N_A^K at some time smaller than \tilde{T}_ε^K and thus belongs to I_ε^K which has been defined in (3.10). We introduce for $u \in \mathbb{R}_+$ the hitting time of $\lfloor u \rfloor$ by the process N_a^K :

$$(B.1) \quad \sigma_u^K := \inf\{t \geq 0, N_a^K(t) = \lfloor u \rfloor\}.$$

Moreover, for $0 < s < 1$, if $\tilde{Z}^{(s)}$ denotes a random walk where up jumps occur with probability $1/(2 - s)$ and down jumps with probability $(1 - s)/(2 - s)$. We denote by $\mathbb{P}_i^{(s)}$ the law of $\tilde{Z}^{(s)}$ when the initial state is $i \in \mathbb{N}$ and introduce for every $a \in \mathbb{Z}_+$ the stopping time

$$(B.2) \quad \tau_a := \inf\{n \in \mathbb{Z}_+, \tilde{Z}_n^{(s)} = a\}.$$

The event $\{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\}$ has two implications on the law of N_a^K on the time interval $[0, T_\varepsilon^K]$:

- (1) Firstly, as \tilde{T}_ε^K is almost surely finite, $\{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\}$ implies that T_ε^K is almost surely finite.
 (2) Secondly it implies that the law ℓ_ε^K of the jump process N_a^K belongs to a family of laws $\mathcal{L}_\varepsilon^K$ defined as follows: $\ell_\varepsilon^K \in \mathcal{L}_\varepsilon^K$ if and only if for every $t \leq T_\varepsilon^K$,

$$\mathbb{P}\left(N_a^K(t+dt) = N_a^K(t) + 1 \mid N_a^K(t)\right) = f_a N_a^K(t) dt + o(dt),$$

$$\left| \mathbb{P}\left(N_a^K(t+dt) = N_a^K(t) - 1 \mid N_a^K(t)\right) - \left(f_a - S_{aA} + \frac{C_{a,a} N_a^K(t)}{K}\right) N_a^K(t) dt \right| \leq 2\varepsilon \frac{C_{A,a} C_{a,A}}{C_{A,A}} N_a^K(t) dt + o(dt).$$

Let (i, j, k) be in \mathbb{Z}_+^3 with $j < k < \lfloor \varepsilon K \rfloor$. During $]\zeta_j^K, T_\varepsilon^K[$ the process N_a necessarily jumps from k to $k+1$. Hence, either it reaches $\lfloor \varepsilon K \rfloor$ before k , either it again jumps from k to $k+1$ and so on. Thus we approximate the probability that there is only one jump from k to $k+1$ by comparing $U_{j,k}^{(K,2)}$ with geometrically distributed random variables.

$$\begin{aligned} \hat{\mathbb{P}}(U_{j,k}^{(K,2)} = 1 \mid U_j^K = i) &\leq \sup_{n_A \in I_\varepsilon^K} \hat{\mathbb{P}}_{(n_A, k+1)}(T_\varepsilon^K < \sigma_k^K \mid T_\varepsilon^K < \sigma_j^K, U_j^K = i) \\ &\leq \sup_{\ell_\varepsilon^K \in \mathcal{L}_\varepsilon^K, n_A \in I_\varepsilon^K} \left\{ \frac{\mathbb{P}_{(n_A, k+1)}(T_\varepsilon^K < \sigma_k^K \mid U_j^K = i)}{\mathbb{P}_{(n_A, k+1)}(T_\varepsilon^K < \sigma_j^K \mid U_j^K = i)} \right\} \leq q_{j,k}^{(s_+(\varepsilon), s_-(\varepsilon))}, \end{aligned}$$

where we use Definition (A.12) for the last inequality. In an analogous way we show that $\hat{\mathbb{P}}(U_{j,k}^{(K,2)} = 1 \mid U_j^K = i) \geq q_{j,k}^{(s_-(\varepsilon), s_+(\varepsilon))}$. We deduce that we can construct two geometrically distributed random variables G_1 and G_2 , possibly on an enlarged space, with respective parameters $q_{j,k}^{(s_+(\varepsilon), s_-(\varepsilon))} \wedge 1$ and $q_{j,k}^{(s_-(\varepsilon), s_+(\varepsilon))}$ such that on the event $\{U_j^K = i\}$,

$$(B.3) \quad G_1 \leq U_{j,k}^{(K,2)} \leq G_2.$$

For the same reasons we obtain $q_{j,k}^{(s_-(\varepsilon), s_+(\varepsilon))} \leq \hat{\mathbb{P}}(U_{j,k}^{(K,2)} = 1) \leq q_{j,k}^{(s_+(\varepsilon), s_-(\varepsilon))} \wedge 1$, and again we can construct two random variables $G'_1 \stackrel{d}{=} G_1$ and $G'_2 \stackrel{d}{=} G_2$ such that

$$(B.4) \quad G'_1 \leq U_{j,k}^{(K,2)} \leq G'_2.$$

Taking $j = 0$ and adding the first part of Equation (A.13) give the first inequality of (7.10). According to Definition (5.1), for ε small enough, $|s_+(\varepsilon) - s_-(\varepsilon)| \leq c\varepsilon$ where c is a finite constant. Hence Equations (B.3), (B.4) and (A.13) entail the existence of a finite constant c such that for ε small enough $|\hat{\mathbb{E}}[U_{j,k}^{(K,2)} \mid U_j^K = i] - \hat{\mathbb{E}}[U_{j,k}^{(K,2)}]| \leq c\varepsilon + (1 - s_-(\varepsilon))^{k+1-j} / s_-^3(\varepsilon)$. Thus according to the first part of Equation (7.10),

$$\begin{aligned} \left| \widehat{\text{Cov}}(U_{j,k}^{(K,2)}, U_j^K) \right| &\leq \sum_{i \in \mathbb{N}^*} i \hat{\mathbb{P}}(U_j^K = i) \left| \hat{\mathbb{E}}[U_{j,k}^{(K,2)} \mid U_j^K = i] - \hat{\mathbb{E}}[U_{j,k}^{(K,2)}] \right| \\ (B.5) \quad &\leq \frac{2}{s_-^2(\varepsilon)} \left(c\varepsilon + \frac{(1 - s_-(\varepsilon))^{k+1-j}}{s_-^3(\varepsilon)} \right), \end{aligned}$$

which ends up the proof of (7.11).

Now we focus on the number of jumps for the process N_A . Definitions (2.4), the definition of $\mathcal{L}_\varepsilon^K$ and the inequality $s_-(\varepsilon) \leq s_+(\varepsilon)$ ensure that if $n_A \in I_\varepsilon^K$,

$$\begin{aligned} \hat{\mathbb{P}}_{(n_A, k)}(N_a(dt) = k+1) &\geq \frac{\mathbb{P}_{k+1}^{(s_-(\varepsilon))}(\tau_{\varepsilon K} < \tau_0)}{\mathbb{P}_k^{(s_+(\varepsilon))}(\tau_{\varepsilon K} < \tau_0)} f_a k dt + o(dt) \\ &= \frac{1 - (1 - s_-(\varepsilon))^{k+1}}{1 - (1 - s_-(\varepsilon))^{\lfloor \varepsilon K \rfloor}} \frac{1 - (1 - s_+(\varepsilon))^{\lfloor \varepsilon K \rfloor}}{1 - (1 - s_+(\varepsilon))^k} f_a k dt \geq s_-^2(\varepsilon) f_a k dt + o(dt), \end{aligned}$$

and

$$\begin{aligned}\hat{\mathbb{P}}_{(n_A, k)}(N_A(dt) \neq n_A) &\leq \left(f_A + D_A + C_{A,A} \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}}\right) + C_{A,a}\varepsilon\right) \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}}\right) K dt + o(dt) \\ &= (2f_A + 3C_{A,a}\varepsilon) \left(\bar{n}_A + 2\varepsilon \frac{C_{A,a}}{C_{A,A}}\right) K dt + o(dt).\end{aligned}$$

Thus for ε small enough:

$$\hat{\mathbb{P}}(N_a(\tau_{m+1}^K) \neq N_a(\tau_m^K) | N_a(\tau_m^K) = k) \geq \frac{s_-^2(\varepsilon)k}{6\bar{n}_A K}.$$

If D_k^K denotes the downcrossing number from k to $k-1$ before T_ε^K , then under the probability $\hat{\mathbb{P}}$, we can bound $U_k^K + D_k^K + H_k^K$ by the sum of $U_k^K + D_k^K$ independent geometrically distributed random variables G_i^K with parameter $s_-^2(\varepsilon)k/6\bar{n}_A K$ and $H_k^K \leq \sum_{1 \leq i \leq U_k^K + D_k^K} (G_i^K - 1)$. Let us remark that if $k \geq 2$, $D_k^K = U_{k-1}^K - 1$, and $D_1^K = 0$. Using the first part of (7.10) twice we get

$$\hat{\mathbb{E}}[H_k^K] \leq \left(\frac{4}{s_-^2(\varepsilon)} - 1\right) \left(\frac{6\bar{n}_A K}{s_-^2(\varepsilon)k} - 1\right),$$

which ends up the proof of the second inequality in (7.10).

As the mutant population size is not Markovian we cannot use symmetry and Strong Markov property to control the dependence of jumps before and after the last visit to a given state as in [25]. Hence we describe the successive excursions of N_a^K above a given level to get the last inequality in (7.10).

Let $Exc(j)$ be the number of excursions above j and $\tilde{U}_{j,k}^{(i)}$ the number of jumps from k to $k+1$ during the i th excursion above j . We first bound uniformly in i the expectation $\hat{\mathbb{E}}[(\tilde{U}_{j,k}^{(i)})^2]$. During an excursion above j , N_a hits $j+1$, but we do not know the value of N_a at this time. Thus we take the maximum value for the probability when n_A belongs to I_ε^K , and on the random event $\{i \leq Exc(j)\}$, $\hat{\mathbb{P}}(\tilde{U}_{j,k}^{(i)} \geq 1) \leq \sup_{n_A \in I_\varepsilon^K} \hat{\mathbb{P}}_{(j+1, n_A)}(\sigma_{k+1}^K < \sigma_j^K | \sigma_j^K < T_\varepsilon^K)$. Then using Coupling (5.3) and Definition (7.1) we obtain

$$\begin{aligned}\hat{\mathbb{P}}(\tilde{U}_{j,k}^{(i)} \geq 1) &\leq \sup_{\ell_\varepsilon^K \in \mathcal{L}_\varepsilon^K, n_A \in I_\varepsilon^K} \frac{\mathbb{P}_{(j+1, n_A)}(\sigma_{k+1}^K < \sigma_j^K, \sigma_j^K < T_\varepsilon^K, T_\varepsilon^K < \infty)}{\mathbb{P}_{(j+1, n_A)}(\sigma_j^K < T_\varepsilon^K, T_\varepsilon^K < \infty)} \\ &\leq \frac{\mathbb{P}_{j+1}^{(s_+(\varepsilon))}(\tau_{k+1} < \tau_j) \mathbb{P}_{k+1}^{(s_-(\varepsilon))}(\tau_j < \tau_{\varepsilon K}) \mathbb{P}_j^{(s_+(\varepsilon))}(\tau_{\varepsilon K} < \tau_0)}{\mathbb{P}_{j+1}^{(s_+(\varepsilon))}(\tau_j < \tau_{\varepsilon K}) \mathbb{P}_j^{(s_-(\varepsilon))}(\tau_{\varepsilon K} < \tau_0)}.\end{aligned}$$

Adding Equation (A.9) we finally get

$$(B.6) \quad \hat{\mathbb{P}}(\tilde{U}_{j,k}^{(i)} \geq 1) \leq \frac{(1 - s_-(\varepsilon))^{k+1-j}}{s_-(\varepsilon)(1 - s_+(\varepsilon))}.$$

Moreover if $\tilde{U}_{j,k}^{(i)} \geq 1$, N_a necessarily hits k after its first jump from k to $k+1$, and before its return to j . Using the same techniques as before we get:

$$\hat{\mathbb{P}}(\tilde{U}_{j,k}^{(i)} = 1 | \tilde{U}_{j,k}^{(i)} \geq 1) \geq \inf_{n_A \in I_\varepsilon^K} \left\{ \hat{\mathbb{P}}_{(n_A, k)}(\sigma_j^K < \sigma_{k+1}^K | \sigma_j^K < T_\varepsilon^K) \right\} \geq \frac{\mathbb{P}_k^{(s_+(\varepsilon))}(\tau_j < \tau_{k+1}) \mathbb{P}_j^{(s_-(\varepsilon))}(\tau_{\varepsilon K} < \tau_0)}{\mathbb{P}_k^{(s_-(\varepsilon))}(\tau_j < \tau_{\varepsilon K}) \mathbb{P}_j^{(s_+(\varepsilon))}(\tau_{\varepsilon K} < \tau_0)},$$

which yields

$$(B.7) \quad \hat{\mathbb{P}}(\tilde{U}_{j,k}^{(i)} = 1 | \tilde{U}_{j,k}^{(i)} \geq 1) \geq s_-(\varepsilon) s_+(\varepsilon) \left(\frac{1 - s_+(\varepsilon)}{1 - s_-(\varepsilon)}\right)^{k-j} \geq s_-^2(\varepsilon) \left(\frac{1 - s_+(\varepsilon)}{1 - s_-(\varepsilon)}\right)^{k-j}.$$

Hence, given that $\tilde{U}_{j,k}^{(i)}$ is non-null, $\tilde{U}_{j,k}^{(i)}$ is smaller than a geometrically distributed random variable with parameter $q = s_-^2(\varepsilon)[(1 - s_+(\varepsilon))/(1 - s_-(\varepsilon))]^{k-j}$. In particular,

$$\mathbb{E}\left[(\tilde{U}_{j,k}^{(i)})^2 \mid \tilde{U}_{j,k}^{(i)} \geq 1\right] \leq \frac{2}{q^2} = \frac{2}{s_-^4(\varepsilon)} \left(\frac{1 - s_-(\varepsilon)}{1 - s_+(\varepsilon)}\right)^{2(k-j)}.$$

Adding Equations (B.6) and recalling that $0 < s_-(\varepsilon) < s_+(\varepsilon) < 1$, we get for every $i \leq Exc(j)$

$$\hat{\mathbb{E}}\left[(\tilde{U}_{j,k}^{(i)})^2\right] \leq \frac{2\lambda_\varepsilon^{k-j}}{s_-^5(\varepsilon)(1 - s_+(\varepsilon))}, \quad \text{where } \lambda_\varepsilon := \frac{(1 - s_-(\varepsilon))^3}{(1 - s_+(\varepsilon))^2} < 1.$$

Finally, noticing that

$$\mathbb{1}_{Exc(j)=l} \hat{\mathbb{E}}\left[(U_{j,k}^{(K,1)})^2\right] \leq \mathbb{1}_{Exc(j)=l} l \sum_{i=1}^{Exc(j)} \hat{\mathbb{E}}\left[(\tilde{U}_{j,k}^{(i)})^2\right] \leq \mathbb{1}_{Exc(j)=l} l^2 \hat{\mathbb{E}}\left[(\tilde{U}_{j,k}^{(1)})^2\right],$$

that $Exc(j) = U_j^K - 1$ and recalling that $\hat{\mathbb{E}}[(U_j^K)^2] \leq 2/s_-^2(\varepsilon)$ end the proof of Equation (7.10).

We end with the proof of (7.12). Definition (A.12), Inequality (B.4) and Equation (A.9) yield:

$$r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{\hat{\mathbb{E}}[U_k^K]}{k+1} \geq r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \left[(k+1) q_{0,k}^{(s_+(\varepsilon), s_-(\varepsilon))} \right]^{-1} = \frac{r_K(A-B)}{s_+(\varepsilon)(1 - (1 - s_-(\varepsilon))^{\lfloor \varepsilon K \rfloor})},$$

with

$$A := \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{1 - (1 - s_-(\varepsilon))^{k+1}}{k+1}, \quad \text{and} \quad B := (1 - s_+(\varepsilon))^{\lfloor \varepsilon K \rfloor} \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{1 - (1 - s_-(\varepsilon))^{k+1}}{(1 - s_+(\varepsilon))^k(k+1)}.$$

For large K , $A = \log(\varepsilon K) + O(1)$. Moreover, Lemma 3.5 in [25] ensures that for $u > 1$ there exists $D(u) < \infty$ such that $\sum_{k=1}^{\lfloor \varepsilon K \rfloor} u^j / (k+1) \leq D(u) u^{\lfloor \varepsilon K \rfloor} / \lfloor \varepsilon K \rfloor$. This implies that $B \leq c / \lfloor \varepsilon K \rfloor$ for a finite c . Finally, by definition, for ε small enough, $|S_{aA} / f_a - s_+(\varepsilon)| \leq c\varepsilon$ for a finite constant c . This yields

$$r_K \sum_{k=1}^{\lfloor \varepsilon K \rfloor - 1} \frac{\hat{\mathbb{E}}[U_k^K]}{k+1} \geq (1 - c\varepsilon) \frac{r_K f_a \log K}{S_{aA}}$$

for a finite c and concludes the proof for the lower bound. The upper bound is obtained in the same way. This concludes the proof of Lemma 7.2. \square

Proof of Lemma 7.5. We use Coupling (5.3) to control the growing of the mutant population during the first period of invasion, and the semi-martingale decomposition in Proposition 1 to bound the fluctuations of M_A . The hitting time of $\lfloor \varepsilon K \rfloor$ and non-extinction event of Z_ε^* are denoted by:

$$T_\varepsilon^{*,K} = \inf\{t \geq 0, Z_\varepsilon^*(t) = \lfloor \varepsilon K \rfloor\}, \quad \text{and} \quad F_\varepsilon^* = \left\{ Z_\varepsilon^*(t) \geq 1, \forall t \geq 0 \right\}, \quad * \in \{-, +\}.$$

From (5.3) we get $\mathcal{D} = \{F_\varepsilon^-, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\} \subset \mathcal{C} = \{T_\varepsilon^K < \infty, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\} \subset \mathcal{F} = \{T_\varepsilon^{+,K} < \infty, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\}$. Lemma A.1 with $\mathcal{G} = \{T_\varepsilon^K > \tilde{T}_\varepsilon^K\} \cup \{T_\varepsilon^K = \infty\}$, and $\mathcal{H} = \{\sup_{t \leq T_\varepsilon^K} |P_{A,b_1}(t) - z_{Ab_1} / z_A| > \sqrt{\varepsilon}\}$ then yields:

$$(B.8) \quad |\mathbb{P}(\mathcal{H}, F_\varepsilon^-, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K) - \mathbb{P}(\mathcal{H}, T_\varepsilon^K < \infty)| \leq |\mathbb{P}(T_\varepsilon^{+,K} < \infty) - \mathbb{P}(F_\varepsilon^-)| + \mathbb{P}(T_\varepsilon^K < \infty, T_\varepsilon^K > \tilde{T}_\varepsilon^K).$$

But thanks to Equation (A.9) we get that $\mathbb{P}(T_\varepsilon^{+,K} < \infty) - \mathbb{P}(F_\varepsilon^-) = s_+(\varepsilon) / (1 - (1 - s_+(\varepsilon))^{\lfloor \varepsilon K \rfloor}) - s_-(\varepsilon)$, and (5.1) and (A.6) lead to:

$$(B.9) \quad |\mathbb{P}(T_\varepsilon^{+,K} < \infty) - \mathbb{P}(F_\varepsilon^-)| + \mathbb{P}(T_\varepsilon^K < \infty, T_\varepsilon^K > \tilde{T}_\varepsilon^K) \leq c\varepsilon.$$

It allows us to focus on the probability $\mathbb{P}(\mathcal{H}, F_\varepsilon^-, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K)$. We recall that $|N_{Ab_1} N_{ab_2} - N_{Ab_2} N_{ab_1}| \leq N_A N_a$, and that Assumption 5 holds. Then (3.1) and (3.11) imply for ε small enough

$$\sup_{t \leq T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \left| P_{A,b_1}(t) - \frac{z_{Ab_1}}{z_A - M_A(t)} \right| \leq r_K f_a T_\varepsilon^K \sup_{t \leq T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K} \left\{ \frac{N_a(t)}{N_A(t)} \right\} \leq \frac{r_K f_a \varepsilon T_\varepsilon^K}{\bar{n}_A - 2\varepsilon C_{A,a} / C_{A,A}} \leq \frac{c\varepsilon T_\varepsilon^K}{\log K}.$$

Moreover, $F_\varepsilon^- \cap \{T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\} \subset F_\varepsilon^- \cap \{T_\varepsilon^K \leq T_\varepsilon^{-,K}\}$. Thus we get

$$\mathbb{P}\left(\sup_{t \leq T_\varepsilon^K} \left| P_{A,b_1}(t) - \frac{z_{Ab_1}}{z_A - M_A(t)} \right| > \frac{\sqrt{\varepsilon}}{2}, F_\varepsilon^-, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\right) \leq \mathbb{P}\left(\frac{c\varepsilon T_\varepsilon^{-,K}}{\log K} > \sqrt{\varepsilon}/2, F_\varepsilon^-\right).$$

Finally, Equation (A.11) ensures that $\lim_{K \rightarrow \infty} T_\varepsilon^{-,K} / \log K = s_-(\varepsilon)^{-1}$ a.s. on the non-extinction event F_ε^- . Thus for $\varepsilon < s_-(\varepsilon)/2c$,

$$(B.10) \quad \lim_{K \rightarrow \infty} \mathbb{P}(\sup_{t \leq T_\varepsilon^K} |P_{A,b_1}(t) - M_A(t)| > \sqrt{\varepsilon}/2, F_\varepsilon^-, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K) = 0.$$

To control the term $|M_A|$, we introduce the sequence of real numbers $t_K = (2f_a \log K) / S_{aA}$:

$$\mathbb{P}\left(\sup_{t \leq T_\varepsilon^K} |M_A(t)| > \frac{\sqrt{\varepsilon}}{2}, F_\varepsilon^-, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K\right) \leq \mathbb{P}\left(\sup_{t \leq T_\varepsilon^K} |M_A(t)| > \frac{\sqrt{\varepsilon}}{2}, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K \wedge t_K\right) + \mathbb{P}(T_\varepsilon^K > t_K, F_\varepsilon^-).$$

Equation (5.1) yields for ε small enough, $t_K \cdot s_-(\varepsilon) / \log K > 3/2$. Thus thanks to (A.11) we get,

$$\lim_{K \rightarrow \infty} \mathbb{P}(T_\varepsilon^K > t_K, F_\varepsilon^-) \leq \lim_{K \rightarrow \infty} \mathbb{P}(T_\varepsilon^{-,K} > t_K, F_\varepsilon^-) = 0.$$

Applying Doob's maximal inequality to the submartingale $|M_A|$ and (3.12) we get:

$$(B.11) \quad \begin{aligned} \mathbb{P}(\sup_{t \leq T_\varepsilon^K} |M_A(t)| > \sqrt{\varepsilon}/2, T_\varepsilon^K \leq \tilde{T}_\varepsilon^K \wedge t_K) &\leq \mathbb{P}(\sup_{t \leq t_K} |M_A(t \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K)| > \sqrt{\varepsilon}/2) \\ &\leq \frac{4}{\varepsilon} \mathbb{E}\left[\langle M_A \rangle_{t_K \wedge T_\varepsilon^K \wedge \tilde{T}_\varepsilon^K}\right] \leq \frac{ct_K}{K}. \end{aligned}$$

Combining Equations (B.8) to (B.11) completes the proof of Lemma 7.5. \square

Acknowledgements: *The author would like to thank Jean-François Delmas and Sylvie Méléard for their help and their careful reading of this paper. She also wants to thank Sylvain Billiard and Pierre Collet for fruitful discussions during her work and several suggestions. This work was partially funded by project MANEGE 'Modèles Aléatoires en Écologie, Génétique et Évolution' 09-BLAN-0215 of ANR (French national research agency) and Chair Modélisation Mathématique et Biodiversité VEOLIA-Ecole Polytechnique-MNHN-FX.*

REFERENCES

- [1] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].
- [2] R. D. Barrett and D. Schluter. Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, 23(1):38–44, 2008.
- [3] N. H. Barton. The effect of hitch-hiking on neutral genealogies. *Genetical Research*, 72(2):123–133, 1998.
- [4] S. Billiard, R. Ferrière, S. Méléard, and V. C. Tran. Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *arXiv preprint arXiv:1310.6274*, 2013.
- [5] N. Champagnat. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116(8):1127–1160, 2006.
- [6] N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical population biology*, 69(3):297–321, 2006.
- [7] N. Champagnat and A. Lambert. Evolution of discrete populations and the canonical diffusion of adaptive dynamics. *The Annals of Applied Probability*, 17(1):102–155, 2007.
- [8] N. Champagnat and S. Méléard. Invasion and adaptive evolution for individual-based spatially structured populations. *Journal of Mathematical Biology*, 55(2):147–188, 2007.
- [9] N. Champagnat and S. Méléard. Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields*, 151(1-2):45–94, 2011.
- [10] P. Collet, S. Méléard, and J. A. Metz. A rigorous model study of the adaptive dynamics of mendelian diploids. *Journal of Mathematical Biology*, pages 1–39, 2011.
- [11] C. Coron. Stochastic modeling of density-dependent diploid populations and extinction vortex. *arXiv preprint arXiv:1207.4920*, to appear in *Advances in Applied Probabilities*, 2012.
- [12] C. Coron. Slow-fast stochastic diffusion dynamics and quasi-stationary distributions for diploid populations. *arXiv preprint arXiv:1309.3405*, 2013.
- [13] R. Durrett. *Probability models for DNA sequence evolution*. Springer, 2008.

- [14] R. Durrett and J. Schweinsberg. Approximating selective sweeps. *Theoretical population biology*, 66(2):129–138, 2004.
- [15] A. Etheridge, P. Pfaffelhuber, and A. Wakolbinger. An approximate sampling formula under genetic hitchhiking. *The Annals of Applied Probability*, 16(2):685–729, 2006.
- [16] S. Ethier and T. Kurtz. Markov processes: Characterization and convergence, 1986, 1986.
- [17] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4):1880–1919, 2004.
- [18] N. Ikeda and S. Watanabe. Stochastic differential equations and diffusion processes. 1989.
- [19] N. L. Kaplan, R. Hudson, and C. Langley. The " hitchhiking effect" revisited. *Genetics*, 123(4):887–899, 1989.
- [20] S. Leocard. Selective sweep and the size of the hitchhiking set. *Advances in Applied Probability*, 41(3):731–764, 2009.
- [21] S. Méléard and V. C. Tran. Trait substitution sequence process and canonical equation for age-structured populations. *Journal of Mathematical Biology*, 58(6):881–921, 2009.
- [22] J. A. Metz, S. A. Geritz, G. Meszéna, F. J. Jacobs, and J. Van Heerwaarden. Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. *Stochastic and spatial structures of dynamical systems*, 45:183–231, 1996.
- [23] B. M. Peter, E. Huerta-Sanchez, and R. Nielsen. Distinguishing between selective sweeps from standing variation and from a de novo mutation. *PLoS genetics*, 8(10):e1003011, 2012.
- [24] M. Prezeworski, G. Coop, and J. D. Wall. The signature of positive selection on standing genetic variation. *Evolution*, 59(11):2312–2323, 2005.
- [25] J. Schweinsberg and R. Durrett. Random partitions approximating the coalescence of lineages during a selective sweep. *The Annals of Applied Probability*, 15(3):1591–1651, 2005.
- [26] J. M. Smith and J. Haigh. The hitch-hiking effect of a favourable gene. *Genet Res*, 23(1):23–35, 1974.
- [27] W. Stephan, T. H. Wiehe, and M. W. Lenz. The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. *Theoretical Population Biology*, 41(2):237–254, 1992.