Survival of a recessive allele in a Mendelian diploid model

Rebecca Neukirch · Anton Bovier

Received: 16 October 2015 / Revised: 8 November 2016 / Published online: 28 November 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract In this paper we analyse the genetic evolution of a diploid hermaphroditic population, which is modelled by a three-type nonlinear birth-and-death process with competition and Mendelian reproduction. In a recent paper, Collet et al. (J Math Biol 67(3):569–607, 2013) have shown that, on the mutation time-scale, the process converges to the Trait-Substitution Sequence of adaptive dynamics, stepping from one homozygotic state to another with higher fitness. We prove that, under the assumption that a dominant allele is also the fittest one, the recessive allele survives for a time of order at least $K^{1/4-\alpha}$, where $K$ is the size of the population and $\alpha > 0$.

Keywords Adaptive dynamics · Population genetics · Mendelian reproduction · Diploid population · Nonlinear birth-and-death process · Genetic variability

Mathematics Subject Classification 60K35 · 92D25 · 60J85

We acknowledge financial support from the German Research Foundation (DFG) through the Hausdorff Center for Mathematics, the Cluster of Excellence ImmunoSensation, and the Priority Programme SPP1590 Probabilistic Structures in Evolution. We thank Loren Coquille for help with the numerical simulations and for fruitful discussions.

Rebecca Neukirch
rebecca.neukirch@iam.uni-bonn.de

Anton Bovier
bovier@uni-bonn.de

1 Institut für Angewandte Mathematik, Rheinische Friedrich-Wilhelms-Universität, Endenicher Allee 60, 53115 Bonn, Germany
1 Introduction

Mendelian diploid models have been studied for over a century in the context of population genetics (see, e.g., Yule 1907; Fisher 1918; Wright 1931; Haldane 1924a, b). Textbook expositions of population genetics are given in, e.g., Crow and Kimura (1970), Nagylaki (1992), Ewens (2004) and Bürger (2000). While population genetics typically deals with models of fixed population size, adaptive dynamics, a variant that has been developed in the 90ies (e.g., Hofbauer and Sigmund 1990; Marrow et al. 1992; Metz et al. 1992), allows for variable population sizes that are controlled by competition kernels that rule the competitive interaction of populations with different phenotypes or geographic locations. Diploid models have been considered in adaptive dynamics already by Kisdi and Geritz (1999).

Starting in the mid-90ies, stochastic individual based models were introduced and investigated that allow for a rigorous derivation of many of the predictions of adaptive dynamics on the basis of convincing models for populations of interacting individuals that incorporate the canonical genetic mechanisms of birth, death, mutation, and competition (see, e.g., Dieckmann and Law 1996; Champagnat et al. 2002, 2008; Champagnat 2006; Fournier and Méléard 2004; Champagnat and Méléard 2011). An important and interesting feature of these models is that various scaling limits when the carrying capacity tends to infinity while mutation rates and mutation step-size tend to zero yield different limit processes on different time-scales. In this way, Champagnat (2006) proves convergence to the Trait Substitution Sequence (TSS) (see, e.g., Dieckmann and Law 1996; Metz et al. 1996) and to the Canonical Equation of Adaptive Dynamics (CEAD). Champagnat and Méléard (2011) also rigorously derive the phenomenon of evolutionary branching under the assumption of coexistence. In a recent paper Baar et al. (2016) the convergence to the CEAD is shown in the simultaneously combined limits of large population, rare mutations and small mutation steps.

In the context of individual based models, so far almost exclusively haploid populations with asexual reproduction were studied. Exceptions are the paper by Collet et al. (2013) where the TSS is derived in a Mendelian diploid model under certain assumptions (that we will discuss below) and, more recently, some papers by Coron et al. (2013), Coron (2014, 2016). In the present paper we pick up this line of research and study a diploid population with Mendelian reproduction similar to the one of Collet et al. (2013), but with one notable difference in the assumptions. Each individual is characterised by a reproduction and death rates which depend on a phenotypic trait (e.g., body size, hair colour, rate of food intake, age of maturity) determined by its genotype, for which there exist two alleles A and a on one single locus. We examine the evolution of the trait distribution of the three genotypes aa, aA and AA under the three basic mechanisms: heredity, mutation and selection. Heredity transmits traits to new offspring and thus ensures the continued existence of the trait distribution. Mutation produces variation in the trait values in the population on which selection is acting. Selection is a consequence of competition for resources or area between individual. Collet et al. (2013) have shown that in the limit of large population and rare mutations, and under a co-dominance assumption of alleles, the suitably time-rescaled process, converges to the TSS model of adaptive dynamics, essentially as shown by Champagnat (2006) in the haploid case. We now reverse the assumption made by
Collet et al. (2013) that the alleles \(a\) and \(A\) are co-dominant and assume instead that \(A\) is the fittest and dominant allele, i.e., the genotypes \(aA\) and \(AA\) have the same phenotype. We show that this has a dramatic effect on the evolution of the population and, in particular, leads to a much prolonged survival of the “unfit” phenotype \(aa\) in the population. More precisely, we prove that the mixed type \(aA\) decays like \(1/t\), in contrast to the exponential decay observed by Collet et al. (2013). This type of behaviour has been observed earlier in the context of population genetic models (see, e.g., Nagylaki 1992, Chapter 4). The main result of the present paper is to show that this fact translated in the stochastic model into survival of the less fit phenotype for a time of order almost \(K^{1/4}\), when \(K\) is the carrying capacity (i.e. the order of the total population size). Let us emphasise that the main difficulty in our analysis is to control the behaviour of the stochastic system over a time horizon that diverges like a power of \(K\). This precludes in particular the use of functional laws of large numbers, or the like. Instead, our proof relies on the stochastic Euler scheme developed by Baar et al. (2016). One could probably give a heuristic derivation of this fact in the context of the diffusion approximation in the one locus two alleles model of population genetics (see, e.g., Ewens 2004), but we are not aware of a reference where this has actually been carried out.

Sexual reproduction in a diploid population amounts for each parent to transmit one of its two alleles to the genotype of the newborn. Hence, unfit alleles can survive in individuals with mixed genotype and individuals with a pure genotype are potentially able to reinvade in the population under certain circumstances, i.e. a new mutant allele \(B\) that appears before the extinction of the \(a\) allele that has strong competition with the \(AA\) population but weak competition with the \(aa\) population may lead to a resurgence of the \(aa\) population at the expense of the \(AA\) population and coexistence of the types \(aa\) and \(BB\). This would increase the genetic variability of the population. In other words, if we choose the mutation time scale in such a way that there remain enough \(a\)-alleles in the population when a new mutation occurs and if the new mutant can coexist with the unfit \(aa\)-individuals, then the \(aa\)-population can recover. Numerical simulations show that this can happen but requires subtle tuning of parameters. This effect will be analysed in a forthcoming publication. Related questions have recently been addressed in haploid models by Billiard and Smadi (2016).

2 Model setup and goals

2.1 Introduction of the model

We consider a Mendelian diploid model introduced by Collet et al. (2013). It models a population of a finite number of individuals with sexual reproduction, where each individual \(i\) is characterised by two alleles, \(u_{i1}^1 u_{i2}^2\), from some allele space \(\mathcal{U} \subset \mathbb{R}\). These two alleles define the genotype of individual \(i\), which in turn defines its phenotype, \(\phi(u_{i1}^1 u_{i2}^2)\), through a function \(\phi: \mathcal{U}^2 \to \mathbb{R}\). We suppress parental effects, thus \(\phi(u_{i1}^1 u_{i2}^2) = \phi(u_{i2}^2 u_{i1}^1)\). The individual-based microscopic Mendelian diploid model is a non-linear stochastic birth-and-death process. Each individual has a Mendelian reproduction rate with mutation and a natural death rate. Moreover, there is an additional
death rate due to ecological competition with the other individuals in the population. The following demographic parameters depend all on the phenotype, but we suppress this from the notation. Let us define

| $f_{u_1u_2} \in \mathbb{R}^+_+$ | The rate of birth (fertility) of an individual with genotype $u_1u_2$ |
| $D_{u_1u_2} \in \mathbb{R}^+_+$ | The rate of natural death of an individual with genotype $u_1u_2$ |
| $K \in \mathbb{N}$ | The parameter which scales the population size |
| $c_{u_1v_1,v_1v_2} \in \mathbb{R}^+_+$ | The competition effect felt by an individual with genotype $u_1u_2$ from an individual with genotype $v_1v_2$ |
| $\mu_K \in \mathbb{R}^+_+$ | The mutation probability per birth event. Here it is independent of the genotype |
| $\sigma > 0$ | The parameter scaling the mutation amplitude |
| $m(u, dh)$ | Mutation law of a mutant allelic trait $u + h \in \mathcal{U}$, born from an individual with allelic trait $u$ |

Scaling the competition function $c$ down by a factor $1/K$ amounts to scaling the population size to order $K$. $K$ is called the carrying capacity. We are interested in asymptotic results when $K$ is very large. We assume rare mutations, i.e. $\mu_K \ll 1$. Hence, if a mutation occurs at a birth event, only one allele changes from $u$ to $u + \sigma h$ where $h$ is a random variable with law $m(u, dh)$ and $\sigma \in [0, 1]$.

At any time $t$, there is a finite number, $N_t$, of individuals, each with genotype in $\mathcal{U}^2$. We denote by $u_1^1u_2^1, \ldots, u_1^n u_2^n$ the genotypes of the population at time $t$. The population, $v_t$, at time $t$ is represented by the rescaled sum of Dirac measures on $\mathcal{U}^2$,

$$v_t = \frac{1}{K} \sum_{i=1}^{N_t} \delta_{u_1^i u_2^i},$$

$v_t$ takes values in

$$\mathcal{M}^K = \left\{ \frac{1}{K} \sum_{i=1}^{n} \delta_{u_1^i u_2^i} \big| n \geq 0, u_1^1 u_2^1, \ldots, u_1^n u_2^n \in \mathcal{U}^2 \right\},$$

where $\mathcal{M}$ denotes the set of finite, nonnegative measures on $\mathcal{U}^2$ equipped with the vague topology. Define $\langle v, g \rangle$ as the integral of the measurable function $g : \mathcal{U}^2 \to \mathbb{R}$ with respect to the measure $v \in \mathcal{M}^K$. Then $\langle v_t, 1 \rangle = \frac{N_t}{K}$ and for any $u_1 u_2 \in \mathcal{U}^2$, the positive number $\langle v_t, 1_{u_1 u_2} \rangle$ is called the density at time $t$ of the genotype $u_1 u_2$. The generator of the process is defined as by Collet et al. (2013): first we define, for the genotypes $u_1 u_2, v_1 v_2$ and a point measure $v$, the Mendelian reproduction operator:

$$(A_{u_1u_2,v_1v_2}F)(v)$$

$$= \frac{1}{4} \left[ F \left( v + \frac{\delta_{u_1v_1}}{K} \right) + F \left( v + \frac{\delta_{u_1v_2}}{K} \right) + F \left( v + \frac{\delta_{u_2v_1}}{K} \right) + F \left( v + \frac{\delta_{u_2v_2}}{K} \right) \right] - F(v),$$
and the Mendelian reproduction-cum-mutation operator:

\[
(M_{u_1u_2,v_1v_2} F)(v) = \frac{1}{8} \int_{\mathbb{R}} \left[ \left( F \left( v + \frac{\delta_{u_1hv_1}}{K} \right) + F \left( v + \frac{\delta_{u_2hv_1}}{K} \right) \right) m_\sigma(u_1, h) + \left( F \left( v + \frac{\delta_{u_1hv_2}}{K} \right) + F \left( v + \frac{\delta_{u_2hv_2}}{K} \right) \right) m_\sigma(u_2, h) + \left( F \left( v + \frac{\delta_{u_1v_1h}}{K} \right) + F \left( v + \frac{\delta_{u_2v_1h}}{K} \right) \right) m_\sigma(v_1, h) + \left( F \left( v + \frac{\delta_{u_1v_2h}}{K} \right) + F \left( v + \frac{\delta_{u_2v_2h}}{K} \right) \right) m_\sigma(v_2, h) \right] dh - F(v).
\]

(2.4)

The process \((v_t)_{t \geq 0}\) is then a \(\mathcal{M}^K\)-valued Markov process with generator \(L^K\), given for any bounded measurable function \(F: \mathcal{M}^K \to \mathbb{R}\) and \(v \in \mathcal{M}^K\) by:

\[
(L^K F)(v) = \int_{\mathcal{U}^2} \left( D_{u_1u_2} + \int_{\mathcal{U}_2} c_{u_1u_2,v_1v_2} \nu(d((v_1v_2))) \right) \left( F \left( v - \frac{\delta_{u_1u_2}}{K} \right) - F(v) \right) K \nu(d(u_1u_2)) + \int_{\mathcal{U}^2} (1 - \mu_K) f_{u_1u_2} \left( \int_{\mathcal{U}_2} \frac{f_{v_1v_2}}{\nu(v)} (A_{u_1u_2,v_1v_2} F)(v) \nu(d(v_1v_2)) \right) K \nu(d(u_1u_2)) + \int_{\mathcal{U}^2} \mu_K f_{u_1u_2} \left( \int_{\mathcal{U}_2} \frac{f_{v_1v_2}}{\nu(v)} (M_{u_1u_2,v_1v_2} F)(v) \nu(d(v_1v_2)) \right) K \nu(d(u_1u_2)).
\]

(2.5)

The first non-linear term describes the competition between individuals. The second and last linear terms describe the birth without and with mutation. There, \(f_{u_1u_2} f_{v_1v_2} K_{\nu(v)}\) is the reproduction rate of an individual with genotype \(u_1u_2\) with an individual with genotype \(v_1v_2\). Note that we assume random mating with multiplicative fertility (i.e. that birth rate is proportional to the product of the fertilities of the mates).

For all \(u_1u_2, v_1v_2 \in \mathcal{U}^2\), we make the following Assumptions (A):

(A1) The functions \(f\), \(D\) and \(c\) are measurable and bounded, which means that there exist \(\bar{f}, \bar{D}, \bar{c} < \infty\) such that

\[
0 \leq f_{u_1u_2} \leq \bar{f}, \quad 0 \leq D_{u_1u_2} \leq \bar{D} \quad \text{and} \quad 0 \leq c_{u_1u_2,v_1v_2} \leq \bar{c}.
\]

(2.6)

(A2) \(f_{u_1u_2} - D_{u_1u_2} > 0\) and there exists \(\zeta > 0\) such that \(\zeta \leq c_{u_1u_2,v_1v_2}\).

(A3) For any \(\sigma > 0\), there exists a function, \(\tilde{m}_\sigma : \mathbb{R} \to \mathbb{R}_+\), such that \(\int \tilde{m}_\sigma(h) dh < \infty\) and \(m_\sigma(u, h) \leq \tilde{m}_\sigma(h)\) for any \(u \in \mathcal{U}\) and \(h \in \mathbb{R}\).

For fixed \(K\), under the Assumptions (A1) + (A3) and assuming that \(\mathbb{E}(\nu_0, 1) < \infty\), Fournier and Méléard (2004) have shown existence and uniqueness in law of a process with infinitesimal generator \(L^K\). For \(K \to \infty\), under more restrictive assumptions, and assuming the convergence of the initial condition, they prove the convergence in \(\mathcal{D}(\mathbb{R}_+, \mathcal{M}^K)\) of the process \(v^K\) to a deterministic process, which is the solution to a
non-linear integro-differential equation. Assumption (A2) ensures that the population does not explode or becomes extinct too fast.

2.2 Goal

We start the process with a monomorphic $aa$-population, where one mutation to an $A$-allele has already occurred. That means, the initial population consists only of individuals with genotype $aa$ except one individual with genotype $aA$. The mutation probability for an individual with genotype $u_1u_2$ is given by $\mu_K$. Hence, the time until the next mutation in the whole population is of order $1/K\mu_K$. Since the time a mutant population needs to invade a resident population is of order $\ln K$ (see, e.g., Champagnat 2006), we set the mutation rate $1/K\mu_K \gg \ln(K)$ in order to be able to consider the fate of the mutant and the resident population without the occurrence of a new mutation. In this setting, the allele space $\mathcal{U} = \{a, A\}$ consists only of two alleles. Our results will imply that if the mutation rate is bigger than $1/K^{1/4-\alpha}$, $\alpha > 0$, then a mutation will occur while the resident $aa$-population is small, but still alive, in contrast to the setting of Collet et al. (2013), where the $a$-allele dies out by time $\ln K$. This different behaviour can be traced to the deterministic system that arises in the large $K$ limit. Figure 1 ($A$-allele fittest and dominant) and Fig. 2 ($a$ and $A$ alleles co-dominant) show simulations of the deterministic systems of the two different models. We see that in the settings of Collet et al. (2013) the mixed type $aA$ dies out exponentially fast, whereas in the model where $A$ is the dominant allele, the mixed type decays much slowly. We will show below that this is due to the fact that, under our hypothesis, the stable fixed point of the deterministic system is degenerate, leading to an algebraic rather than exponential approach to the fixed point. The main task is to prove that this translates into a survival of the unfit allele in the stochastic model for a time of order $K^\beta$. We show that this is indeed the case, with $\beta = 1/4 - \alpha$. This implies that for mutation rates of order $1/K \ln K$, a further mutant will occur in the $AA$ population before the $aa$ allele is extinct.

2.3 Assumptions on the model

Let $N_{uv}(t)$ be the number of individuals with genotype $uv \in \{aa, aA, AA\}$ in the population at time $t$ and set $n_{uv}(t) \equiv 1/K N_{uv}(t)$.

**Definition 1** The equilibrium size of a monomorphic $uu$-population, $u \in \{a, A\}$, is the fixed point of a 1-dim Lotka–Volterra equation and is given by

$$\bar{n}_u = \frac{f_{uu} - D_{uu}}{c_{uu,uu}}. \quad (2.7)$$

**Definition 2** For any $u, v \in \{a, A\}$,

$$S_{uv,uu} = f_{uv} - D_{uv} - c_{uv,uu} \bar{n}_v \quad (2.8)$$
is called the *invasion fitness* of a mutant $uv$ in a resident $uu$-population, where $\tilde{n}_u$ is given by (2.7).

We assume that the dominant $A$-allele defines the phenotype of an individual, i.e. $AA$ and $Aa$ individuals have the same phenotype. In particular, the fertility and the natural death rates are the same for $aA$- and $AA$-individuals. For simplicity, we assume that the competition rates are the same for all the three different genotypes. To sum up, we make the following Assumptions (B) on the rates:

(B1) $f_{aa} = f_{aA} \equiv f_{AA} =: f$,
(B2) $D_{AA} \equiv D_{aA} =: D$ but $D_{aa} = D + \Delta$,
(B3) $c_{u_1u_2,v_1v_2} =: c$, $\forall u_1u_2, v_1v_2 \in \{aa, aA, AA\}$.

**Remark 1** We choose constant fertilities and constant competition rates for simplicity. What is really needed, is that the fitness of the $aA$ and $AA$ types are equal and higher than that of the $aa$ type.

Observe that, under Assumptions (B),

$$S_{aA,aa} = S_{AA,aa} = f - D - c\tilde{n}_{aa} = f - D - c\frac{f - D - \Delta}{c} = \Delta,$$

$$S_{aa,aA} = S_{aA,AA} = f - D - \Delta - c\tilde{n}_{AA} = f - D - \Delta - c\frac{f - D}{c} = -\Delta. \quad (2.9)$$

Therefore, the $aA$-individuals are as fit as the $AA$-individuals and both are fitter than the $aa$-individuals. In our model, an individual chooses a partner uniformly at random for reproduction, and, according the Mendelian laws, each individual transmits one allele, chosen uniformly at random from its genotype, to the offspring’s genotype. For example, if we want to produce an individual with genotype $aa$, there are four possible combinations for the parents: $aa \leftrightarrow aa$, $aa \leftrightarrow aA$, $aA \leftrightarrow aa$ and $aA \leftrightarrow aA$. The first combination results in an $aa$-individual with probability $\frac{1}{2}$, the second and third one with probability $\frac{1}{4}$ and the last one with probability $\frac{1}{4}$. Therefore we have the following birth rates:

$$b_{aa}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = \frac{f(N_{aa}(t) + \frac{1}{2}N_{aA}(t))^2}{N_{aa}(t) + N_{aA}(t) + N_{AA}(t)}.$$

$$b_{aA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = \frac{2f(N_{aa}(t) + \frac{1}{2}N_{aA}(t))(N_{AA}(t) + \frac{1}{2}N_{aA}(t))}{N_{aa}(t) + N_{aA}(t) + N_{AA}(t)},$$

$$b_{AA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = \frac{f(N_{AA}(t) + \frac{1}{2}N_{aA}(t))^2}{N_{aa}(t) + N_{aA}(t) + N_{AA}(t)}. \quad (2.10)$$

The death rates are the sum of the natural death and the competition:

$$d_{aa}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = N_{aa}(t) \left( D + \Delta + c(n_{aa}(t) + n_{aA}(t) + n_{AA}(t)) \right),$$

$$d_{aA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = N_{aA}(t) \left( D + c(n_{aa}(t) + n_{aA}(t) + n_{AA}(t)) \right),$$

$$d_{AA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = N_{AA}(t) \left( D + c(n_{aa}(t) + n_{aA}(t) + n_{AA}(t)) \right). \quad (2.11)$$
In the sequel, the *sum process*, $\Sigma(t)$, defined by
\[ \Sigma(t) = n_{aa}(t) + n_{aA}(t) + n_{AA}(t), \]
plays an important role. A simple calculation shows that the sum process jumps up to $1/K$ (resp. down) with rate $b_\Sigma$ (resp. $d_\Sigma$) given by
\[ b_\Sigma(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = f_K \Sigma(t), \]
\[ d_\Sigma(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = D_K \Sigma(t) + \Delta N_{aa}(t). \]

### 3 Main theorems

In the sequel we denote by $\tau_n$, $n \geq 1$, the ordered sequence of times when a mutation occurs in the population. We assume $\tau_0 = 0$ and make the following Assumption (C) on the mutation rate $\mu_K$:
\[ (C) \quad \ln(K) \leq \frac{1}{K \mu_K} \ll K^{1/4-\alpha}. \]

We recall one result from Collet et al. (2013) (Proposition D.2) which carries over to our setting: it is shown that if the resident population $n_{aa}(t)$ is in a $\delta$-neighbourhood of its equilibrium $\bar{n}_a$, then $n_{aa}(t)$ stays in this neighbourhood for an exponentially long time, as long as the mutant population is smaller than $\delta$. The proof of this result is based on large deviation estimates (see Freidlin and Wentzell 1984).

**Proposition 3.1** (Proposition D.2 in Collet et al. 2013) Let $\text{supp}(\nu^K_0) = \{aa\}$ and let $\tau_1$ denote the first mutation time. For any sufficiently small $\delta > 0$, if $\langle \nu^K_0, \mathbb{1}_{aa} \rangle$ belongs to the $\delta/2$-neighbourhood of $\bar{n}_a$, then the time of exit of $\langle \nu^K_0, \mathbb{1}_{aa} \rangle$ from the $\delta$-neighbourhood of $\bar{n}_a$ is bigger than $e^{V_K \wedge \tau_1}$, for $V > 0$, with probability converging to 1. Moreover, there exists a constant $M$, such that, for any sufficiently small $\delta > 0$, this remains true, if the death rate of an individual with genotype aa,
\[ D + c_K \langle \nu^K_1, \mathbb{1}_{aa} \rangle, \]
is perturbed by an additional random process that is uniformly bounded by $M\delta$.

We start the population process when $n_{aa}$ is in a $\delta/2$-neighbourhood of its equilibrium, $\bar{n}_a$, and there is one individual with genotype $aA$. The first theorem says that there is a positive probability that the mutant population fixates in the resident $aa$-population and the second theorem gives the time for the invasion of the mutant population and a lower bound on the survival time of the recessive $a$-allele. Define
\[ \tau^\text{mut}_\delta = \inf\{t \geq 0 : 2n_{AA}(t) + n_{aA}(t) \geq \delta\}, \]
\[ \tau^\text{mut}_0 = \inf\{t \geq 0 : 2n_{AA}(t) + n_{aA}(t) = 0\}. \]

**Theorem 3.1** (Proposition D.4 in Collet et al. 2013) Let $(z_K)$ be a sequence of integers such that $z_K/K$ converges to $\bar{n}_a$, for $K \to \infty$. Then
Survival of a recessive allele in a Mendelian diploid model

\[
\lim_{\delta \to 0} \lim_{K \to \infty} \mathbb{P} \left[ \sum_{\substack{aA \delta \rightarrow 0 \delta \rightarrow \infty}} \mathbb{P} \left( \tau^\text{mut}_{\delta} < \tau^\text{mut}_0 \right) = \frac{\Delta}{f} \right].
\] (3.5)

where we recall that \( \Delta \) is the invasion fitness of a mutant \( aA \) in a resident \( aa \)-population (cf. 2.9).

We now state the main results of this paper:

**Theorem 3.2** Consider the model verifying Assumptions A and B. Let \( \delta > \epsilon, \alpha > 0 \) and

\[
\tau^\text{hit}_{\eta} \equiv \inf \{ t \geq \tau^\text{mut}_{\delta} : n_{aA}(t) \leq \eta \}. \tag{3.6}
\]

Define \( \tau^\text{sur}_{\epsilon} \equiv \tau^\text{hit}_{\epsilon K} - \tau^\text{hit}_{\epsilon} \). Then, conditional on survival of the mutant, i.e., on the event \( \{ \tau^\text{mut}_{\delta} < \tau^\text{mut}_0 \} \), with probability converging to one as \( K \uparrow \infty \), the following statements hold:

(i) \( \tau^\text{hit}_{\epsilon} = O(\ln K) \), and

(ii) \( \tau^\text{sur}_{\epsilon} = O(K^{1/4 - \alpha}) \).

**Remark 2** As long as there are \( aA \)-individuals in the population, the smaller \( aa \)-population does not die out, since the \( aA \)-population always gives birth to \( aa \)-individuals. For smaller values of the power \( 1/4 - \alpha \) in (ii), the natural fluctuations of the big \( AA \)-population are too high: the death rate of \( n_{aa}(t) \) can be too large due to the competition felt by \( n_{AA}(t) \) and could induce the death of the \( aa \)-population and hence also of the \( aA \)-population.

The next theorem states that if the mutation rate satisfies Assumption C, then a new mutation to a (possibly fitter) allele, \( B \), occurs while some \( a \)-alleles are still alive. This mutation will happen after the invasion of the mutant population and when the mixed type \( aA \)-population already decreased to a small level again. More precisely,

**Theorem 3.3** Assume that Assumption C is satisfied. Then, with probability converging to one,

\[
\tau^\text{hit}_{\epsilon} \wedge \tau_1 = \tau^\text{hit}_{\epsilon} \quad \text{and} \quad \tau_1 \wedge \tau^\text{hit}_1 = \tau_1. \tag{3.7}
\]

The interest in this result lies in the fact that a new mutant allele \( B \) that appears before the extinction of the \( a \)-allele that has strong competition with the \( AA \)-population but weak competition with the \( aa \)-population may lead to a resurgence of the \( aa \)-population at the expense of the \( AA \)-population and coexistence of the types \( aa \) and \( BB \). Numerical simulations, which are objects of a following publication, show that this can happen but requires subtle tuning of parameters. Since the proofs of the main Theorems (3.1, 3.2, 3.3) have several parts and are quite technical, we first give an outline of them before we turn to the details (Sect. 5.2).
3.1 Outline of the proofs

3.1.1 Heuristics leading to the main theorems

The basis of the main theorems is the observation of the different behaviour of the limiting deterministic system, when A is the fittest and dominant one and when the alleles are co-dominant (Collet et al. 2013). More precisely, they have dissimilar long-term behaviour (cf. Figs. 1, 2).

By analysing the systems one gets that both have the same fixed points $n_{aa} \equiv (\bar{n}_a, 0, 0)$ and $n_{AA} \equiv (0, 0, \bar{n}_A)$. The computation of the eigenvalues of the Jacobian matrix at the fixed point $n_{aa}$ yields in both models two negatives and one positive eigenvalues. Hence in both systems the fixed point $n_{aa}$ is unstable. In contrast, the eigenvalues at the fixed point $n_{AA}$ are all negative in the system of Collet et al. (2013) whereas in this model there is one zero eigenvalue. This leads to the different long term behaviour towards the stable fixed point $n_{AA}$. In Collet et al. (2013) model the $aA$-population dies out exponentially fast whereas in this model the degenerated eigenvalue corresponds to a decay of $n_{AA}(t)$ like a function $f(t) = \frac{1}{t}$. The goal is to show that the stochastic system behaves like the deterministic system.

3.1.2 Organisation of the proofs

The main theorems describe the invasion of a mutant in the resident population, and the survival of the recessive allele. This invasion can be divided into three phases, in a similar way as in Champagnat (2006), Collet et al. (2013) or Baar et al. (2016) (cf.
Fig. 2  Collet et al. model: $a$ and $A$ co-dominant

Fig. 3  The three phases of the proof (Simulation by Loren Coquille)

Fig. 3) (the general idea that an invasion can be divided in these phases is much older, see, e.g., Rouhani and Barton 1987):

**Phase 1** Fixation of the mutant population,
**Phase 2** Invasion of the mutant population,
**Phase 3** Survival of the recessive allele.
The first two phases are similar to the ones in Collet et al. (2013), whereas the last phase will be analysed in eight steps. Technically, the analysis uses tools developed by Baar et al. (2016) and classical potential methods (see, e.g., Bovier and den Hollander 2015).

**Settings for the steps**

- **Step 1** Upper bound on $\Sigma(t)$,
- **Step 2** Upper bound on $n_{aa}(t)$,
- **Step 3** Lower bound on $\Sigma(t)$,
- **Step 4** Upper and lower bound on $n_{AA}(t)$,
- **Step 5** Decay of $n_{aA}(t)$,
- **Step 6** Decay time of $n_{aA}(t)$,
- **Step 7** Decay and decay time of $n_{aa}(t)$,
- **Step 8** Growth and growth time of $\Sigma(t)$,
- **Total decay time of** $n_{aA}(t)$.

**Phase 1: Fixation of the mutant population** In the first phase we show that there is a positive probability that the fitter mutant population $A(t) \equiv n_{aA}(t) + 2n_{AA}(t)$ fixates in the resident population. More precisely, as long as the mutant population size is smaller than a fixed $\delta$, the resident $aa$-population stays close to its equilibrium $\bar{n}_a$ (Proposition 3.1) and its dynamics are nearly the same as before since the influence of the mutant population is negligible. We can approximate the dynamics of the mutant population $A(t)$ by a rescaled birth and death process and can show that the probability that this branching process increases to a $\delta$-level is close to its survival probability and hence also the probability that the mutant population $A(t) \equiv n_{aA}(t) + 2n_{AA}(t)$ grows up to a size $\delta$. This is the content of Theorem 3.1.

**Phase 2: Invasion of the mutant population** The fixation (Phase 1) ends with a macroscopic mutant population of size $\delta$. In the second phase the mutant population invades the resident population and suppresses it. By the Large Population Approximation (Theorem 4.1, Fournier and Méléard 2004) the behaviour of the process is now close to the solution of the deterministic system (4.1) with the same initial condition on any finite time interval, when $K$ tends to infinity. Thus, we get from the analysis of the dynamical system in Sect. 4 that any solution starting in a $\delta$-neighbourhood of $(\bar{n}_a, 0, 0)$ converges to an $\epsilon$-neighbourhood of $(0, 0, \bar{n}_A)$ in finite time ($t_2$ in Fig. 3).

Since we see in the dynamical system that as soon as the $AA$-population is close to its equilibrium, the $aA$-population decays like $\frac{1}{t}$, we only proceed until $n_{aA}$ decreases to an $\epsilon$-level to ensure that the duration of this phase is still finite.

In Champagnat (2006), it is shown that the duration of the first phase is of order $O(\ln K)$ and that the time for the second phase is bounded. Thus the time needed by the $aA$-population to reach again the $\epsilon$-level after the fixation is of order $O(\ln K)$ (cf. Theorem 3.2 (i)). From Proposition 3.1 we get that the resident $aa$-population stays in a $\delta$-neighbourhood of its equilibrium $\bar{n}_a$ an exponentially long time $\exp(VK)$ as long as the mutant population is smaller than $\delta$. Thus we can approximate the rate of mutation until this exit time by $\mu_K f K \bar{n}_a$. Hence the waiting time for mutation to occur is of order $\frac{1}{K \mu_K}$. Champagnat (2006) proved that there is also no accumulation of
mutations in the second phase. More precisely, he shows that, for any initial condition, the probability of a mutation on any bounded time interval is very small:

**Lemma 3.1** (Lemma 2 (a) in Champagnat 2006) Assume that the initial condition of \( \nu_t \) satisfies \( \sup_K \mathbb{E}(\langle \nu_0, 1 \rangle) < \infty \). Then, for any \( \eta > 0 \), there exists an \( \epsilon > 0 \) such that, for any \( t > 0 \),

\[
\limsup_{K \to \infty} \mathbb{P}^{\nu_0}_{\mathcal{K}} \left( \exists n \geq 0 : \frac{t}{K \mu_K} \leq \tau_n \leq \frac{t + \epsilon}{K \mu_K} \right) < \eta, \tag{3.8}
\]

where \( \tau_n \) are the ordered sequence of times when a mutation occurs, defined in the beginning of Sect. 3.

Using Lemma 3.1, we get that, for fixed \( \eta > 0 \), there exists a constant, \( \eta > \rho > 0 \), such that, for sufficiently large \( K \),

\[
\mathbb{P}^{\delta_{aa} + \frac{1}{K} \delta_{aA}} \left( \tau_1 < \frac{\rho}{K \mu_K} \right) < \delta, \tag{3.9}
\]

where \( \tau_1 \) is the time of the next mutation. Thus, the next mutation occurs with high probability not before a time \( \frac{\rho}{K \mu_K} \). Hence, under the assumption that

\[
\ln(K) \ll \frac{1}{K \mu_K}, \tag{3.10}
\]

(cf. left inequality of (3.1)) there appears no mutation before the first and second phase are completed.

**Phase 3: Survival of the recessive allele** The last phase starts as soon as \( n_{aA}(t) \) has decreased to an \( \epsilon \)-level. This phase is different from the one in Champagnat (2006) and Collet et al. (2013), since the analysis of the deterministic system in Sect. 4 reveals that \( n_{aA}(t) \) decreases only like a function \( f(t) = \frac{1}{t} \), in contrast to the exponential decay in Collet et al. (2013). Thus, we may expect that the time to extinction will not be \( \mathcal{O}(\ln K) \) anymore, and the recessive allele \( a \) will survive in the population for a much longer time. This is a situation similar to the one encountered by Baar et al. (2016), where it was necessary to show that the stochastic system remains close to a deterministic one over times of order \( K^\alpha \) due to the fact that the evolutionary advantage of the mutant population vanishes like a negative power of \( K \). Just as in that case, we cannot use the Law of Large Numbers, but we adopt the stochastic Euler scheme from Baar et al. (2016) to show that the behaviour of the deterministic and the stochastic systems remain close for a time of order \( K^{1/4-\alpha} \).

Let us put this scheme on a mathematically footing (Fig. 4):

**Settings for the steps** We define stopping times depending on \( n_{aA}(t) \) in such a way that we can control the other processes \( n_{aa}(t), n_{AA}(t), \) and \( \Sigma(t) \) on the resulting time intervals. Fix \( \epsilon > 0 \) and \( \vartheta > 0 \) such that \( \epsilon < \frac{\Delta}{2} < \vartheta < \Delta \). We set
Fig. 4 Steps of the proof (Simulation by Loren Coquille)

\[ x = \left( \frac{f + \vartheta}{f + \Delta} \right)^{\frac{1}{2}}, \]  

(3.11)

and, for \(0 \leq i \leq \left\lfloor \frac{-\ln(\epsilon K^{1/4-\alpha})}{\ln(x)} \right\rfloor\), with \(\alpha > 0\), we define the stopping times on \(n_{aA}(t)\) by

\[
\begin{align*}
\tau_{i-}^{aA} &\equiv \begin{cases} 
\tau_{i-}^{hit}, & \text{for } i = 0, \\
\inf\{t \geq \tau_{i-}^{(i-1)}: n_{aA}(t) \leq x^i \epsilon\}, & \text{else,}
\end{cases} \\
\tau_{i+}^{aA} &\equiv \inf\{t \geq \tau_{i-}^{(i-)}: n_{aA}(t) \geq x^i \epsilon + x^{2i} \epsilon^2\}.
\end{align*}
\]

(3.12)

(3.13)

During the time intervals \(t \in [\tau_{i-}^{aA}, \tau_{i+}^{aA} \wedge \tau_{i+}^{(i+1)-}]\), \(n_{aA}(t) \in [x^{i+1} \epsilon, x^i \epsilon + x^{2i} \epsilon^2]\).

The upper bound on \(i\) is chosen in such a way that

\[ x^i \epsilon \geq K^{-1/4+\alpha}. \]

(3.14)

The following eight steps will be iterated from \(i = 0\) to \(i = \left\lfloor \frac{-\ln(\epsilon K^{1/4-\alpha})}{\ln(x)} \right\rfloor\).

**Remark 3** Since in Phase 3 the biggest contribution to the birth rate of \(n_{aa}(t)\) is given by the combination of two \(aA\)-individuals, \(n_{aa}\) behaves like \(n_{aA}^2\). We let \(n_{aA}\) decrease only until \(K^{-1/4+\alpha}\). Afterwards \(n_{aa}\) would be of smaller order than \(K^{-1/2}\) and the natural fluctuations of the big \(AA\)-population, of order \(K^{-1/2}\), would induce the death
of the $aa$-population due to competition. Since $n_{aa}$ contributes to the birth of the $aA$-population we also lose the control over this.

**Step 1: Upper bound on $\Sigma(t)$** We show that, on the time interval $t \in [\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK^a}]$, there exists a constant, $M_\Sigma > 0$, such that the sum process $\Sigma(t)$ does not exceed the level $\bar{n}_A + 3M_\Sigma(x^2i^2)^{1+\alpha}$, with high probability:

**Proposition 3.2** For all $M > 0$ and $0 \leq i \leq \left\lfloor -\ln(e^{1/4-\alpha})/\ln(x) \right\rfloor$, let

$$\tau_{\Sigma,M}^{\alpha} \equiv \inf \{ t > \tau_{aA}^{i-} : \Sigma(t) - \bar{n}_A \geq 3M(x^2i^2)^{1+\alpha} \}.$$  \hspace{1cm} (3.15)

Then there exists a constant $M_\Sigma > 0$ such that

$$\mathbb{P}[\tau_{\Sigma,M}^{\alpha} < \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK^a}] = o(K^{-1}).$$  \hspace{1cm} (3.16)

To prove Proposition 3.2, we define the difference process between $\Sigma(t)K$ and $\bar{n}_AK$ and couple it with a birth–death-immigration process. We show that this process jumps up with probability less than $\frac{1}{2}$ and show that the probability that the process reaches the level $3M_\Sigma(x^2i^2)^{1+\alpha}K$ before going to zero, is very small. Then we show that the process returns many times to zero until it reaches the level $3M_\Sigma(x^2i^2)^{1+\alpha}K$ and calculate the time for one such return.

**Remark 4** This is only a coarse bound on the sum process $\Sigma(t)$ but with our initial conditions we are not able to get a finer one. After Step 7 we have enough information to refine it but for the iteration this upper bound suffices.

**Step 2: Upper bound on $n_{aa}(t)$** An upper bound on $n_{aa}(t)$ is obtained similarly as in Step 1. Let

$$\gamma_{\Delta} = \frac{f + \Delta}{4\bar{n}_A(f + \Delta)}.$$  \hspace{1cm} (3.17)

We show that, on the time interval $t \in [\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK^a}]$, there exists a constant, $M_{aa} > 0$, such that the $aa$-population does not exceed the level $\gamma_{\Delta}x^2i^2 + 3M_{aa}(x^2i^2)^{1+\alpha}$, with high probability:

**Proposition 3.3** For all $M > 0$ and $0 \leq i \leq \left\lfloor -\ln(e^{1/4-\alpha})/\ln(x) \right\rfloor$, let

$$\tau_{\alpha,aa,M} \equiv \inf \{ t > \tau_{aA}^{i-} : n_{aa}(t) - \gamma_{\Delta}x^2i^2 \geq 3M(x^2i^2)^{1+\alpha} \}.$$  \hspace{1cm} (3.18)

Then there exists a constant, $M_{aa} > 0$, such that

$$\mathbb{P}[\tau_{\alpha,aa,M} < \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK^a}] = o(K^{-1}).$$  \hspace{1cm} (3.19)

The proof is similar to the one of Proposition 3.2.
Step 3: Lower bound on $\Sigma(t)$ With Propositions 3.2 and 3.3, we can bound $\Sigma(t)$ from below. We show that on the time interval $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{i(1)}- \wedge e^{VK^a}\right]$ the sum process does not drop below $\tilde{n}_A - \frac{A + \partial}{c n_A} y_\Delta x^2 i^2 \epsilon^2 = -3 M \Sigma(x^2 i^2 \epsilon^2)^{1+\alpha}$, with high probability:

**Proposition 3.4** For all $M > 0$ and $0 \leq i \leq \left\lfloor \frac{-\ln(e K^{1/4-\alpha})}{\ln(x)} \right\rfloor$, let

$$\tau_{\Sigma,M}^{i} \equiv \inf \left\{ t > \tau_{aA}^{i-} : \Sigma(t) - \left(\tilde{n}_A - \frac{A + \partial}{c n_A} y_\Delta x^2 i^2 \epsilon^2\right) \leq -3 M (x^2 i^2 \epsilon^2)^{1+\alpha}\right\}.$$  

Then there exists a constant, $M_\Sigma > 0$, such that

$$P[\tau_{\Sigma,M}^{i} < \tau_{aA}^{i+} \wedge \tau_{aA}^{i(1)}- \wedge e^{VK^a}] = o(K^{-1}).$$  

(3.20)

The proof is similar to those of Propositions 3.2 and 3.3.

Step 4: Lower and upper bound on $n_{AA}(t)$ Since we now have bounded the processes $n_{aa}(t), n_{A}(t)$ and $\Sigma(t)$ from above and below (for $n_{aa}(t)$ it suffices to set the lower bound to zero), it is easy to get a lower and an upper bound on $n_{AA}(t)$ on the time interval $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{i(1)}- \wedge e^{VK^a}\right]$. There exists a constant, $M_{AA} > 0$, such that with high probability $n_{AA}(t)$ does not drop below $\tilde{n}_A - x^i \epsilon - M_{AA} x^2 i^2 \epsilon^2$ (Proposition 3.5), and does not exceed the level $\tilde{n}_A - x^{i(1)} \epsilon - M_{AA} (x^2 i^2 \epsilon^2)^{1+\alpha}$ (Proposition 3.6):

**Proposition 3.5** For all $M > 0$ and $0 \leq i \leq \left\lfloor \frac{-\ln(e K^{1/4-\alpha})}{\ln(x)} \right\rfloor$, let

$$\tau_{AA,M}^{2i} \equiv \inf \left\{ t > \tau_{aA}^{i-} : n_{AA}(t) - (\tilde{n}_A - x^i \epsilon) \leq -M x^2 i^2 \epsilon^2\right\}.$$  

Then there exists a constant $M_{AA} > 0$ such that

$$P[\tau_{AA,M}^{2i} < \tau_{aA}^{i+} \wedge \tau_{aA}^{i(1)}- \wedge e^{VK^a}] = o(K^{-1}).$$  

(3.23)

**Proposition 3.6** For all $M > 0$ and $0 \leq i \leq \left\lfloor \frac{-\ln(e K^{1/4-\alpha})}{\ln(x)} \right\rfloor$, let

$$\tau_{AA,M}^{i} \equiv \inf \left\{ t > \tau_{aA}^{i-} : n_{AA}(t) - (\tilde{n}_A - x^{i+1} \epsilon) \geq M (x^2 i^2 \epsilon^2)^{1+\alpha}\right\}.$$  

Then there exists a constant, $M_{AA} > 0$, such that

$$P[\tau_{AA,M}^{i} < \tau_{aA}^{i+} \wedge \tau_{aA}^{i(1)}- \wedge e^{VK^a}] = o(K^{-1}).$$  

(3.25)

Step 5: Decay of $n_{aA}(t)$ We now have upper and lower bounds for all the single processes, for $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{i(1)}- \wedge e^{VK^a}\right]$. Using these bounds, we prove that $n_{aA}(t)$ has the tendency to decrease on a given time interval. We show that, with high probability, $n_{aA}(t)$, restarted at $x^i \epsilon$ (i.e. we set $\tau_{aA}^{i-} = 0$), hits the level $x^{i+1} \epsilon$ before it reaches the level $x^i \epsilon + x^2 i^2 \epsilon^2$. 

© Springer
Proposition 3.7 There exists a constant $C > 0$ such that, for all $0 \leq i \leq \left\lfloor \frac{-\ln(e K^{1/4-\alpha})}{\ln(x)} \right\rfloor$

$$\mathbb{P}\left[ \tau_{aA}^{i+1} < \tau_{aA}^{(i+1)-} \mid n_{aA}(0) = x^i \epsilon \right] \leq K^{1/4-\alpha} \exp(-C K^{1/4+3\alpha}).$$ \hspace{1cm} (3.26)

For the proof we couple $n_{aA}(t)$ with majorising and minorising birth–death-immigration processes and show that these processes jump up with probability less than $\frac{1}{2}$. This way we prove that with high probability $n_{aA}(t)$ reaches $x^{i+1} \epsilon$ before going back to $x^i \epsilon + x^{2i} \epsilon^2$.

Step 6: Decay of $n_{aA}(t)$ This is the step where we see that $n_{aA}(t)$ decays like a function $f(t) = \frac{1}{t}$. Precisely, it is shown that the time which the $aA$-population needs to decrease from $x^i \epsilon$ to $x^{i+1} \epsilon$ is of order $\frac{1}{x^i \epsilon}$.

Proposition 3.8 Let

$$\theta_i(aA) \equiv \inf\{t \geq 0 : n_{aA}(t) \leq x^{i+1} \epsilon \mid n_{aA}(0) = x^i \epsilon\},$$ \hspace{1cm} (3.27)

be the decay time of $n_{aA}(t)$ on the time interval $t \in [\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK^{a}}]$. Then there exist positive constants, $C_l, C_u$, and a constant $M > 0$, such that for all $0 \leq i \leq \left\lfloor \frac{-\ln(e K^{1/4-\alpha})}{\ln(x)} \right\rfloor$

$$\mathbb{P}\left[ \frac{C_u}{x^i \epsilon} \geq \theta_i(aA) \geq \frac{C_l}{x^i \epsilon} \right] \geq 1 - \exp(-MK^{1/2+2\alpha}).$$ \hspace{1cm} (3.28)

To prove this proposition we calculate an upper and a lower bound on the decay time of the majorising resp. minorising processes obtained in Step 5 which are of the same order. Precisely, we estimate the number of jumps the processes make until they reach $x^{i+1} \epsilon$, and the time of one jump.

Step 7: Decay and decay time of $n_{aa}(t)$ To carry out the iteration, we have to ensure that, on a given time interval, the $aa$-population decreases below the upper bound needed for the next iteration step. We show that $n_{aa}(t)$ decreases from $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$ to $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$, and stays smaller than $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$ when $n_{aa}(t)$ reaches $x^{i+1} \epsilon$.

Proposition 3.9 For $t \in [\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK^{a}}]$, the process $n_{aa}(t)$ decreases from $x^i \epsilon$ to $x^{i+1} \epsilon$ and $n_{aa}(t)$ decreases from $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$ below the $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$-level.

The proof of this proposition has three parts: first, as in Step 5, we show that $n_{aa}(t)$ has the tendency to decrease and that it reaches $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$ before going back to $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$. The second part is similar to Step 6, where we estimate the number of jumps and the duration of one jump of the process. In the last part we show, as in Step 2, that the process stays below $\gamma_{\Delta x^2 i^2 \epsilon^2 + M_{aa}(x^2 i^2 \epsilon^2)^{1+\alpha}}$ until $n_{aA}(t)$ hits the level $x^{i+1} \epsilon$ and the next iteration step starts.
Step 8: Growth and growth time of $\Sigma(t)$ Similarly to Step 7, we also have to ensure that the sum process increases from the level $\tilde{n}_A - \frac{\Delta + \theta}{\epsilon_n A} \gamma A x^{2i} \epsilon^2$ to $\tilde{n}_A - \frac{\Delta + \theta}{\epsilon_n A} \gamma A x^{2i+2} \epsilon^2$ on a given time interval and is greater than $\tilde{n}_A - \frac{\Delta + \theta}{\epsilon_n A} \gamma A x^{2i+2} \epsilon^2 - M \Sigma(x^{2i+2} \epsilon^2)^{1+\alpha}$ when the $aA$-population reaches the level $x^{i+1} \epsilon$.

**Proposition 3.10** While $n_{aA}$ decreases from $x^i \epsilon$ to $x^{i+1} \epsilon$, the sum process $\Sigma(t)$ increases from $\tilde{n}_A - \frac{\Delta + \theta}{\epsilon_n A} \gamma A x^{2i} \epsilon^2 - M \Sigma(x^{2i} \epsilon^2)^{1+\alpha}$ to $\tilde{n}_A - \frac{\Delta + \theta}{\epsilon_n A} \gamma A x^{2i+2} \epsilon^2$ and stays above $\tilde{n}_A - \frac{\Delta + \theta}{\epsilon_n A} \gamma A x^{2i+2} \epsilon^2 - M \Sigma(x^{2i+2} \epsilon^2)^{1+\alpha}$ until the $aA$-population hits the $x^{i+1} \epsilon$-level.

The proof uses the same three parts as described in the proof of Proposition 3.9.

**Final step: Total decay time of $n_{aA}(t)$** We iterate Step 1 to 8 until $i = \left\lfloor \frac{-\ln(\epsilon K^{1/4-\alpha})}{\ln(\alpha)} \right\rfloor$, the value for which $n_{aA}(t)$ is of order $K^{-1/4+\alpha}$. Finally, we sum up the decay time of the $aA$-population in each iteration step and get the desired result (Theorem 3.2 (ii)). Moreover, we ensure the upper bound on the mutation probability $\mu_K$ in Assumption (C) (3.1) is satisfied.

## 4 The deterministic system

### 4.1 The large population approximation

The main ingredient for the second phase is the deterministic system, since we know from Fournier and Méléard (2004) or Collet et al. (2013) that, for large populations, the behaviour of the stochastic process is close to the solution of a deterministic equation. Thus we analyse it here.

**Proposition 4.1** (Proposition 3.2 in Collet et al. 2013) Let $T > 0$ and $C \subset \mathbb{R}_+^3$ be compact. Assume that the initial condition $\frac{1}{K}(N_{aa}^0, N_{aA}^0, N_{AA}^0)$ converges almost surely to a deterministic vector $(x_0, y_0, z_0) \in C$ when $K$ goes to infinity. Let $(x(t), y(t), z(t)) = \phi(t; (x_0, y_0, z_0))$ denote the solution to

\[
\dot{\phi}(t; (x_0, y_0, z_0)) = \begin{pmatrix}
\tilde{b}_{aa}(x(t), y(t), z(t)) - \tilde{d}_{aa}(x(t), y(t), z(t)) \\
\tilde{b}_{aA}(x(t), y(t), z(t)) - \tilde{d}_{aA}(x(t), y(t), z(t)) \\
\tilde{b}_{AA}(x(t), y(t), z(t)) - \tilde{d}_{AA}(x(t), y(t), z(t))
\end{pmatrix}
=: X(x(t), y(t), z(t)),
\]  

(4.1)

where

\[
\tilde{b}_{aa}(x(t), y(t), z(t)) = \frac{(f_{aa}x(t) + \frac{1}{2}f_{AA}y(t))^2}{(f_{aa}x(t) + f_{A}y(t) + f_{AA}z(t))},
\]
\[
\tilde{a}_{aa}(x(t), y(t), z(t)) = \frac{c_{aa,aa}x(t)}{c_{aa,aa}x(t) + c_{aa,aA}y(t) + c_{aa,AA}z(t)},
\]
and similar expression for the $aA$- and $AA$-type. Then, for all $T > 0$,
\[ \lim_{K \to \infty} \sup_{t \in [0, T]} |n_{uv}(t) - \phi_{uv}(t; (x_0, y_0, z_0))| = 0, \text{ a.s.,} \tag{4.2} \]
for all $uv \in \{aa, aA, AA\}$.

Thus, to understand the behaviour of the process we have to analyse the deterministic system (4.1) above. The vector field (4.1) of the model we consider is given by
\[
X(x, y, z) = X_\Delta(x, y, z) = \begin{pmatrix}
    f \left( \frac{(x + \frac{1}{2}y)^2}{x + y + z} \right) - (D + \Delta + c(x + y + z))x \\
    2f \left( \frac{(x + \frac{1}{2}y)(z + \frac{1}{2}y)}{x + y + z} \right) - (D + c(x + y + z))y \\
    f \left( \frac{(z + \frac{1}{2}y)^2}{x + y + z} \right) - (D + c(x + y + z))z
\end{pmatrix}, \tag{4.3}
\]
which has some particular properties:

**Theorem 4.1** Assume (A)+(B) and let $\epsilon > 0$, then

(i) the vector field (4.3) has the unstable fixed point $n_{aa} \equiv (\bar{n}_a, 0, 0)$ and the stable fixed point $n_{AA} \equiv (0, 0, \bar{n}_A)$,

(ii) the Jacobian matrix at the unstable fixed point $n_{aa}$, $D X_\Delta(n_{aa})$, has two negative and one positive eigenvalues,

(iii) the Jacobian matrix at the stable fixed point $n_{AA}$, $D X_\Delta(n_{AA})$, has two negative and one zero eigenvalues,

(iv) for $0 < \varrho < \varphi$, and as soon as the $aA$-population decreased to an $\epsilon$-level, then
\[
\frac{2\bar{n}_A(f + \Delta)}{(f + \Delta) + \varrho} \leq n_{aA}(t) \leq \frac{2\bar{n}_A(f + \Delta)}{(f + \Delta - \varrho)t + \frac{2\bar{n}_A(f + \Delta)}{\epsilon}}. \tag{4.4}
\]

There is also a biological explanation for the behaviour of $n_{aA}(t)$ described in Theorem 4.1 (iv). Since the $A$-allele is the fittest and dominant one and because of the phenotypic viewpoint the $aA$-population is as fit as the $AA$-population and both die with the same rate. The $aA$-population only decreases because of the disadvantage in reproduction due to the less fit, decreasing $aa$-population. Observe that Theorem 4.1 (i) + (ii) also holds in the model of Collet et al. (2013) (cf. Proposition 3.3 therein) but the Jacobian matrix of their fixed point $n_{AA}$ has three negative eigenvalues and thus they get the exponential decay of $n_{aA}(t)$. The behaviour of solutions of the deterministic system can be analysed using the following result of Collet et al. (2013):

**Theorem 4.2** (Theorem C.2 in Collet et al. 2013) Let $\zeta = u_A - u_a$ be the variation of the allelic trait. Suppose it is non zero and of small enough modulus. If $\zeta \frac{dS_{aA,aa}}{d\zeta}(0) > 0$ then the fixed point $n_{aa}$ is unstable and we have fixation for the macroscopic dynamics.

More precisely, there exists an invariant stable curve $\Gamma_\zeta$ which joins $n_{aa}$ to $n_{AA}$. Moreover there exists an invariant tubular neighbourhood $\mathcal{V}$ of $\Gamma_\zeta$ such that the orbit of any initial condition in $\mathcal{V}$ converges to $n_{AA}$.

If $\zeta \frac{dS_{aA,aa}}{d\zeta}(0) < 0$ the fixed point $n_{aa}$ is stable and the mutant disappears in the macroscopic dynamics.
Their proof works as follows. First they consider the unperturbed version $X_0$ of the vector field (4.1) in the case of neutrality between the alleles $A$ and $a$. That is $f_{u_1 u_2} = f$, $D_{u_1 u_2} = D$ and $C_{u_1 u_2, v_1 v_2} = c$, for $u_1 u_2, v_1 v_2 \in \{aa, aA, AA\}$. They get that this system has a line of fixed points $\Gamma_0$ which is transversally hyperbolic. Afterwards they consider the system $X_\zeta$ with small perturbations $\zeta$. From Theorem 4.1 in Hirsch et al. (1977) (p. 39) they deduce that there exists an attractive and invariant curve $\Gamma_\zeta$, converging to $\Gamma_0$, as $\zeta \to 0$. Hence, there is a small enough tubular neighbourhood $\mathcal{V}$ of $\Gamma_0$ such that $\Gamma_\zeta$ is contained in $\mathcal{V}$ and attracts all orbits with initial conditions in $\mathcal{V}$ (cf. Fig. 3). To show that the orbit of any initial condition on $\Gamma_\zeta$ converges to one of the two fixed points $n_{aa}$ or $n_{AA}$, one have to ensure that the vector field does not vanish on $\Gamma_\zeta$, except for these two fixed points. Since the curve $\Gamma_\zeta$ is attractive, it is equivalent to look for the fixed points in the tubular $\mathcal{V}$. For finding the zero points in $\mathcal{V}$, Collet et al. (2013) use linear combinations of the left eigenvectors of $DX_0(\Gamma_0(v))'$, with the perturbed vector field $X_\zeta$. First they quote to zero the two linear combinations with the eigenvectors which span the stable affine subspace. By the implicit function Theorem, they get a curve which contains all possible zeros in $\mathcal{V}$ (cf. Proposition B.2 in Collet et al. 2013). Then they consider the last linear combination and look for the points on the received curve where it vanish. Under the conditions that the derivative of the third linear combination at the point $\bar{n}_A$ is non zero and does not vanishes between the fixed points $\pm \bar{n}_A$ they get that $X_\zeta$ has only two zeros in a tubular neighbourhood of $\Gamma_0$ (cf. Theorem B.4 in Collet et al. 2013). We have to do some extra work to get the same result for the model with dominant $A$-allele since the derivative of the third linear combination in our model, described above, is zero at the point $\bar{n}_A$. But by an easy calculation (see (5.31) and (5.32)) we can indeed prove that this point is an isolated zero and we can deduce from Theorem 4.2 the following corollary, which is the main result we need about the dynamical system.

**Corollary 4.1** Let $\Delta \neq 0$ small enough.

(i) The attracting and invariant curve $\Gamma_\Delta$ of the perturbed vector field $X_\Delta$ (4.1) contained in the positive quadrant, is the piece of unstable manifold between the equilibrium points $n_{aa}$ and $n_{AA}$.

(ii) There exists an invariant tubular neighbourhood $\mathcal{V}$ of $\Gamma_\Delta$ such that the orbit of any initial condition in $\mathcal{V}$ converges to the equilibrium point $n_{AA}$.

Hence, if we start the process in a neighbourhood of the unstable fixed point $n_{aa}$, it will leave this neighbourhood in finite time and converge to a neighbourhood of the stable fixed point $n_{AA}$.

5 Proofs of Theorem 4.1 and the main theorems

5.1 Analysis of the deterministic system

Because of Proposition 4.1 we have to analyse the deterministic system (4.1) (a simulation is shown in Fig. 1).
Proof of Theorem 4.1 In the following we consider the differential equations of \(n_{aa}(t), n_{A}(t)\) and \(n_{AA}(t)\), given by (4.3):

\[
\dot{n}_{aa}(t) = f\left(\frac{n_{aa}(t) + \frac{1}{2}n_{A}(t)}{\Sigma(t)}\right)^2 - n_{aa}(t)(D + \Delta + c \Sigma(t)),
\]

(5.1)

\[
\dot{n}_{A}(t) = 2f\left(\frac{n_{aa}(t) + \frac{1}{2}n_{A}(t)}{\Sigma(t)}\right)\left(\frac{n_{AA}(t) + \frac{1}{2}n_{A}(t)}{\Sigma(t)}\right) - n_{A}(t)(D + c \Sigma(t)),
\]

(5.2)

\[
\dot{n}_{AA}(t) = f\left(\frac{n_{AA}(t) + \frac{1}{2}n_{A}(t)}{\Sigma(t)}\right)^2 - n_{AA}(t)(D + c \Sigma(t)).
\]

(5.3)

The fixed points

By summing (5.1) to (5.3) first, and two times (5.3) and (5.2), we see that the vector field (4.3) vanishes for the points \(n_{aa}\) and \(n_{AA}\). The Jacobian matrix at the fixed point \(n_{aa}\) is given by

\[
DX_{\Delta}((\bar{n}_a, 0, 0)) = \begin{pmatrix} -f + D + \Delta & -f + D + \Delta & -2f + D + \Delta \\ 0 & \Delta & 2f \\ 0 & 0 & -f + \Delta \end{pmatrix}.
\]

(5.4)

The matrix has the three eigenvalues \(\lambda_1 = -(f - D - \Delta), \lambda_2 = \Delta\) and \(\lambda_3 = -(f - \Delta)\). For \(\Delta\) small enough and from Assumption (A2) we know that \(\lambda_1, \lambda_3 < 0\), whereas \(\lambda_2 > 0\). Thus the fixed point \(n_{aa}\) is unstable.

The Jacobian matrix at the fixed point \(n_{AA}\) is given by

\[
DX_{\Delta}((0, 0, \bar{n}_A)) = \begin{pmatrix} -f - \Delta & 0 & 0 \\ 2f & 0 & 0 \\ -2f + D & -f + D & -f + D \end{pmatrix},
\]

(5.5)

it has the three eigenvalues \(\lambda_1 = -f - \Delta < 0, \lambda_2 = 0\) and \(\lambda_3 = -(f - D) < 0\), from Assumption (A2). The fact that one of the eigenvalues is zero is the main novel feature of this system compared to that count in Collet et al. (2013). Because of the zero eigenvalue, \(n_{AA}\) is a non-hyperbolic equilibrium point of the system and linearization fails to determine its stability properties. Instead, we use the result of the center manifold theory (Hirsch et al. 1977; Perko 2001) that asserts that the qualitative behaviour of the dynamical system in a neighbourhood of the non-hyperbolic critical point \(n_{AA}\) is determined by its behaviour on the center manifold near \(n_{AA}\).

Theorem 5.1 (The Local Center Manifold Theorem 2.12.1 in Perko 2001) Let \(f \in C^r(E)\), where \(E\) is an open subset of \(\mathbb{R}^n\) containing the origin and \(r \geq 1\). Suppose that \(f(0) = 0\) and \(Df(0)\) has \(c\) eigenvalues with zero real parts and \(s\) eigenvalues with negative real parts, where \(c + s = n\). Then the system \(\dot{z} = f(z)\) can be written in diagonal form

\[
\dot{x} = Cx + F(x, y)
\]

\[
\dot{y} = Py + G(x, y),
\]

(5.6)
where \( z = (x, y) \in \mathbb{R}^c \times \mathbb{R}^s \), \( C \) is a \( c \times c \)-matrix with \( c \) eigenvalues having zero real parts, \( P \) is a \( s \times s \)-matrix with \( s \) eigenvalues with negative real parts, and \( F(0) = G(0) = 0 \), \( DF(0) = DG(0) = 0 \). Furthermore, there exists \( \delta > 0 \) and a function, \( h \in C^r(N_\delta(0)) \), where \( N_\delta(0) \) is the \( \delta \)-neighbourhood of \( 0 \), that defines the local center manifold and satisfies:

\[
Dh(x)[Cx + F(x, h(x))] - Ph(x) - G(x, h(x)) = 0, \tag{5.7}
\]

for \( |x| < \delta \). The flow on the center manifold \( W^c(0) \) is defined by the system of differential equations

\[
\dot{x} = Cx + F(x, h(x)), \tag{5.8}
\]

for all \( x \in \mathbb{R}^c \) with \( |x| < \delta \).

The fact that the center manifold of our system near \( n_{AA} \) has dimension one, simplifies the problem of determining the stability and the qualitative behaviour of the flow on it near the non-hyperbolic critical point. The Local Center Manifold Theorem shows that the non-hyperbolic critical point \( n_{AA} \) is indeed a stable fixed point and that the flow on the center manifold near the critical point behaves like a function \( 1/t \). This can be seen as follows: assume that \( n_{AA}(t) \) has decreased to a level \( \epsilon \). Let \( n_{AA}(t) = z(t), n_{aA}(t) = y(t) \) and \( n_{aa}(t) = x(t) \). By the affine transformation \( n_{AA} \mapsto n_{AA} - \bar{n}_A \) we get a translated system

\[
Y(z, y, x) = \begin{pmatrix}
  f \left( \frac{(z + \bar{n}_A + \frac{1}{2}y)^2}{z + y + x + \bar{n}_A} \right) - (D + c(z + y + x + \bar{n}_A))(z + \bar{n}_A) \\
  2f \left( \frac{(z + \bar{n}_A + \frac{1}{2}y)(x + \frac{1}{2}y)}{z + y + x + \bar{n}_A} \right) - (D + c(z + y + x + \bar{n}_A))y \\
  f \left( \frac{(x + \frac{1}{2}y)^2}{z + y + x + \bar{n}_A} \right) - (D + \Delta + c(z + y + x + \bar{n}_A))x
\end{pmatrix}, \tag{5.9}
\]

which has a critical point at the origin. The Jacobian matrix of \( Y \) at the fixed point \((0, 0, 0)\) is given by

\[
DY((0, 0, 0)) = \begin{pmatrix}
  -(f - D) & -(f - D) & -(2f - D) \\
  0 & 0 & 2f \\
  0 & 0 & -(f + \Delta)
\end{pmatrix}, \tag{5.10}
\]

which has the eigenvalues \( \lambda_1 = -(f - D), \lambda_2 = 0 \) and \( \lambda_3 = -(f + \Delta) \) with corresponding eigenvectors

\[
EV_1 = \begin{pmatrix}
  1 \\
  0
\end{pmatrix}, \quad EV_2 = \begin{pmatrix}
  -1 \\
  1
\end{pmatrix}, \quad \text{and} \quad EV_3 = \begin{pmatrix}
  fD + \Delta(2f - D) \\
  (f + \Delta)(D + \Delta)
\end{pmatrix}. \tag{5.11}
\]
We can find a transformation

\[
T = \begin{pmatrix}
1 & 1 & \frac{D}{D + \Delta}
\end{pmatrix},
\]

(5.12)

that transforms \(DY((0, 0, 0))\) into a block matrix. The change of variables

\[
\begin{pmatrix}
\frac{\dot{z}}{y} \\
\frac{\dot{y}}{x}
\end{pmatrix} = T \begin{pmatrix}
\frac{z}{y} \\
\frac{x}{y}
\end{pmatrix},
\]

(5.13)

casts the system into the form

\[
\dot{\tilde{z}} = \dot{\tilde{z}} + \frac{D}{D + \Delta} \tilde{x}, \quad \dot{\tilde{y}} = \dot{\tilde{y}} + \frac{2f}{f + \Delta} \tilde{x}, \quad \dot{\tilde{x}} = \tilde{x}.
\]

(5.14)

Let \(h(\tilde{y})\) be the local center manifold. We approximate \(h(\tilde{y}) = (h_1(\tilde{y})\tilde{y}^2 + O(\tilde{y}^3)\) and substitute the series expansions into the center manifold equation (5.7)

\[
\begin{pmatrix}
2h_1 \tilde{y} \\
2h_2 \tilde{y}
\end{pmatrix} \dot{h}(h_1 \tilde{y}^2, \tilde{y}, h_2 \tilde{y}^2) = \begin{pmatrix}
\dot{\tilde{z}}(h_1 \tilde{y}^2, \tilde{y}, h_2 \tilde{y}^2) \\
\dot{\tilde{x}}(h_1 \tilde{y}^2, \tilde{y}, h_2 \tilde{y}^2)
\end{pmatrix}.
\]

(5.15)

To determine \(h_1\) and \(h_2\) we compare the coefficients of the same powers of \(\tilde{y}\). We first consider \(\dot{\tilde{y}}(h_1 \tilde{y}^2, \tilde{y}, h_2 \tilde{y}^2)\) and get that the coefficient of \(\tilde{y}\) is zero. The first coefficients which are not zero at the right hand side are the ones of \(\tilde{y}^2\). Thus we have to compare these with the coefficient of \(\tilde{y}^2\) on the left hand side which is zero as mentioned above. Hence, the coefficient of \(\dot{\tilde{x}}(h_1 \tilde{y}^2, \tilde{y}, h_2 \tilde{y}^2)\) of \(\tilde{y}^2\), which is \(-h_2(f + \Delta) + \frac{f}{4nA}\), equals zero and we get \(h_2 = \frac{f}{4nA(f + \Delta)}\). From the coefficient of \(\dot{\tilde{z}}(h_1 \tilde{y}^2, \tilde{y}, h_2 \tilde{y}^2)\) of \(\tilde{y}^2\), which is \(Dh_1 - \frac{f(\Delta(D + \Delta))}{4nA(f + \Delta)}\), we get \(h_1 = \frac{f\Delta(D + \Delta)}{4nA(f + \Delta)(D + \Delta)}\). Thus the local center manifold is given by

\[
h(\tilde{y}) = \left(\frac{f\Delta^2}{4nA D(f + \Delta)(D + \Delta)}\right) \tilde{y}^2 + O(\tilde{y}^3).
\]

(5.16)

Substitution of this result into \(\dot{\tilde{y}}\) yields the flow on the local center manifold

\[
\dot{\tilde{y}} = -\frac{f\Delta}{2nA(f + \Delta)} \tilde{y}^2 + O(\tilde{y}^3).
\]

(5.17)

We can bound the solution of this equation with initial condition \(y(0) = \epsilon\) by

\[
\frac{2\bar{n}_A(f + \Delta)}{(f \Delta + \epsilon)t + \frac{2nA(f + \Delta)}{\epsilon}} \leq \tilde{y}(t) \leq \frac{2\bar{n}_A(f + \Delta)}{(f \Delta - \epsilon)t + \frac{2nA(f + \Delta)}{\epsilon}}.
\]

(5.18)
for $0 < \varrho < f \Delta$. Thus we see that $n_{AA}$ is a stable fixed point and $n_{AA}(t)$ approaches this fixed point like a function $\frac{1}{2}$.

In contrast to the model of Collet et al. (2013), we see in the proof that $n_{AA}(t)$ does not dies out exponentially fast. Instead it decays like a function $f(t) = \frac{1}{2}$, as soon as $n_{AA} + n_{AA} \approx \bar{n}_A$ and thus survives a longer time in the population (cf. Figs. 1, 2). Up to now, we know how the deterministic system evolves near its fixed points. Namely:

- if we start the process in a neighbourhood of the unstable fixed point $n_{aa}$, it will leave this neighbourhood in finite time.
- whereas, if the process is in a neighbourhood of the stable fixed point $n_{AA}$, it converges to this point $n_{AA}$, but the convergence is slower than in the model of Collet et al. (2013).

We now turn to the analysis of the behaviour between these points.

**Behaviour between the fixed points** We show that the deterministic system (4.1) moves from a neighbourhood of the unstable fixed point $n_{aa}$ to a neighbourhood of the stable fixed point $n_{AA}$ (Corollary 4.1).

**Proof of Corollary 4.1** The proof is similar to the proof of Theorem C.2 in Collet et al. (2013). It only differs in the last step. The general unperturbed vector field $X_0$ in the case of neutrality between the $a$- and $A$-alleles is given by:

$$X_0 = \begin{pmatrix}
    f\left(\frac{x+\frac{1}{2}y}{x+y+z}\right)^2 - (D + c(x + y + z))x \\
    2f\left(\frac{x+\frac{1}{2}y+z+\frac{1}{2}y}{x+y+z}\right)^2 - (D + c(x + y + z))y \\
    f\left(\frac{z+\frac{1}{2}y}{x+y+z}\right)^2 - (D + c(x + y + z))z
\end{pmatrix}, \quad (5.19)$$

The content of Theorem B.1 in Collet et al. (2013) is that $X_0 (5.19)$ has a line of fixed points given by,

$$\Gamma_0(v) = \begin{pmatrix}
    \frac{(v-\bar{n}_A)^2}{4\bar{n}_A} \\
    -\frac{v^2-\bar{n}_A^2}{2\bar{n}_A} \\
    \frac{(v+\bar{n}_A)^2}{4\bar{n}_A}
\end{pmatrix}, \quad v \in [-\bar{n}_A, \bar{n}_A], \quad (5.20)$$

and that the differential of the vector field $X_0$ at each point of the curve $\Gamma_0$, $DX_0(\Gamma_0(v))$, has the three eigenvectors

$$e_1(v) = \begin{pmatrix}
    \frac{(v-\bar{n}_A)^2}{4\bar{n}_A} \\
    -\frac{v^2-\bar{n}_A^2}{2\bar{n}_A} \\
    \frac{(v+\bar{n}_A)^2}{4\bar{n}_A}
\end{pmatrix}, \quad e_2(v) = \begin{pmatrix}
    \frac{v-\bar{n}_A}{2\bar{n}_A} \\
    -\frac{v}{\bar{n}_A} \\
    \frac{v+\bar{n}_A}{2\bar{n}_A}
\end{pmatrix}, \quad e_3(v) = \frac{1}{2\bar{n}_A} \begin{pmatrix}
    1 \\
    -2 \\
    1
\end{pmatrix}, \quad (5.21)$$

with respective eigenvalues $-(f-D) < 0$, $0$, and $-f < 0$. $DX_0(\Gamma_0(v))^t$ has the three eigenvalues, $-f + D$, 0, and $-f$, with corresponding eigenvectors.
Survival of a recessive allele in a Mendelian diploid model

Fig. 5  The curves $\Gamma_0$, $\Gamma_{\Delta}$ and the tube $\mathcal{V}$ in the perturbed vector field $X_\Delta$ (Simulation by Loren Coquille)

$$\beta_1(v) = \frac{1}{\bar{n}_A} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \beta_2(v) = \begin{pmatrix} -\frac{v+\bar{n}_A}{\bar{n}_A} \\ \frac{v}{\bar{n}_A} \\ \frac{v-\bar{n}_A}{\bar{n}_A} \end{pmatrix}, \quad \beta_3(v) = \begin{pmatrix} \frac{(v+\bar{n}_A)^2}{2\bar{n}_A} \\ \frac{v^2-\bar{n}_A^2}{2\bar{n}_A} \\ \frac{(v-\bar{n}_A)^2}{2\bar{n}_A} \end{pmatrix}, \quad (5.22)$$

which satisfy, for any $i, j \in \{1, 2, 3\}$ and any $v$,

$$\langle \beta_i(v), e_j(v) \rangle = \delta_{i,j}. \quad (5.23)$$

Next we analyse the asymptotics of the flow associated to the perturbed vector field, $X_\Delta$ (4.3), as $t \to \infty$. From Theorem B.1 in Collet et al. (2013) we know that the curve of fixed points, $\Gamma_0$, is transversally hyperbolic and invariant for the vector field $X_0$. Thus, Theorem 4.1 in Hirsch et al. (1977) implies that, for small enough $\Delta$, there is an attractive curve, $\Gamma_{\Delta}$ that is invariant under $X_\Delta$. Moreover, $\Gamma_{\Delta}$ is regular and converges to $\Gamma_0$, as $\Delta \to 0$. Hence, there is a small tubular neighbourhood, $\mathcal{V}$, of $\Gamma_0$ such that $\Gamma_{\Delta}$ is contained in $\mathcal{V}$ and attracts all orbits with initial conditions in $\mathcal{V}$ (cf. Fig. 5).

We want to study the flow associated to the vector field $X_\Delta$. From the remarks above we know that the curve $\Gamma_{\Delta}$ is attractive for this flow. Thus, it suffices to analyse the flow on the curve. Precisely, we want to show that the vector field does not vanish on $\Gamma_{\Delta}$ except for the two fixed points, $n_{aa}$ and $n_{AA}$. Thus the orbit of any initial condition on $\Gamma_{\Delta}$ converges to one of the two fixed points. It is easier to look for the fixed points in the tube $\mathcal{V}$, which is equivalent to looking for fixed points on $\Gamma_{\Delta}$ because of the
The determinant of the Jacobian matrix of the transformation $(v, a)$ is a diffeomorphism which maps $[−\tilde{\Delta}, \tilde{\Delta}]$ to a closed neighbourhood of the points $\pm \tilde{\Delta}$. There is a number $1/123$ (Proposition B.2 in Collet et al. 2013) and zeros of this linear combination on the curve above.

The third linear combination of the components vanishes. Since Hirsch gives us only a local statement, we consider the curve in local frames. A point $(x, y, z) \in \mathcal{V}$ is represented by the parametrisation

$$M(v, r, s) = \Gamma_0(v) + re_1(v) + se_3(v) = (1 + r)\Gamma_0(v) + s\frac{d^2\Gamma_0(v)}{dv^2}, \quad (5.24)$$

with $v \in [−\tilde{n}_A − \delta, \tilde{n}_A + \delta]$ and $r, s \in [−\delta, \delta]$, with $\delta > 0$ chosen small enough. The determinant of the Jacobian matrix of the transformation $(v, r, s) \mapsto (x, y, z) = M(v, r, s)$ is $-\frac{e_1}{2}$ and thus invertible and does not vanish if $0 < \delta < 1$. Moreover, it is a diffeomorphism which maps $[−\tilde{n}_A − \delta, \tilde{n}_A + \delta] \times [−\delta, \delta]^2$ to a closed neighbourhood of $\mathcal{V}$.

For finding the zero points in $\mathcal{V}$, we use linear combinations of the left eigenvectors $\beta_i$, $i \in \{1, 2, 3\}$, with the perturbed vector field $X_\Delta$. First we look for zeros of the two linear combinations of $X_\Delta$ with the eigenvectors $\beta_1$ and $\beta_3$ which spans the stable affine subspace. By the implicit function Theorem, we obtain a curve which contains all possible zeros in $\mathcal{V}$. Then we consider the last linear combination of $X_\Delta$ with $\beta_2$ and zeros of this linear combination on the curve above.

**Proposition 5.1** (Proposition B.2 in Collet et al. 2013) For any $\delta > 0$ small enough, there is a number $\Delta_0 = \Delta_0(\delta)$ such that, for any $\Delta \in [−\Delta_0, \Delta_0]$, there is a smooth curve $\mathcal{Z}_\Delta = (r_\Delta(v), s_\Delta(v)) \subset \mathbb{R}^2$, depending smoothly on $\Delta$ and converging to 0 when $\Delta$ tends to zero such that, for any $v \in [−\tilde{n}_A − \delta, \tilde{n}_A + \delta]$, we have

$$\langle \beta_1(v), X_\Delta(M(v, r_\Delta(v), s_\Delta(v))) \rangle = \langle \beta_3(v), X_\Delta(M(v, r_\Delta(v), s_\Delta(v))) \rangle = 0. \quad (5.25)$$

Moreover, if a point $(v, r, s)$ with $v \in [−\tilde{n}_A − \delta, \tilde{n}_A + \delta]$, $r$ and $s$ small enough is such that

$$\langle \beta_1(v), X_\Delta(M(v, r, s)) \rangle = \langle \beta_3(v), X_\Delta(M(v, r, s)) \rangle = 0, \quad (5.26)$$

then $(r, s) = (r_\Delta(v), s_\Delta(v))$.

Next, we look for the points of the resulting curve (obtained from Proposition 5.1) where the third linear combination of the components vanishes. Since

$$\langle \beta_2(v), X(0, M(v, r_0(v), s_0(v))) \rangle = \langle \beta_2(v), X(0, M(v, 0, 0)) \rangle \quad = \langle \beta_2(v), X(0, \Gamma_0(v)) \rangle = 0, \quad (5.27)$$

this function vanishes for $\Delta = 0$ in $v$ and we apply the Malgrange preparation theorem (Golubitsky and Guillemin 1973), which provides a representation for the linear combination near $\Delta = 0$:

$$\langle \beta_2(v), X(\Delta, M(v, r_\Delta(v), s_\Delta(v))) \rangle = \Delta^2 h(\Delta, v) + \Delta g(v), \quad (5.28)$$

where $h, g$ are two smooth functions. To show that the third linear combination indeed vanishes only in small neighbourhoods of the points $±\tilde{n}_A$, Collet et al. (2013) use a representation for the function $g$ which is independent of the perturbation $\Delta$. 

\[ Springer \]
Lemma 5.1 (Lemma B.3 in Collet et al. 2013) The function $g$ in (5.28) is given by

$$g(v) = \langle \beta_2(v), \partial_X X_0(\Gamma_0(v)) \rangle. \quad (5.29)$$

Then they ensure that this function has only two zeros, which implies, because of the representation (5.28), for $|\Delta| \neq 0$ small enough, that the perturbed vector field $X_\Delta$ has only the two known fixed points $n_{aa}$ and $n_{AA}$.

Theorem 5.2 (Theorem B.4 in Collet et al. 2013) Assume the function

$$g(v) = \langle \beta_2(v), \partial_X X_0(\Gamma_0(v)) \rangle, \quad (5.30)$$

satisfies $\frac{dg}{dv}(\pm \bar{n}_A) \neq 0$ and does not vanish in $(-\bar{n}_A, \bar{n}_A)$. Then, for $|\Delta| \neq 0$ small enough, the vector field $X_\Delta$ has only two zeros in a tubular neighbourhood of $\Gamma_0$. These zeros are $(n_{aa}(\Delta), 0, 0)$ and $(0, 0, n_{AA}(\Delta))$, with $n_{aa}(\Delta)$ and $n_{AA}(\Delta)$ regular near $\Delta = 0$ and $n_{aa}(0) = n_{AA}(0) = \bar{n}_A$.

While the hypothesis $\frac{dg}{dv}(\pm \bar{n}_A) \neq 0$ does not hold here, the conclusion of Theorem 5.2 remains true. Namely, we have from (4.3) that

$$\partial_X X_\Delta(x, y, z) = \begin{pmatrix} -x \\ 0 \\ 0 \end{pmatrix}, \quad (5.31)$$

and thus

$$g(v) = \langle \beta_2(v), \partial_X X_0(\Gamma_0(v)) \rangle = \begin{pmatrix} \frac{v + \bar{n}_A}{\bar{n}_A} \\ -\frac{v}{\bar{n}_A} \\ -\frac{v - \bar{n}_A}{\bar{n}_A} \end{pmatrix} \cdot \begin{pmatrix} -\frac{(v - \bar{n}_A)^2}{4\bar{n}_A^2} \\ 0 \\ 0 \end{pmatrix} = \frac{(v + \bar{n}_A)(v - \bar{n}_A)^2}{4\bar{n}_A^2}. \quad (5.32)$$

Obviously, $g(\pm \bar{n}_A) = 0$ and $g$ has no other zeros, in particular, it does not vanish in $(-\bar{n}_A, \bar{n}_A)$. Hence, it follows from the representation (5.28) that there is a $\delta > 0$ such that, for $\Delta$ small enough,

$$\langle \beta_2(v), X_\Delta(M(v, r_\Delta(v), s_\Delta(v))) \rangle$$

has only two zeros in $v \in [-\delta - \bar{n}_A, \bar{n}_A + \delta]$. From Theorem 4.1 we get the existence of these two fixed points, which have to be the points $n_{aa}$ and $n_{AA}$. Finally, we need the following lemma, which is analogous to Theorem C.1 in Collet et al. (2013).

Lemma 5.2 (a) The local stable manifold of the unstable fixed point $n_{aa} = (\bar{n}_a, 0, 0)$ intersects the closed positive quadrant only along the line $y = z = 0$.

(b) The local unstable manifold is contained in the curve $\Gamma_\Delta$.

Proof We start by proving (a). From Theorem 4.1 we get the hyperbolicity, thus we can apply Theorem 4.1 in Hirsch et al. (1977). The Jacobian matrix $DX_\Delta((\bar{n}_a, 0, 0))$
(cf. (5.4)) has the three eigenvalues $\lambda_1 = -(f - D - \Delta)$, $\lambda_2 = \Delta$ and $\lambda_3 = -(f - \Delta)$ with corresponding eigenvectors

\[
EV(\lambda_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_1(-\bar{n}_A) 
\]

\[
EV(\lambda_2) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \frac{\Delta}{f - D} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_2(-\bar{n}_A) + O(\Delta) 
\]

\[
EV(\lambda_3) = \frac{1}{2\bar{n}_A} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + \frac{\Delta}{2\bar{n}_A(f - D)} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_3(-\bar{n}_A) + O(\Delta). 
\]

Let $E_{aa}^s(\Delta)$ the two dimensional affine stable subspace spanned by the eigenvectors $EV(\lambda_1)$ and $EV(\lambda_3)$ with origin in $n_{aa}$:

\[
E_{aa}^s(\Delta) = \{ x \in \mathbb{R}^3 | x = (\bar{n}_a, 0, 0)^t + s EV(\lambda_1) + t EV(\lambda_3), \forall s, t \in \mathbb{R} \}. 
\]

Again, by the lamination and permanence condition in Theorem 4.1 in Hirsch et al. (1977), we get the existence of a stable manifold $W_{aa}^{s,loc}$ and an unstable manifold $W_{aa}^{u,loc}$ of the fixed point $n_{aa}$. The local stable manifold $W_{aa}^{s,loc}$ is a piece of regular manifold tangent in $n_{aa}$ to the subspace $E_{aa}^s(\Delta)$. We see that the $x$-axis is invariant for $X_{\Delta}$ and is contained in $W_{aa}^{s,loc}$. From (5.36), we get that $E_{aa}^s(\Delta)$ intersects the closed positive quadrant only along the line $y = z = 0$. Hence, the same is true for $W_{aa}^{s,loc}$ since it is a piece of the subspace $E_{aa}^s(\Delta)$. This shows (a). To show (b), we show that the local unstable manifold $W_{aa}^{u,loc}$ is contained in the closed positive quadrant. This follows because $W_{aa}^{u,loc}$ is tangent to the linear unstable direction in $EV(\lambda_2)$ in $n_{aa}$, which points into the positive quadrant. From Theorem 4.1 in Hirsch et al. (1977) we get that the invariant curve, $\Gamma_{\Delta}$, is unique, thus $W_{aa}^{u,loc} \subset \Gamma_{\Delta}$ and (b) follows by the invariance of the positive quadrant under the flow.

The preceding steps conclude the proof of Corollary 4.1. 

\[
\Box
\]

5.2 Proof of the main theorems in Sect. 3

We carry out the proofs of the main theorems (Sect. 3) in full detail.

The mutant process $A(t) = 2n_{AA}(t) + n_{aA}(t)$ jumps up (resp. down) by rate $b_A$ (resp. $d_A$) given by:

\[
b_A = \frac{2fK}{\Sigma(t)} \left( (n_{AA}(t) + \frac{1}{2}n_{aA}(t))^2 + (n_{AA}(t) + \frac{1}{2}n_{aA}(t)) (n_{aa}(t) + \frac{1}{2}n_{aA}(t)) \right) \\
= fK(2n_{AA}(t) + n_{aA}(t)) = A(t)Kf, 
\]

\[
d_A = 2n_{AA}(t)K(D + c\Sigma(t)) + n_{aA}(t)K(D + c\Sigma(t)) = A(t)K(D + c\Sigma(t)). 
\]
5.2.1 Phase 1: fixation of the mutant population

Recall the stopping times (3.13) and (3.12), when the mutant population $A(t)$ increased to a $\delta$-level and its stopping time of extinction. We show Theorem 3.1:

Proof of Theorem 3.1 We start the population process with a monomorphic $aa$-population which stays in a $\delta/2K$-neighbourhood of its equilibrium $\bar{n}_a K$ and one mutant with genotype $aA$, i.e. $\tau_0 = 0$. Because of (3.1), there will be no further mutation in the process.

Proposition 3.1 states that if the resident population $n_{aa}$ is in a $\delta/2$-neighbourhood of its equilibrium $\bar{n}_a$, then $n_{aa}$ will stay in a $\delta$-neighbourhood for an exponentially long time as long as the mutant population is less than $\delta$. Hence we get that, as long as the mutant population is smaller than $\delta$, the time the process $n_{aa}(t)$ needs to exit from its domain $\bar{n}_a$ is bigger than $e^{VK}$ with probability converging to 1, for some $V > 0$ (cf. Champagnat (2006)) and the dynamics of the mutant population are negligible for $n_{aa}(t)$.

With this knowledge we analyse the fate of the mutants for $t < t_{\bar{\tau}}^{mut} \wedge e^{VK}$. We use the comparison results of birth and death processes (Theorem 2 in Champagnat 2006) to bound the mutant process from below and above.

We denote by $\preceq$ the following stochastic dominant relation: if $P_1$ and $P_2$ are the laws of $\mathbb{R}$-valued processes, we will write $P_1 \preceq P_2$ if we can construct on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ two processes $X^1$ and $X^2$ such that the law of the processes is $\mathcal{L}(X^i) = P_i$, $i \in \{1, 2\}$ and for all $t > 0$, $\omega \in \Omega$:

First we construct a process $A_l(t) \preceq A(t)$ which is the minorising process of the mutant process. This process has the birth and death rates:

$$b_l(t) = A_l(t)Kf, \quad d_l(t) = A_l(t)K[D + c(\bar{n}_a + 2\delta)]. \quad (5.39)$$

$A(t) \preceq A_u(t)$ is the majorising process with rates:

$$b_u(t) = A_u(t)fK, \quad d_u(t) = A_u(t)K[D + c(\bar{n}_a - \delta)]. \quad (5.40)$$

We define the stopping times

$$T_{n/K}^l \equiv \inf \left\{ t \geq 0 : A_l(t) = \frac{n}{K} \right\}, \quad T_{n/K}^u \equiv \inf \left\{ t \geq 0 : A_u(t) = \frac{n}{K} \right\} \quad (5.41)$$

$$\Theta_a \equiv \inf \left\{ t \geq 0 : |n_{aa}(t) - \bar{n}_a| > \delta \right\}, \quad (5.42)$$

which are the first times that the processes $A_l$, resp. $A_u$ reach the level $n/K$ and the exit time of $n_{aa}(t)$ from the domain $[\bar{n}_a - \delta, \bar{n}_a + \delta]$.

Note that both processes $A_l(t)$ and $A_u(t)$ are super-critical. In the following we use the results for super-critical branching processes proven by Champagnat (2006):

Lemma 5.3 (Theorem 4 (b) in Champagnat (2006)) Let $b, d > 0$. For any $K \geq 1$ and any $z \in \mathbb{N}/K$, let $P^K_z$ the law of the $\mathbb{N}/K$-valued Markov linear birth and death process $(\omega_t, t \geq 0)$ with birth and death rates $b$ and $d$ and initial state $z$. Define, for any $\rho \in \mathbb{R}$, on $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$, the stopping time

\[ \square \] Springer
$$T_\rho = \inf\{t \geq 0 : \omega_t = \rho\}. \quad (5.43)$$

Let \((t_K)_{K \geq 1}\) be a sequence of positive numbers such that \(\ln K \ll t_K\). If \(b > d\), for any \(\epsilon > 0\),

\[
\begin{align*}
&\text{(a) } \lim_{K \to \infty} \mathbb{P}^{K}_{1 \frac{K}{\rho}} (T_0 \leq t_K \wedge T_{[\varepsilon K]/K}) = \frac{d}{b}, \\
&\text{(b) } \lim_{K \to \infty} \mathbb{P}^{K}_{1 \frac{K}{\rho}} (T_{[\varepsilon K]/K} \leq t_K) = 1 - \frac{d}{b}.
\end{align*}
\]

(5.44) (5.45)

With respect to Theorem 3.3, we prove the theorem for arbitrary \(\mu_K\). From (3.9) we know that the next mutation occurs with high probability not before a time \(\frac{\rho_{K}}{K\mu_K}\). Then

\[
\begin{align*}
\mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} &\left( T_{[\delta K]/K}^l < T_0^l \wedge \frac{\rho_{K}}{K\mu_K} \wedge \Theta_a \right) \leq \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( \tau_0^{mut} < \tau_{[\delta K]/K}^{mut} \wedge \frac{\rho_{K}}{K\mu_K} \wedge \Theta_a \right) \\
&\leq \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_{[\delta K]/K}^l < T_0^l \wedge \frac{\rho_{K}}{K\mu_K} \wedge \Theta_a \right).
\end{align*}
\]

(5.46)

Using Proposition 3.1, there exists \(V > 0\) such that, for sufficiently large \(K\),

\[
\mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( \tau_1 \wedge e^{VK} < \Theta_a \right) \geq 1 - \delta.
\]

(5.47)

With (5.47) and for \(K\) large enough we can estimate,

\[
\begin{align*}
\mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} &\left( T_{[\delta K]/K}^l < T_0^l \wedge \frac{\rho_{K}}{K\mu_K} \wedge \Theta_a \right) \geq \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_{[\delta K]/K}^l < T_0^l \wedge \frac{\rho_{K}}{K\mu_K} \wedge e^{VK} \right) - \delta \\
&\geq \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( \left\{ T_{[\delta K]/K}^l < T_0^l \wedge \frac{\rho_{K}}{K\mu_K} \right\} \cap \left\{ T_0^l \geq \frac{\rho_{K}}{K\mu_K} \right\} \right) - \delta \\
&\geq \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_{[\delta K]/K}^l < \frac{\rho_{K}}{K\mu_K} \right) - \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_0^l < \frac{\rho_{K}}{K\mu_K} \right) - \delta \\
&\geq \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_{[\delta K]/K}^l < \frac{\rho_{K}}{K\mu_K} \right) - \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_0^l < \frac{\rho_{K}}{K\mu_K} \right) - \delta
\end{align*}
\]

(5.48)

The extinction time for a binary branching process when \(b \neq d\) is given by (cf. page 109 in Athreya and Ney 2011)

\[
\mathbb{P}_n(T_0 \leq t) = \left( \frac{d(1 - e^{-(b-d)t})}{b - de^{-(b-d)t}} \right)^n,
\]

(5.49)

for any \(t \geq 0\) and \(n \in \mathbb{N}\). Under our condition on \(\mu_K\) (3.1) in Theorem 3.3, this implies that

\[
\lim_{K \to \infty} \mathbb{P}^{1 \frac{K}{\rho}}_{\frac{K}{\rho}} \left( T_0^l \leq \frac{\rho_{K}}{K\mu_K} \right) = 0.
\]

(5.50)
Together with Lemma 5.3, we get

$$\lim_{K \to \infty} \mathbb{P}_1 \left[ \tau_0^{mut} \leq \tau_0^{mut} \wedge \frac{\rho}{K \mu_K} \wedge \Theta_a \right] \geq 1 - \frac{d_l}{b_l} - \delta = \frac{\Delta}{f} - \left( \frac{c(M + 1)}{f} + 1 \right) \delta.$$  \hspace{1cm} (5.51)

If we next consider the upper bound of (5.46), we see that

$$P_1 \left[ T_u^{\delta}/K < T_0^{u} \wedge \frac{\rho}{K \mu_K} \wedge \Theta_a \right] \leq P_1 \left[ T_u^{\delta}/K < T_0^{u} \wedge \frac{\rho}{K \mu_K} \right].$$  \hspace{1cm} (5.52)

Similar as in (5.51)

$$\lim_{K \to \infty} \mathbb{P}_1 \left[ \tau_0^{mut} \leq \tau_0^{mut} \wedge \frac{\rho}{K \mu_K} \wedge \Theta_a \right] = \frac{\Delta}{f} + \frac{cM \delta}{f}.$$  \hspace{1cm} (5.53)

Now, we let \( \delta \to 0 \) and get the desired result.

\[ \square \]

5.2.2 Phase 2: invasion of the mutant

If the mutant invades in the resident \( aa \)-population then the first phase ends with a macroscopic mutant population. Especially, we know that the \( aA \)-population is of order \( \delta \) due to its advantage in recombination in contrast to the \( AA \)-population. Applying the Large Population Approximation (Theorem 4.1, Fournier and Méléard 2004) with this initial condition, we get that the behaviour of the process is close to the solution of the deterministic system (4.1), when \( K \) tends to infinity. We approximate the population process by the solution of the dynamical system (4.1). Result (4.2) in Proposition 4.1 is known (see Champagnat and Méléard 2011 or Collet et al. 2013). We use this result only until the \( aA \)-population decreases to an \( \epsilon \)-level.

5.2.3 Phase 3: survival of the recessive \( a \)-Allele

This phase starts as soon as the \( aA \)-population hits the \( \epsilon \) value. At this time we restart the population-process, that means we set the time to zero. The analysis of the deterministic dynamical system up to this point shows that we get the following initial conditions:

\[
n_{aA}(0) = \epsilon, \hspace{1cm} (5.54)
\]
\[
n_{aa}(0) \leq \frac{f}{4 \bar{n}_A(f + \Delta)} \epsilon^2 + M_{aa} \epsilon^{2+\alpha}, \hspace{1cm} (5.55)
\]
\[
|n_{AA}(0) - (\bar{n}_A - \epsilon)| \leq M_{AA} \epsilon^2, \hspace{1cm} (5.56)
\]
\[
|\Sigma(0) - (\bar{n}_A + \frac{\Delta}{c\bar{n}_A} \gamma \epsilon^2)| \leq M_{\Sigma} \epsilon^{2+\alpha}, \hspace{1cm} (5.57)
\]

where \( M_i, i \in \{aa, AA, \Sigma\} \) are constants.
In the following stopping times denoted by \( \tau \) are always stopping times on rescaled processes, whereas stopping times denoted by \( T \) are the stopping times of the corresponding non-rescaled processes.

We transform the birth and death rates of the processes \( N_{aa}(t), N_{Aa}(t), N_{AA}(t) \) and the sum process \( \Sigma(t)K \) in such a way that we can consider them as the birth and death rates of linear birth–death-immigration processes:

\[
\begin{align*}
    b_{\Sigma}(t) & = f\Sigma(t)K, \\
    d_{\Sigma}(t) & = D\Sigma(t)K + c\Sigma^2(t)K + \Delta N_{aa}(t), \\
    b_{aa}(t) & = fN_{aa}(t) \left(1 - \frac{n_{AA}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}n_{aa}^2(t), \\
    d_{aa}(t) & = N_{aa}(t)(D + \Delta + c\Sigma(t)), \\
    b_{aA}(t) & = fN_{aA}(t) \left(1 - \frac{n_{AA}(t)}{2\Sigma(t)}\right) + 2fN_{aa}(t)\frac{n_{AA}(t)}{\Sigma(t)}, \\
    d_{aA}(t) & = N_{aA}(t)(D + c\Sigma(t)), \\
    b_{AA}(t) & = fN_{AA}(t) \left(1 - \frac{n_{aa}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}n_{aa}^2(t), \\
    d_{AA}(t) & = N_{AA}(t)(D + c\Sigma(t)).
\end{align*}
\]

We proceed as described in the outline. Recall the settings for the steps (3.11), (3.12), (3.13) and (3.14).

**Step 1** We prove the upper bound of the sum process (Proposition 3.2). For this we construct a process, the difference process, which records the drift from the sum process away from the upper bound \( \tilde{n}_AK \).

**Proof of Proposition 3.2** The difference process \( X_{t}^{\mu,\Sigma} \) between \( \Sigma(t)K \) and \( \tilde{n}_AK \) is a branching process with the same rates as \( \Sigma(t)K \). Set

\[
\begin{align*}
    X_{t}^{\mu,\Sigma} & = \Sigma(t)K - \tilde{n}_AK, \\
    T_{0}^{X_{t}^{\mu,\Sigma}} & = \inf\{t \geq 0 : X_{t}^{\mu,\Sigma} = 0\}, \\
    T_{a,M}^{X_{t}^{\mu,\Sigma}} & = \inf\{t \geq 0 : X_{t}^{\mu,\Sigma} \geq 3M(\lambda^2\epsilon^2)^{1+\alpha}K\}. \tag{5.62}
\end{align*}
\]

This is a process in continuous time. For the following we need the discrete process \( Y_{n}^{\mu,\Sigma} \) associated to \( X_{t}^{\mu,\Sigma} \). To obtain this process we introduce a sequence of stopping times \( \vartheta_n \) which records the times, when \( X_{t}^{\mu,\Sigma} \) makes a jump. Formally, \( \vartheta_n \) is the smallest time \( t \) such that \( X_{t}^{\mu,\Sigma} \neq X_{s}^{\mu,\Sigma} \), for all \( \vartheta_{n-1} \leq s < t \). We set \( X_{\vartheta_n}^{\mu,\Sigma} = Y_{n}^{\mu,\Sigma} \).

This discretisation has probability less than \( \frac{1}{2} \) to make an upward jump:

**Lemma 5.4** For \( n \in \mathbb{N} \) such that \( \vartheta_n \in [\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK}a_k] \) and \( 1 \leq k \leq 3M(\lambda^2\epsilon^2)^{1+\alpha}K \) there exists a constant \( C_0 > 0 \) such that

\[
\mathbb{P}[Y_{n+1}^{\mu,\Sigma} = k+1|Y_{n}^{\mu,\Sigma} = k] \leq \frac{1}{2} - C_0kK^{-1} \equiv \rho_{\Sigma}(k). \tag{5.63}
\]
**Proof** With the rates \((5.58)\), we get by some straightforward computations

\[
\mathbb{P}[Y_{n+1}^\Sigma = k + 1 | Y_n^\Sigma = k] = \frac{b_k}{b_k + d_k} \\
= \frac{(\hat{n}_A K + k) f}{(\hat{n}_A K + k) [f + D + c(\hat{n}_A + k K^{-1})] + \Delta N_{aa}(t)} \\
= \frac{1}{2} + \frac{1}{2}D - \frac{1}{2}c\hat{n}_A - \frac{1}{2}ckK^{-1} - \frac{1}{2}\frac{\Delta N_{aa}(t)}{\hat{n}_A K + k}
\]

\[\leq \frac{1}{2} - C_0kK^{-1}.\]  \hfill (5.64)

To obtain a Markov chain we couple the process \(Y_n^\Sigma\) with a process \(Z_n^\Sigma\) via:

\begin{enumerate}
\item \(Z_0^\Sigma = Y_0^\Sigma \lor 0\)
\item \(\mathbb{P}[Z_{n+1}^\Sigma = k + 1 | Y_n^\Sigma < Z_n^\Sigma = k] = p_k(k),\) \hfill (5.65)
\item \(\mathbb{P}[Z_{n+1}^\Sigma = k - 1 | Y_n^\Sigma < Z_n^\Sigma = k] = 1 - p_k(k),\) \hfill (5.66)
\item \(\mathbb{P}[Z_{n+1}^\Sigma = k + 1 | Y_n^\Sigma = Z_n^\Sigma = k] = 1,\) \hfill (5.67)
\item \(\mathbb{P}[Z_{n+1}^\Sigma = k - 1 | Y_n^\Sigma = Z_n^\Sigma = k] = 1 - \frac{p_k(k) - \mathbb{P}[Y_{n+1}^\Sigma = k + 1 | Y_n^\Sigma = k]}{1 - \mathbb{P}[Y_{n+1}^\Sigma = k + 1 | Y_n^\Sigma = k]},\) \hfill (5.68)
\end{enumerate}

Observe that by construction \(Z_n^\Sigma \geq Y_n^\Sigma\), a.s. and that \(\mathbb{P}[Z_{n+1}^\Sigma = 1 | Z_n^\Sigma = 0] = 1\).

The marginal distribution of \(Z_n^\Sigma\) is the desired Markov chain with transition probabilities

\[
\mathbb{P}[Z_{n+1}^\Sigma = k + 1 | Z_n^\Sigma = k] = p_k(k)\mathbb{P}[Y_n^\Sigma < Z_n^\Sigma | Z_n^\Sigma = k] + \mathbb{P}[Y_n^\Sigma = k + 1 | Y_n^\Sigma = k]\mathbb{P}[Y_n^\Sigma = Z_n^\Sigma | Z_n^\Sigma = k]
\]

\[
+ \frac{p_k(k) - \mathbb{P}[Y_{n+1}^\Sigma = k + 1 | Y_{n+1}^\Sigma = k]}{1 - \mathbb{P}[Y_{n+1}^\Sigma = k + 1 | Y_{n+1}^\Sigma = k]}\mathbb{P}[Y_n^\Sigma = Z_n^\Sigma | Z_n^\Sigma = k](1 - \mathbb{P}[Y_n^\Sigma = k + 1 | Y_n^\Sigma = k])
\]

\[= p_k(k)\mathbb{P}[Y_n^\Sigma < Z_n^\Sigma | Z_n^\Sigma = k] + \mathbb{P}[Y_n^\Sigma = Z_n^\Sigma | Z_n^\Sigma = k] = p_k(k),\] \hfill (5.71)

and invariant measure

\[
\mu(n) = \frac{\prod_{k=1}^{n-1} \left( \frac{1}{2} - C_0kK^{-1} \right)}{\prod_{k=1}^{n} \left( \frac{1}{2} + C_0kK^{-1} \right)}, \text{ for } n \geq 2,
\]  \hfill (5.73)
with \( \mu(0) = 1 \) and \( \mu(1) = \frac{1}{2 + C_0 K^{-1}} \). We want to calculate the probability that the
Markov chain \( Z_n^{\mu, \Sigma} \), starting at a point \( zK \), reaches first \( 3M_{\Sigma}(x^{2i}e^2)^{1+\alpha}K \) before going to zero, which is the equilibrium potential of a one dimensional chain (see
Bovier 2006). Using Equation 7.1.57 in the book by Bovier and den Hollander (2015),
Chapter 7.1.4, we get (for \( K \) large enough)

\[
\mathbb{P}_{zK}[T_{\alpha, M_{\Sigma}}^{Z_n^{\mu, \Sigma}} < T_0^{Z_0^{\mu, \Sigma}}] = \frac{\sum_{n=1}^{zK} \frac{1}{1 - p_{\Sigma}(n)} \frac{1}{\mu(n)}}{\sum_{n=1}^{zK} \frac{1}{1 - p_{\Sigma}(n)} \frac{1}{\mu(n)}}
\]

\[
= \frac{\sum_{n=1}^{zK} \prod_{k=1}^{n-1} \frac{1}{1 + 2C_0kK^{-1}}}{\sum_{n=1}^{zK} \prod_{k=1}^{n-1} \frac{1}{1 - 2C_0kK^{-1}}}
\]

\[
= \frac{\sum_{n=1}^{zK} \exp \left( \sum_{k=1}^{n-1} \ln \left( \frac{1 + 2C_0kK^{-1}}{1 - 2C_0kK^{-1}} \right) \right)}{\sum_{n=1}^{zK} \exp \left( \sum_{k=1}^{n-1} 4C_0kK^{-1} \right)}
\]

\[
\leq \frac{\sum_{n=1}^{zK} \exp \left( \sum_{k=1}^{n-1} 4C_0kK^{-1} - \mathcal{O}(kK^{-2}) \right)}{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} \right)}
\]

\[
\leq \frac{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} - \mathcal{O}(nK^{-2}) \right)}{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} \right)}
\]

\[
\leq \frac{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} - \mathcal{O}(nK^{-2}) \right)}{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} \right)}
\]

\[
\leq \frac{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} - \mathcal{O}(nK^{-2}) \right)}{\sum_{n=1}^{zK} \exp \left( 2C_0n^2K^{-1} - 2C_0nK^{-1} \right)}
\]

\[
\leq \exp \left( 2C_0z^2K \right)
\]

\[
\leq \exp \left( 8C_0M_{\Sigma}(x^{2i}e^2)^{1+\alpha}K - 4C_0M_{\Sigma}(x^{2i}e^2)^{1+\alpha} - \mathcal{O}((x^{6i}e^6)^{1+\alpha}K) \right)
\]

\[
\leq \exp(-2C_0K(2M_{\Sigma}(x^{2i}e^2)^{1+\alpha} - z^2)).
\]

We denote by \( R \) the number of times that the process \( Z_n^{\mu, \Sigma} \) returns to zero before it
reaches \( 3M_{\Sigma}(x^{2i}e^2)^{1+\alpha}K \). Let \( R_z^k = \mathbb{P}_{zK}[R = k] \) be the probability that this number is \( k \) when starting in \( zK \). We define the times of the \( n \)-th returns to zero:

\[
T_0^1 = \inf \{ t > 0 : Z_t^{\mu, \Sigma} = 0 \}, \quad T_0^n = \inf \{ t > T_0^{n-1} : Z_t^{\mu, \Sigma} = 0 \}.
\]

We then have

\[
R_z^k = \mathbb{P}_{zK}[T_0 < T_{\alpha, M_{\Sigma}}^{Z_n^{\mu, \Sigma}}](1 - \mathbb{P}_0[T_{\alpha, M_{\Sigma}}^{Z_n^{\mu, \Sigma}} < T_0])^{k-1}\mathbb{P}_0[T_{\alpha, M_{\Sigma}}^{Z_n^{\mu, \Sigma}} < T_0] \leq (1 - A)^{k-1}A,
\]
where

\[ A \equiv \mathbb{P}_0[T_{\alpha, M}^\Sigma < T_0] = \sum_{i \geq 1} p(0, i) \mathbb{P}_i[T_{\alpha, M}^\Sigma < T_0] = p(0, 1) \mathbb{P}_1[T_{\alpha, M}^\Sigma < T_0] \leq \exp(-\tilde{C}_0 K M^2 \Sigma(x^4 i^4)^{1+\alpha}) \leq \exp(-\tilde{C}_0 M^2 \Sigma K^{3 \alpha}), \] (5.77)

for some finite positive constant \( \tilde{C}_0 \) and \( p(0, i) = P[Z_n^\Sigma = 0 | Z_{n+1}^\Sigma = i] \). Then

\[ \mathbb{P}[R \leq N] \leq \sum_{i=1}^{N} R_i^i \leq \sum_{i=1}^{N} (1 - A)^{i-1} A = 1 - (1 - A)^N. \] (5.78)

We choose, e.g., \( N \sim \frac{1}{K^2 A} \), so that \( \mathbb{P}[R \leq N] = o(K^{-1}) \). Let \( I_0 \equiv T_1^0 \) and \( I_n \equiv T_n^0 - T_{n-1}^0 \) the time the process needs for return to zero. The \((I_j)_{j \in \mathbb{N}}\) are i.i.d. random variables and it holds:

\[ \sum_{n=1}^{R} I_n \leq T_{\alpha, M}^\Sigma \leq \sum_{n=1}^{R+1} I_n. \] (5.79)

The underlying process, the sum process \( \Sigma(t) \) (2.12), of \( Z_n^\Sigma \) jumps with a rate

\[ \lambda_\Sigma = f \Sigma(t) K + D \Sigma(t) K + c \Sigma(t)^2 K + \Delta N_{aa}(t) \leq C_\lambda K. \] (5.80)

Since the Markov chain \( Z_n^\Sigma \) has to jump at least one time, it holds that, for all \( j \in \mathbb{N} \), \( I_j > W \), a.s., where \( W \sim \exp(C_\lambda K) \). Thus

\[ \mathbb{P}[I_j < y] \leq \mathbb{P}[W < y] = 1 - \exp(-C_\lambda K y). \] (5.81)

We have

\[ \mathbb{P}[T_{\alpha, M}^\Sigma < \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge e^{VK_a}] \leq \mathbb{P}[T_{\alpha, M}^\Sigma < e^{VK_a}] = \mathbb{P}[T_{\alpha, M}^\Sigma < e^{VK_a} \cap [R > N]] + \mathbb{P}[T_{\alpha, M}^\Sigma < e^{VK_a} \cap [R \leq N]] \leq \mathbb{P}[T_{\alpha, M}^\Sigma < e^{VK_a} \cap [R > N]] + \mathbb{P}[R \leq N]. \] (5.82)

First we estimate \( \mathbb{P}[T_{\alpha, M}^\Sigma < e^{VK_a} \cap [R > N]] \). Since \( T_{\alpha, M}^\Sigma \geq \sum_{n=1}^{R} I_n \), it holds that if \( \frac{n}{2} \) of the \( I_j \) are greater than \( \frac{2}{n} e^{VK_a} \), then \( T_{\alpha, M}^\Sigma \geq e^{VK_a} \).
\[ \mathbb{P} \left[ T^{X, u \Sigma}_{\alpha, M \Sigma} < e^{VK^\alpha} \cap \{ R > N \} \right] \]
\[ \leq \sum_{n=N}^{\infty} \mathbb{P} \left[ T^{X, u \Sigma}_{\alpha, M \Sigma} < e^{VK^\alpha} \cap \{ R = n \} \right] \]
\[ \leq \sum_{n=N}^{\infty} \mathbb{P} \left[ \sum_{j=1}^{n} I_j < e^{VK^\alpha} \right] \leq \sum_{n=N}^{\infty} \mathbb{P} \left[ \sum_{j=1}^{n} \mathbb{1}_{\left\{ I_j < \frac{e^{VK^\alpha}}{n} \right\}} > \frac{n}{2} \right]. \tag{5.83} \]

We have \( p_n \equiv \mathbb{P} \left[ I_j < \frac{2}{n} e^{VK^\alpha} \right] \leq \mathbb{P} \left[ W < \frac{2}{n} e^{VK^\alpha} \right] = 1 - \exp \left( -\frac{2C_{\lambda} K e^{VK^\alpha}}{n} \right). \) The number of random variables \( I_j \) that are greater than \( \frac{2}{n} e^{VK^\alpha} \) is binomial distributed with parameters \( p_n, n. \)

\[ \sum_{n=N}^{\infty} \mathbb{P} \left[ \sum_{j=1}^{n} \mathbb{1}_{\left\{ I_j < \frac{2}{n} e^{VK^\alpha} \right\}} > \frac{n}{2} \right] \]
\[ \leq \sum_{n=N}^{\infty} 4^n p_n^2 \leq \frac{(16 p_N)^{N/2}}{1 - 4 p_N^{1/2}}, \tag{5.84} \]

where we used that, in the range of summation, \( p_n \leq p_N. \) Then, for \( K \) large enough, \( 4 p_N^{1/2} \leq 1/2, \) and

\[ (16 p_N)^{N/2} \leq \left( 16 \left( 1 - \exp \left( -\frac{2C_{\lambda} K e^{VK^\alpha}}{N} \right) \right) \right)^{N/2} \leq \left( 16 \left( \frac{2C_{\lambda} K e^{VK^\alpha}}{N} \right) \right)^{N/2} = (16(2C_{\lambda} A K^3 e^{VK^\alpha}))^{N/2}. \tag{5.85} \]

Recalling that \( A = e^{-O(K^{3\alpha})}, \) one sees that (5.84) is bounded by \( o(e^{-K^{2\alpha}}). \) Hence we get

\[ \mathbb{P} \left[ T^{X, u \Sigma}_{\alpha, M \Sigma} < \tau_{iA}^+ \wedge \tau_{(i+1)-A}^+ \wedge e^{VK^\alpha} \right] \leq \mathbb{P} \left[ T^{X, u \Sigma}_{\alpha, M \Sigma} < e^{VK^\alpha} \cap \{ R > N \} \right] + \mathbb{P} \left[ R \leq N \right] \]
\[ = o \left( \frac{1}{K} \right). \tag{5.86} \]

This concludes the proof of Proposition (3.2).

\[ \square \]

**Step 2** We derive a rough upper bound on the process \( n_{aa} \) (Proposition 3.3). Recall that

\[ \gamma_{\Delta} \equiv \frac{f + \Delta}{4 \bar{n}_A (f + \Delta)}. \tag{5.87} \]

**Proof of Proposition 3.3** The proof is similar to the one of Proposition 3.2. Again, we define the difference process \( X_i^{\Delta i} \) between \( N_{aa} \) and \( \gamma_{\Delta x^{2i}} e^{2K}. \) This is a branching process with the same rates as \( n_{aa}. \) Set
\[ X_t^{aa} = N_{aa}(t) - \gamma_{\Delta} x^{2i} \epsilon^2 K \]  
(5.88)

\[ T_{0}^{X,aa} = \inf\{t \geq 0 : X_t^{aa} = 0\} \]  
(5.89)

\[ T_{\alpha, M_{aa}}^{X, aa} = \inf\{t \geq 0 : X_t^{aa} \geq 3M_{aa}(x^{2i} \epsilon^2)^{1+\alpha} K\}. \]  
(5.90)

Let \( Y_n^{aa} \) be the discretisation of \( X_t^{aa} \), obtained as described in Step 1.

**Lemma 5.5** For \( n \in \mathbb{N} \) such that \( \vartheta_n \in [\tau_{a_A}^{i-}, \tau_{a_A}^{i+} \wedge \tau_{a_A}^{(i+1)-} \wedge e^v K_{a}^i] \) and \( 1 \leq k \leq 3M_{aa}(x^{2i} \epsilon)_{1+\alpha} K \), there exists a constant \( C_0 > 0 \) such that

\[ \mathbb{P}[Y_{n+1}^{aa} = k + 1 | Y_n^{aa} = k] \leq \frac{1}{2} - C_0 \equiv p_{aa}. \]  
(5.91)

**Proof** In the following we use Proposition 3.1 for the first iteration step, since the mutant population \( n_{AA} \) increased to an \( \epsilon \)-neighbourhood of its equilibrium \( \bar{n}_A \) and the other two populations decreased to an \( \epsilon \) order. Thus the influence of the small \( a_A \)- and \( aa \)-populations is negligible for the dynamics of \( n_{AA} \) and the \( AA \)-population will stay in the \( \bar{n}_A \)-neighbourhood an exponentially long time. Now for \( i^{th} \) iteration-step we use the finer bounds of \( n_{AA}(t) \) (Propositions 3.5, 3.6) and \( \Sigma(t) \) (Proposition 3.4) estimated in the \( (i - 1)^{th} \) iteration-step before. By (5.59), we have

\[ \mathbb{P}[Y_{n+1}^{aa} = k + 1 | Y_n^{aa} = k] \]

\[ = \frac{(\gamma_{\Delta} x^{2i} \epsilon^2 K + k) f \left( 1 - \frac{n_{AA}(t)}{\Sigma(t)} \right) + \frac{f K}{4 \Sigma(t)} n_{aa}^2(t)}{(\gamma_{\Delta} x^{2i} \epsilon^2 K + k) \left( f \left( 1 - \frac{n_{AA}(t)}{\Sigma(t)} \right) + D + \Delta + c \Sigma(t) \right) + \frac{f K}{4 \Sigma(t)} n_{aa}^2(t)} \]

\[ = \frac{1}{2} + \frac{1}{2} \frac{f}{\Sigma(t)} n_{AA}(t) + \frac{f K}{8 \Sigma(t) (\gamma_{\Delta} x^{2i} \epsilon^2 K + k)} - \frac{1}{2} D - \frac{1}{2} \Delta - \frac{1}{2} c \Sigma(t) + \frac{f K}{4 \Sigma(t)} n_{aa}^2(t) \]  
(5.92)

Using Propositions 3.6, 3.4, and 3.2 and (3.17), one sees that the numerator in the second summand of (5.92) is bounded from above by

\[
\begin{align*}
\frac{f}{2} & - \frac{f(n_{A} - x^i \epsilon)}{2(n_{A} + 3M_{\Sigma}(x^{2i} \epsilon)^{1+\alpha})} + \frac{f(x^i \epsilon + x^{2i} \epsilon^2)^2}{8\gamma_{\Delta} x^{2i} \epsilon^2(n_{A} - \frac{\Delta + c \gamma_{\Delta}}{\alpha_{\gamma_{\Delta}}}) \gamma_{\Delta} x^{2i} \epsilon^2 + 3M_{\Sigma}(x^{2i} \epsilon)^{1+\alpha})} \\
& - \frac{D + \Delta + c \bar{n}_A}{2} + O(x^{2i} \epsilon^2) \\
& \leq - \frac{f}{2} + \frac{f}{8n_{A} \gamma_{\Delta}} + O(x^i \epsilon) = - \frac{f + \Delta}{2f + \Delta} + O(x^i \epsilon).
\end{align*}
\]  
(5.93)

Since \( \epsilon < \frac{\Delta}{2} \), there exists a constant \( C_0 > 0 \) such that

\[ \mathbb{P}[Y_{n+1}^{aa} = k + 1 | Y_n^{aa} = k] \leq \frac{1}{2} - C_0 \equiv p_{aa}. \]  
(5.94)
As in Step 1 we couple \( Y_n^{aa} \) with a process \( Z_n^{aa} \) via:

1. \( Z_{0}^{aa} = Y_{0}^{aa} \vee 0 \) \hspace{1cm} (5.95)
2. \( \mathbb{P}[Z_{n+1}^{aa} = k+1|Y_{n}^{aa} < Z_{n}^{aa} = k] = p_{aa}, \) \hspace{1cm} (5.96)
3. \( \mathbb{P}[Z_{n+1}^{aa} = k-1|Y_{n}^{aa} < Z_{n}^{aa} = k] = 1 - p_{aa}, \) \hspace{1cm} (5.97)
4. \( \mathbb{P}[Z_{n+1}^{aa} = k+1|Y_{n}^{aa} = k] = 1, \) \hspace{1cm} (5.98)
5. \( \mathbb{P}[Z_{n+1}^{aa} = k-1|Y_{n}^{aa} = k] = 1 - \frac{p_{0}}{\mathbb{P}[Y_{n+1}^{aa} = k+1|Y_{n}^{aa} = k]} \) \hspace{1cm} (5.99)
6. \( \mathbb{P}[Z_{n+1}^{aa} = k-1|Y_{n}^{aa} = k] = 1 - \frac{p_{0}}{\mathbb{P}[Y_{n+1}^{aa} = k+1|Y_{n}^{aa} = k]} \) \hspace{1cm} (5.100)

Observe that by construction \( Z_{n}^{aa} \geq Y_{n}^{aa} \), a.s. The marginal distribution of \( Z_{n}^{aa} \) is the desired Markov chain with transition probabilities

\[
\mathbb{P}[Z_{n+1}^{aa} = k+1|Z_{n}^{aa} = k] = p_{aa} = 1 - \mathbb{P}[Z_{n+1}^{aa} = k-1|Z_{n}^{aa} = k] 
\] \hspace{1cm} (5.101)

and invariant measure

\[
\mu(n) = \frac{\prod_{k=1}^{n-1} \left( \frac{1}{2} - C_{0} \right)}{\prod_{k=1}^{n} \left( \frac{1}{2} + C_{0} \right)} = \left( \frac{1}{2} - C_{0} \right)^{n-1} \cdot \left( \frac{1}{2} + C_{0} \right)^{n}. \] \hspace{1cm} (5.102)

The remainder of the proof is a complete re-run of the proof of Proposition 3.2 and we skip the details.

**Step 3** We estimate the lower bound on \( \Sigma(t) \), for \( t \in \left[ \tau_{i-}^{\alpha}, \tau_{i+}^{\alpha} \land \tau_{i+}^{\alpha-} \land e^{VK_{\alpha}} \right] \).

**Proof of Proposition 3.4** The proof is similar to those of Propositions 3.2 and 3.3. We only perform the crucial steps. This time the difference process is given by the difference of \( \Sigma(t) \) and \( \bar{\nu}_{A} - \frac{\Delta+\phi}{c_{1}} \gamma_{\Delta} x^{2i} e^{i} \). Let

\[
X_{i}^{\Sigma} = \Sigma(t) K - \left( \bar{\nu}_{A} - \frac{\Delta+\phi}{c_{1}} \gamma_{\Delta} x^{2i} e^{i} \right) K 
\] \hspace{1cm} (5.103)

\[
T_{o}^{X,i} \equiv \inf \{ t \geq 0 : X_{i}^{\Sigma} = 0 \} \] \hspace{1cm} (5.104)

\[
T_{a,M_{\Sigma}}^{X,i} \equiv \inf \{ t \geq 0 : X_{i}^{\Sigma} \leq -3 M_{\Sigma} (x^{2i} e^{i})^{1+\alpha} K \} \] \hspace{1cm} (5.105)

As described in Step 1 we construct the discrete process \( Y_{n}^{\Sigma} \) associated to \( X_{i}^{\Sigma} \). We show that \( Y_{n}^{\Sigma} \) jumps down with a probability less than \( \frac{1}{2} \).

**Lemma 5.6** For \( n \in \mathbb{N} \) such that \( \theta_{n} \in [\tau_{i-}^{\alpha}, \tau_{i+}^{\alpha} \land \tau_{i+}^{\alpha-} \land e^{VK_{\alpha}} \) and \( 1 \leq k \leq 3 M_{\Sigma} (x^{2i} e^{i})^{1+\alpha} K \) there exists a constant \( C_{0} > 0 \) such that

\[
\mathbb{P}[Y_{n+1} = -k-1|Y_{n} = -k] \leq \frac{1}{2} - C_{0} x^{2i} e^{i} \equiv p_{\Sigma} \] \hspace{1cm} (5.106)
Proof Using the rates of the sum process (5.58) and the upper bound on $n_{aa}$ (Proposition 3.3), this is a simple computation and we skip the details. \hfill \Box

As in Step 1, to obtain a Markov chain we couple the process $Y_{n}^{l\Sigma}$ with a process $Z_{n}^{l\Sigma}$ via:

\begin{enumerate}
\item $Z_{0} = Y_{0} \lor 0,$ \hspace{1cm} (5.107)
\item $\mathbb{P}[Z_{n+1}^{l\Sigma} = -k + 1|Y_{n}^{l\Sigma} > Z_{n}^{l\Sigma} = -k] = 1 - p_{\Sigma},$ \hspace{1cm} (5.108)
\item $\mathbb{P}[Z_{n+1}^{l\Sigma} = -k - 1|Y_{n}^{l\Sigma} > Z_{n}^{l\Sigma} = -k] = p_{\Sigma},$ \hspace{1cm} (5.109)
\item $\mathbb{P}[Z_{n+1}^{l\Sigma} = -k - 1|Y_{n}^{l\Sigma} = -k - 1, Y_{n}^{l\Sigma} = Z_{n}^{l\Sigma} = -k] = 1,$ \hspace{1cm} (5.110)
\item $\mathbb{P}[Z_{n+1}^{l\Sigma} = -k - 1|Y_{n}^{l\Sigma} = -k + 1, Y_{n}^{l\Sigma} = Z_{n}^{l\Sigma} = -k] = 1 - \frac{p_{\Sigma} - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k|Y_{n}^{l\Sigma} = -k]}{1 - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k|Y_{n}^{l\Sigma} = -k]},$ \hspace{1cm} (5.111)
\item $\mathbb{P}[Z_{n+1}^{l\Sigma} = -k + 1|Y_{n}^{l\Sigma} = -k + 1, Y_{n}^{l\Sigma} = Z_{n}^{l\Sigma} = -k] = 1 - \frac{p_{\Sigma} - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k + 1|Y_{n}^{l\Sigma} = -k]}{1 - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k + 1|Y_{n}^{l\Sigma} = -k]}.$ \hspace{1cm} (5.112)
\end{enumerate}

Observe that by construction $Z_{n}^{l\Sigma} \preceq Y_{n}^{l\Sigma}$, a.s.. The marginal distribution of $Z_{n}^{l\Sigma}$ is the desired Markov chain with transition probabilities

$$\mathbb{P}[Z_{n+1}^{l\Sigma} = k + 1|Z_{n}^{l\Sigma} = k] = p_{\Sigma},$$

$$\mathbb{P}[Z_{n+1}^{l\Sigma} = k - 1|Z_{n}^{l\Sigma} = k] = 1 - p_{\Sigma},$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1} (\frac{1}{2} - C_{0}x^{2i} \varepsilon^{2})}{\prod_{k=1}^{n} (\frac{1}{2} + C_{0}x^{2i} \varepsilon^{2})} = \left(\frac{1}{2} - C_{0}x^{2i} \varepsilon^{2}\right)^{n-1} \left(\frac{1}{2} + C_{0}x^{2i} \varepsilon^{2}\right)^{n}.$$ \hspace{1cm} (5.115)

The remainder of the proof follows along the lines of the proof given in Step 1. We prove that the process returns to zero many times before it hits $3M_{\Sigma}(x^{2i} \varepsilon^{2})^{1+\alpha}$ and calculate the duration of one zero-return to get the desired result. \hfill \Box

Step 4 With Propositions 3.2–3.4 and the settings we are able to calculate a lower (Step 4.1) and an upper bound (Step 4.2) for $n_{AA}(t)$, for $t \in \left[\tau_{aA}^{-i}, \tau_{aA}^{i+} \land \tau_{aA}^{(i+1)-} \land e^{V^{kA}}\right]$.

Step 4.1 We now prove Proposition 3.5, the lower bound on $n_{AA}$.

Proof of Proposition 3.5 From Proposition 3.4 we know that $\Sigma(t) \geq \tilde{n}_{A} - \frac{\Delta + \vartheta}{cn_{A}} \nu_{\Delta} x^{2i} \varepsilon^{2} - 3M_{\Sigma}(x^{2i} \varepsilon^{2})^{1+\alpha}$. With the upper bound in Proposition 3.3 for $n_{aa}$ and the settings for the steps used for the $aA$-population, we get

$$n_{AA}(t) = \Sigma(t) - n_{aA}(t) - n_{aa}(t) \geq \tilde{n}_{A} - \frac{\Delta + \vartheta}{cn_{A}} \nu_{\Delta} x^{2i} \varepsilon^{2} - x^{i} \varepsilon - (1 + \gamma_{\Delta})x^{2i} \varepsilon^{2} - 3(M_{\Sigma} + M_{aa})(x^{2i} \varepsilon^{2})^{1+\alpha}$$

$$\geq \tilde{n}_{A} - x^{i} \varepsilon + O(x^{2i} \varepsilon^{2}).$$ \hspace{1cm} (5.116)
Step 4.2 We prove Proposition 3.6, the upper bound on $n_{AA}$.

Proof of Proposition 3.6 From Proposition 3.2 we know that $\Sigma(t) \leq \tilde{n}_A + 3M \Sigma(x^2i \epsilon^2)^{1+\alpha}$. With the lower bound on $n_{AA}(t)$ defined in the settings we get

$$n_{AA}(t) = \Sigma(t) - n_{aa}(t) - n_{aa}(t) \leq \tilde{n}_A + 3M \Sigma(x^2i \epsilon^2)^{1+\alpha} - x^{i+1} \epsilon$$

$$\leq \tilde{n}_A - x^{i+1} \epsilon + O((x^2i \epsilon^2)^{1+\alpha}). \tag{5.117}$$

\[ \Box \]

Step 5 Up to now we have estimated upper and lower bounds for all single processes: $\Sigma(t), n_{aa}(t), n_{AA}(t)$ and $n_{AA}(t)$, for $t \in [\tau_{aA}^i, \tau_{aA}^i \wedge \tau_{aA}^i \wedge \epsilon V K^a]$. Now, we prove that $n_{AA}(t)$ has the tendency to decrease on the time intervals defined in the settings. For this we restart $n_{AA}$ when the process hits $x^i \epsilon$ and show that with high probability the $aA$-population decreases to $x^{i+1} \epsilon$ before it exceeds again the $x^i \epsilon + x^2i \epsilon^2$-value (Proposition 3.7). For this we couple $n_{AA}(t)$ with a process which minorises it and on one which majorises it and show that these processes decrease to $x^{i+1} \epsilon$ before going back to $x^i \epsilon + x^2i \epsilon^2$.

Proof of Proposition 3.7 As before let $Y_n^{aA}$ (cf. Step 1) be the associated discrete process to $N_n^{aA}$. We start by coupling $Y_n^{aA}$ with a Markov chain $Z_n$ such that $Z_n \nRightarrow Y_n^{aA}$, a.s.

Lemma 5.7 For $n \in \mathbb{N}$ such that $\vartheta_n \in [\tau_{aA}^i, \tau_{aA}^i \wedge \tau_{aA}^i \wedge \epsilon V K^a]$ there exists a constant $C_0 > 0$ such that

$$\mathbb{P}[Y_{n+1}^{aA} = k + 1 | Y_n^{aA} = k] \leq \frac{1}{2} - C_0 x^{i+1} \epsilon \equiv p_{aA}^\epsilon. \tag{5.118}$$

Proof For $t < \tau_{aA}^i \wedge \epsilon V K^a$, we have

$$\mathbb{P}[Y_{n+1}^{aA} = k + 1 | Y_n^{aA} = k] = \frac{f k \left( 1 - \frac{kK^{-1}}{\Sigma(t)} \right) + 2 f N_{aa}(t) \frac{n_{AA}(t)}{\Sigma(t)} k [D + c \Sigma(t)]}{f k \left( 1 - \frac{kK^{-1}}{\Sigma(t)} \right) + 2 f N_{aa}(t) \frac{n_{AA}(t)}{\Sigma(t)} k} \leq \frac{1}{2} + \frac{1 f \left( 1 - \frac{kK^{-1}}{\Sigma(t)} \right) + 2 f N_{aa}(t) \frac{n_{AA}(t)}{\Sigma(t)} k [D + c \Sigma(t)]}{f \left( 1 - \frac{kK^{-1}}{\Sigma(t)} \right) + 2 f N_{aa}(t) \frac{n_{AA}(t)}{\Sigma(t)} k [D + c \Sigma(t)]}. \tag{5.119}$$

As in the previous steps, we bound the nominator of the second summand in (5.119) using Propositions 3.2, 3.4, and 3.6, from above by

$$\frac{1}{2} f \frac{1}{4n_A} k K^{-1} + \frac{f (n_A - x^{i+1} \epsilon)}{kK^{-1} (n_A - \frac{\Delta + \vartheta}{cA} \gamma_d x^2i \epsilon^2)} \gamma_d x^2i \epsilon^2 - D_2 = - \frac{\vartheta}{2} \left( \tilde{n}_A - \frac{\Delta + \vartheta}{cA} \gamma_d x^2i \epsilon^2 \right) + O((x^2i \epsilon^2)^{1+\alpha})$$

$$\leq - \frac{f \vartheta - \frac{\Delta}{2}}{4\tilde{n}_A f + \vartheta} x^{i+1} \epsilon + O(x^2i \epsilon^2). \tag{5.120}$$
This term is negative since $\frac{A}{2} < \vartheta$. Hence, we get

$$\mathbb{P}[Y_{n+1}^a = k + 1 | Y_n^a = k] \leq \frac{1}{2} - \frac{\vartheta - \frac{A}{2}}{8n_A(f + \vartheta)} x_{i+1}^i \epsilon + O(\epsilon^2). \quad (5.121)$$

To obtain a Markov chain we couple the process $Y_n^a$ with a process $Z_n^u$ via:

1. $Z_0^u = Y_0^a$
2. $\mathbb{P}[Z_{n+1}^u = k + 1 | Y_n^a < Z_n^u = k] = p_{uA}^u$.
3. $\mathbb{P}[Z_{n+1}^u = k - 1 | Y_n^a < Z_n^u = k] = 1 - p_{uA}^u$, (5.124)
4. $\mathbb{P}[Z_{n+1}^u = k + 1 | Y_n^a = Z_n^u = k] = 1$. (5.125)
5. $\mathbb{P}[Z_{n+1}^u = k + 1 | Y_n^a = Z_n^u = k] = \frac{P_{uA}^u - \mathbb{P}[Y_{n+1}^a = k + 1 | Y_n^a = k]}{1 - \mathbb{P}[Y_{n+1}^a = k + 1 | Y_n^a = k]}$. (5.126)
6. $\mathbb{P}[Z_{n+1}^u = k - 1 | Y_n^a = Z_n^u = k] = 1 - \frac{P_{uA}^u - \mathbb{P}[Y_{n+1}^a = k + 1 | Y_n^a = k]}{1 - \mathbb{P}[Y_{n+1}^a = k + 1 | Y_n^a = k]}$. (5.127)

Observe that by construction $Z_n^u \succeq Y_n^a$, a.s.. The marginal distribution of $Z_n^u$ is the desired Markov chain with transition probabilities

$$\mathbb{P}[Z_{n+1}^u = k + 1 | Z_n^u = k] = p_{aA}^u. \quad (5.128)$$
$$\mathbb{P}[Z_{n+1}^u = k - 1 | Z_n^u = k] = 1 - p_{aA}^u, \quad (5.129)$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1} \left( \frac{1}{2} - C_0 x_{i+1}^i \epsilon \right)}{\prod_{k=1}^{n} \left( \frac{1}{2} + C_0 x_{i+1}^i \epsilon \right)} = \frac{(\frac{1}{2} - C_0 x_{i+1}^i \epsilon)^{n-1}}{(\frac{1}{2} + C_0 x_{i+1}^i \epsilon)^n}. \quad (5.130)$$

We define the stopping times

$$T_{i+}^Z \equiv \inf \{ \vartheta_n \geq 0 : Z_n \geq x^i \epsilon K + x^{2i} \epsilon^2 K \}, \quad (5.131)$$
$$T_{(i+1)-}^Z \equiv \inf \{ \vartheta_n \geq 0 : Z_n \leq x_{i+1}^i \epsilon K \}. \quad (5.132)$$

For $x_{i+1}^i \epsilon \leq z < x^i \epsilon$, we get as before the following bound on the harmonic function

$$\mathbb{P}_{zK}[T_{i+}^Z < T_{(i+1)-}^Z] = \sum_{n=x_{i+1}^i + 1}^{zK} \frac{1 + 2C_0 x_{i+1}^i \epsilon}{1 - 2C_0 x_{i+1}^i \epsilon} \left( \frac{1 + 2C_0 x_{i+1}^i \epsilon}{1 - 2C_0 x_{i+1}^i \epsilon} \right)^{n-1} \leq K^{1/4 - \alpha} \exp \left( - \tilde{C} K^{1/4 + 3\alpha} \right).$$
Now we couple $Y_n^{aA}$ with a Markov chain $Z_n^l$ such that $Z_n^l \preceq Y_n^{aA}$, a.s.

**Lemma 5.8** For $n \in \mathbb{N}$ such that $\vartheta_n \in [\tau_{aA}^{i-}, \tau_{aA}^{i+} \land \tau_{aA}^{(i+1)-} \land e^V K^a]$ there exists a constant $C_1 > 0$ such that

$$\mathbb{P}[Y_{n+1} = k + 1 | Y_n^{aA} = k] \geq \frac{1}{2} - C_1 x^i \epsilon \equiv p_{aA}^l.$$  (5.133)

**Proof** The proof is completely analogous to the proof of Lemma 5.7 and we skip the details. \qed

To obtain a Markov chain we couple the process $Y_n^{aA}$ with a process $Z_n^l$ via:

1. $Z_0^l = Y_0$
2. $\mathbb{P}[Z_{n+1}^l = k + 1 | Y_n^{aA} > Z_n^l = k] = p_{aA}^l$.
3. $\mathbb{P}[Z_{n+1}^l = k - 1 | Y_n^{aA} > Z_n^l = k] = 1 - p_{aA}^l$.
4. $\mathbb{P}[Z_{n+1}^l = k - 1 | Y_{n+1} = k - 1, Y_n^{aA} = Z_n^l = k] = 1$.
5. $\mathbb{P}[Z_{n+1}^l = k | Y_{n+1} = k + 1, Y_n^{aA} = Z_n^l = k] = \frac{\mathbb{P}[Y_{n+1} = k + 1 | Y_n^{aA} = k] - p_{aA}^l}{\mathbb{P}[Y_{n+1} = k + 1 | Y_n^{aA} = k]}$.
6. $\mathbb{P}[Z_{n+1}^l = k + 1 | Y_{n+1} = k + 1, Y_n^{aA} = Z_n^l = k] = 1 - \frac{\mathbb{P}[Y_{n+1} = k + 1 | Y_n^{aA} = k] - p_{aA}^l}{\mathbb{P}[Y_{n+1} = k + 1 | Y_n^{aA} = k]}$.  (5.138)

Observe that by construction $Z_n^l \preceq Y_n^{aA}$, a.s. The marginal distribution of $Z_n^l$ is the desired Markov chain with transition probabilities

$$\mathbb{P}[Z_{n+1}^l = k + 1 | Z_n^l = k] = p_{aA}^l,$$  (5.140)
$$\mathbb{P}[Z_{n+1}^l = k - 1 | Z_n^l = k] = 1 - p_{aA}^l.$$  (5.141)

Similar to the upper process, we can show that the lower process reaches $x^{i+1} \epsilon K$ before returning to $x^i \epsilon K + x^{2i} \epsilon^2 K$, with high probability. This concludes the proof of the proposition. \qed

**Step 6** In this step we calculate the time which $n_{aA}(t)$ needs to decrease from $x^i \epsilon$ to $x^{i+1} \epsilon$ (Proposition 3.8).

**Proof of Proposition 3.8** Let $Z_n^l \preceq Y_n^{aA} \preceq Z_n^u$ be defined as in the step before and $Y_0 = Z_0^l = Z_0^u = x^i \epsilon K$.

Recalling (5.60), we get

$$\lambda_{aA}(t) = f N_{aA}(t) \left(1 - \frac{n_{aA}(t)}{2 \Sigma(t)}\right) + 2 f N_{aA}(t) \frac{n_{A}(t)}{\Sigma(t)} + N_{aA}(t)[D + c \Sigma(t)]$$
$$\geq 2 f x^{i+1} \epsilon K + O(x^{2i} \epsilon^2 K) \equiv C_\lambda x^{i+1} \epsilon K \equiv \lambda_{aA}^l.$$  (5.142)
Let $n_* := \inf \{n \geq 0 : Y_n^{aA} - Y_0^{aA} \leq -(1-x)x^i \epsilon K\}$ be the random variable which counts the number of jumps $Y_n^{aA} - Y_0^{aA}$ makes until it is smaller than $-(1-x)x^i \epsilon K$. The time between two jumps of $n_{aA}(t)$ is given by $\tau_{m-1} - \tau_m$. It holds that $J^u_m \geq J^l_m$, where $J^u_m$ (resp. $J^l_m$) are i.i.d. exponential distributed random variables with parameter $\lambda^u_{aA}$ (resp. $\lambda^l_{aA}$). We want to estimate bounds for the times that the processes $Z^u_n$, resp. $Z^l_n$, need to decrease from $x^i \epsilon K$ to $x^{i+1} \epsilon K$. Thus we show, for constants $C_u, C_l > 0$, that

\begin{align}
(i) \quad \Pr \left[ \sum_{m=1}^{n_*} J^u_m > \frac{2C_u}{C_{\lambda}x^{i+1} \epsilon} \right] & \leq \exp(-M K^{1/2+2\alpha}), \\
(ii) \quad \Pr \left[ \sum_{m=1}^{n_*} J^l_m < \frac{C_l}{2C_{\lambda}x^i \epsilon} \right] & \leq \exp(-M K^{1/2+2\alpha}).
\end{align}

We start by showing (5.144). We need to find $N$ such that, with high probability, $n_* \leq N$. To do this, we use the majorising process $Z^u_n$. Let $W_k^u$ be i.i.d. random variables with

\begin{align}
\Pr[W_k^u = 1] &= \frac{1}{2} - C_0 x^{i+1} \epsilon, \\
\Pr[W_k^u = -1] &= \frac{1}{2} + C_0 x^{i+1} \epsilon, \text{ and } \mathbb{E}[W_k^u] = -2C_0 x^{i+1} \epsilon.
\end{align}

$W_k^u$ records a birth or a death event of the process $Z^u_n$. From Lemma 5.7 we get

\begin{align}
\Pr[n_* \leq N] &\geq \Pr \left[ \exists n \leq N : \sum_{k=1}^{n} W_k \leq - \left( (1-x)x^i \epsilon K \right) \right] \\
&\geq \Pr \left[ \sum_{k=1}^{N} W_k \leq - \left( (1-x)x^i \epsilon K \right) \right] \\
&\geq 1 - \Pr \left[ \sum_{k=1}^{N} (W_k - \mathbb{E}W_k) \geq 2NC_0 x^{i+1} \epsilon - \left( (1-x)x^i \epsilon K \right) \right].
\end{align}

By Hoeffding’s inequality and choosing $N = \frac{1-x}{C_0x}K =: C_u K$, we get

\begin{align}
\Pr[n_* \leq C_u K] &\geq 1 - \exp \left( -\frac{C_0(x^i \epsilon)^2 K x(1-x)}{2} \right) \\
&\geq 1 - \exp(-K^{1/2+2\alpha} C_0 (1-x)x/2),
\end{align}
where we used that $\lambda^i \epsilon \geq K^{-1/4+\alpha}$. Thus
\[
\mathbb{P}\left[ \sum_{m=1}^{n_*} J^u_m > \frac{2C_u}{C_\lambda x^i + 1} \epsilon \right] \leq \mathbb{P}\left[ \sum_{m=1}^{C_u K} J^u_m > \frac{2C_u}{C_\lambda x^i + 1} \epsilon \right] + \exp\left(-K^{1/2+2\alpha} C_0 x (1-x)/2\right).
\] (5.149)

By applying the exponential Chebyshev inequality we get
\[
\mathbb{P}\left[ \sum_{m=1}^{C_u K} J^u_m > \frac{2C_u}{C_\lambda x^i + 1} \epsilon \right] \leq \mathbb{P}\left[ \sum_{m=1}^{C_u K} J^u_m > \frac{2C_u}{C_\lambda x^i + 1} \epsilon \right] \leq \exp\left(-C_u K/2\right). \] (5.150)

Next we show (5.145). For this we need to find $N$ such that $\mathbb{P}[n_* \leq N]$ is very small. For this we use the process $Z^l$. Let $W^l_k$ be i.i.d. random variables which record a birth or a death event of the process $Z^l_n$. They satisfy
\[
\mathbb{P}[W^l_k = 1] = \frac{1}{2} - C_1 x^i \epsilon, \quad \mathbb{P}[W^l_k = -1] = \frac{1}{2} + C_1 x^i \epsilon \quad \text{and} \quad \mathbb{E}[W^l_k] = -2C_1 x^i \epsilon. \] (5.151)

Note that from Lemma 5.8
\[
\mathbb{P}[n_* \leq N] \leq \mathbb{P}\left[ \inf\left\{ n \geq 0 : Z^l_n - Z^l_0 \leq -\left[ (1-x)x^i \epsilon K \right] \right\} \leq N \right]
= \mathbb{P}\left[ \exists n \leq N : \sum_{k=1}^{n} (W^l_k - \mathbb{E}W^l_k) \leq 2n C_1 x^i \epsilon - \left[ (1-x)x^i \epsilon K \right] \right]
\leq \sum_{n=0}^{N} \mathbb{P}\left[ \sum_{k=1}^{n} (W^l_k - \mathbb{E}W^l_k) \leq 2n C_1 x^i \epsilon - \left[ (1-x)x^i \epsilon K \right] \right].
\] (5.152)

If we choose $N = \frac{1-x}{4C_1} K =: C_1 K$, using Hoeffding’s inequality, we get, for all $n \leq N$,
\[
\mathbb{P}\left[ \sum_{k=1}^{n} (W^l_k - \mathbb{E}W^l_k) \geq \left[ (1-x)x^i \epsilon K \right] - \frac{1}{2}(1-x)x^i \epsilon K \right]
\leq \exp(-K^{1/2+2\alpha} C_1 (1-x)/2).
\] (5.154)

Thus
\[
\mathbb{P}\left[ \sum_{m=1}^{n_*} J^l_m < \frac{C_l}{2C_\lambda x^i \epsilon} \right]
\leq \mathbb{P}\left[ \sum_{m=1}^{C_l K} J^l_m < \frac{C_l}{2C_\lambda x^i \epsilon} \right] + C_l K \exp(-K^{1/2+2\alpha} C_1 (1-x)/2).
\] (5.155)
It holds that
\[
\mathbb{P} \left[ \sum_{m=1}^{C_l K} J_m^l > \frac{C_l}{2C_x x^i \epsilon} \right] = 1 - \mathbb{P} \left[ \sum_{m=1}^{C_l K} J_m^l < \frac{C_l}{2C_x x^i \epsilon} \right] = 1 - \mathbb{P} \left[ -\sum_{m=1}^{C_l K} J_m^l > -\frac{C_l}{2C_x x^i \epsilon} \right].
\] (5.156)

A simple use of the exponential Chebyshev inequality shows that
\[
\mathbb{P} \left[ \sum_{m=1}^{C_l K} J_m^l < \frac{C_l}{2C_x x^i \epsilon} \right] \leq \exp(-C_l K / 2).
\] (5.157)

Thus we have that \(\mathbb{P} \left[ \sum_{m=1}^{n_a} J_m^l < \frac{C_l}{2C_x x^i \epsilon} \right] \leq \exp(-M K^{1/2+2\alpha})\), for some constant \(M > 0\).

**Step 7** In this step it is shown that \(n_{aa}(t)\) decreases under the upper bound \(\gamma \Delta x^{2i+2} \epsilon^2 + M_{aa}(x^{2i+2} \epsilon^2)^{1+\alpha}\), which we need to proceed the next iteration step, in at least the time \(n_{aA}(t)\) needs to decrease from \(x^i \epsilon\) to \(x^{i+1} \epsilon\) (Proposition 3.9). We set the time to zero when \(n_{aA}(t)\) hits \(x^i \epsilon\). Hence \(n_{aa}(0) \leq \gamma \Delta x^{2i+2} \epsilon^2 + M_{aa}(x^{2i+2} \epsilon^2)^{1+\alpha}\). Let
\[
\theta_i^+(aa) \equiv \inf \{ t \geq 0 : n_{aa}(t) \geq \gamma \Delta x^{2i+2} \epsilon^2 + 3M_{aa}(x^{2i+2} \epsilon^2)^{1+\alpha} \},
\] (5.158)
\[
\theta_i^-(aa) \equiv \inf \{ t \geq 0 : n_{aa}(t) \leq \gamma \Delta x^{2i+2} \epsilon^2 \}.
\] (5.159)

For the proof we proceed in three parts. First we show that \(n_{aa}(t)\) has the tendency to decrease. For this we construct a majorising process for \(n_{aa}(t)\) and show that this process decreases on the given time interval. This process is used in the second part to estimate an upper bound on the time which the \(aa\)-population needs for the decay from \(\gamma \Delta x^{2i+2} \epsilon^2 + M_{aa}(x^{2i+2} \epsilon^2)^{1+\alpha}\) to \(\gamma \Delta x^{2i+2} \epsilon^2\). As result we will get that \(n_{aa}(t)\) reaches the next upper bound before the \(aA\)-population decreases to \(x^{i+1} \epsilon\). Thus in the third part we ensure that \(n_{aa}(t)\) stays below the bound until \(n_{aA}(t)\) reaches \(x^{i+1} \epsilon\).

**Part 1** We show

**Proposition 5.2** For \(t \in [\tau_{iA}^-, \tau_{iA}^+ \wedge \tau_{iA}^{(i+1)-} \wedge \epsilon V K^\alpha] \) there are constants \(\tilde{C}, \tilde{\tilde{C}} > 0\) such that
\[
\mathbb{P}[\theta_i^+(aa) < \theta_i^-(aa) | n_{aa}(0) \leq \gamma \Delta x^{2i} \epsilon^2 + M_{aa}(x^{2i} \epsilon^2)^{1+\alpha}] \leq \tilde{C} K^{\alpha/2} \exp(-\tilde{\tilde{C}} K^{1/2+\alpha}).
\] (5.160)

In this part the same strategy as in Step 5 is used. We couple \(n_{aa}(t)\) with a process which majorises it and show that this process decreases from \(\gamma \Delta x^{2i+2} \epsilon^2 + M_{aa}(x^{2i+2} \epsilon^2)^{1+\alpha}\) to \(\gamma \Delta x^{2i+2} \epsilon^2\) with high probability.
As before, let \( Y_{n}^{aa} \) be the discretisation of \( N_{aa}(t) \). We start with the construction of the upper process.

**Lemma 5.9** For \( n \in \mathbb{N} \) such that \( \theta_{n} \in \left[ \tau_{aA}^{i-}, \tau_{aA}^{i+} \land \tau_{aA}^{(i+1)-} \land e^{VK}\right] \) there exists a constant \( C_{u} > 0 \) such that

\[
\mathbb{P}[Y_{n+1}^{aa} = k + 1|Y_{n}^{aa} = k] \leq \frac{1}{2} - C_{u} = p_{u}. \tag{5.161}
\]

**Proof** We know that, for \( t \in \left[ \tau_{aA}^{i-}, \tau_{aA}^{i+} \land \tau_{aA}^{(i+1)-} \land e^{VK}\right], n_{aA}(t) \in [x^{i+1}\epsilon, x^{i}\epsilon + x^{2i}\epsilon^{2}] \). Again with (5.59) we have

\[
\mathbb{P}[Y_{n+1}^{aa} = k + 1|Y_{n}^{aa} = k] = \frac{f_{k} \left( 1 - \frac{n_{aA}(t)}{\Sigma(t)} \right) + \frac{f_{K}}{4\Sigma(t)} n_{aA}^{2}(t)}{f_{k} \left( 1 - \frac{n_{aA}(t)}{\Sigma(t)} \right) + \frac{f_{K}}{4\Sigma(t)} n_{aA}^{2}(t) + k[D + \Delta + c \Sigma(t)]} \\
\leq \frac{1}{2} + \frac{\frac{1}{2} f - \frac{f_{K}}{8\Sigma(t)} n_{aA}(t) + \frac{f_{K}}{8\Sigma(t)} n_{aA}(t) - \frac{1}{2} D - \frac{1}{2} c \Sigma(t) - \frac{1}{2} \Delta}{f \left( 1 - \frac{n_{aA}(t)}{\Sigma(t)} \right) + \frac{f_{K}}{4\Sigma(t)} n_{aA}^{2}(t) + [D + \Delta + c \Sigma(t)]}.
\]

Using Propositions 3.2, 3.4 and 3.6, we bound the numerator in the second summand from above by

\[
- \frac{f(\bar{n}_{A} - x^{i}\epsilon)}{2\bar{n}_{A}} + \frac{f(1 + 2x^{i}\epsilon)}{8\bar{n}_{A}y_{A}\Delta x^{2}} - \frac{\Delta}{2} + O(x^{2i}\epsilon^{2}) \\
\leq - \frac{f(f + \Delta)(\theta - \frac{\Delta}{2}) + \Delta \theta(f + \Delta)}{2(f + \frac{\Delta}{2})(f + \theta)} + O(x^{i}\epsilon). \tag{5.163}
\]

Since \( \frac{\Delta}{2} < \theta \) there exists a constant \( C_{u} > 0 \) such that

\[
\mathbb{P}[Y_{n+1} = k + 1|Y_{n} = k] \leq \frac{1}{2} - C_{u} = p_{u}. \tag{5.164}
\]

\( \Box \)

By replacing \( p_{aa} \) by \( p_{u} \), we couple \( Y_{n}^{aa} \) in the same way with a process \( Z_{n}^{u} \) as it was done in Step 2. Observe that by construction \( Z_{n}^{u} \rhd Y_{n}^{aa} \), a.s.. The marginal distribution of \( Z_{n}^{u} \) is the desired Markov chain with transition probabilities

\[
\mathbb{P}[Z_{n+1}^{u} = k + 1|Z_{n}^{u} = k] = p_{u}, \tag{5.165}
\]

\[
\mathbb{P}[Z_{n+1}^{u} = k - 1|Z_{n}^{u} = k] = 1 - p_{u}, \tag{5.166}
\]

and invariant measure

\[
\mu(n) = \frac{\prod_{k=1}^{n} \left( \frac{1}{2} - C_{u} \right)}{\prod_{k=1}^{n} \left( \frac{1}{2} + C_{u} \right)} = \left( \frac{1}{2} - C_{u} \right)^{n-1} \left( \frac{1}{2} + C_{u} \right)^{n}. \tag{5.167}
\]
We define the stopping times
\[
T_i^+(aa) \equiv \inf \{ t \geq 0 : Z_n^a \geq \gamma_{\Delta} x^{2i+2} e^2 K + 3 M_{aa} (x^{2i} e^2)^{1+\alpha} K \}, \quad (5.168)
\]
\[
T_i^-(aa) \equiv \inf \{ t \geq 0 : Z_n^a \leq \gamma_{\Delta} x^{2i+2} e^2 K \}. \quad (5.169)
\]

Again with the formula of the equilibrium potential we estimate for \( \gamma_{\Delta} x^{2i+2} e^2 K \leq z K < \gamma_{\Delta} x^{2i} e^2 K + M_{aa} (x^{2i} e^2)^{1+\alpha} K \)
\[
\P_{z K} [T_i^+(aa) < T_i^-(aa)] = \frac{\sum_{n=\gamma_{\Delta} x^{2i+2} e^2 K}^{z K} (1+2 C_u) \alpha K} {\sum_{n=\gamma_{\Delta} x^{2i+2} e^2 K+1}^{\gamma_{\Delta} x^{2i} e^2 K+3 M_{aa} (x^{2i} e^2)^{1+\alpha} K} (1+2 C_u) \alpha K} \leq \bar{C} K^{\alpha/2} \exp(-\bar{C} K^{1/2+\alpha}). \quad (5.170)
\]

**Part 2** Similarly to Step 6, we calculate an upper bound on the time which \( n_{aa}(t) \) needs at most to decrease from \( \gamma_{\Delta} x^{2i} e^2 + M_{aa} (x^{2i} e^2)^{1+\alpha} \) to \( \gamma_{\Delta} x^{2i+2} e^2 \).

**Proposition 5.3** Let
\[
\theta_i(aa) \equiv \inf \{ t \geq 0 : n_{aa}(t) \leq \gamma_{\Delta} x^{2i+2} e^2 | n_{aa}(0) = \gamma_{\Delta} x^{2i} e^2 + M_{aa} (x^{2i} e^2)^{1+\alpha} \}, \quad (5.171)
\]
be the decay time of \( n_{aa}(t) \), for \( t \in [\tau_A^i, \tau_{aA}^i \land \tau_{AA}^{i+1} \land e^{\gamma K^{\alpha}}] \). Then there exist finite, positive constants \( C^a_u \) and \( M \) such that for all \( 0 \leq i \leq \lfloor \frac{\ln(K^{1/4-\alpha})}{\ln(x)} \rfloor \)
\[
\P[\theta_i(aa) > C^a_u] \leq \exp(-M K^{1/2+2\alpha}). \quad (5.172)
\]

**Proof** The proof works like the one of Proposition 3.8. We calculate an upper bound on the decay time of the majorising process \( Z_n^a \). Let \( Y_n^{aa} \) and \( Z_n^a \) be defined as in the step before and let \( W^a \) be i.i.d. random variables with
\[
\P[W^a = 1] = \frac{1}{2} - C_u, \quad \P[W^a = -1] = \frac{1}{2} + C_u \quad \text{and} \quad \E[W^a] = -2 C_u. \quad (5.173)
\]

The \( W^a \)'s record a birth or a death event of \( Z_n^a \). Similar as in Step 6, we choose \( N = \frac{\gamma_{\Delta}(1-x^2)}{C_u} x^{2i} e^2 K =: \bar{C} x^{2i} e^2 K \) and show that \( \P[n_\ast \leq \frac{\gamma_{\Delta}(1-x^2)}{C_u} x^{2i} e^2 K] \geq 1 - \exp(-K^{1/2+2\alpha} C_u \gamma_{\Delta}(1-x^2)/2). \) It holds
\[
\lambda_{aa}(t) = f N_{aa}(t) \left( 1 - \frac{N_{aa}(t)}{2 N_{aa}(t)} \right) + f \frac{K}{4 \Sigma(t)} n_{aA}(t) + N_{aa}(t) \{ D + \Delta + c \Sigma(t) \}
\geq \frac{f}{4 n_{aA}} x^{2i+2} e^2 K + (f + \Delta) \gamma_{\Delta} x^{2i+2} e^2 K
\geq \frac{2 f + \frac{\Delta}{4 n_{aA}} x^{2i+2} e^2 K \equiv C_{\lambda} x^{2i+2} e^2 K \equiv \lambda_{aa}. \quad (5.174)
\]

\( \textcopyright \ Springer \)
Again, let \( \tau_m - \tau_{m-1} \) be the time between two jumps of \( n_{aa}(t) \) and let \( J_m^{aa} \) be i.i.d. exponential random variables with parameter \( \lambda'_{aa} \). As in Step 6, using the exponential Chebyshev inequality, we get that

\[
P\left[ \sum_{m=1}^{\tilde{C} x^{2i}2^K} J_m^{aa} > \frac{2\tilde{C}}{C\lambda x^2} \right] \leq \exp(-\tilde{C} K^{1/2+2\alpha}) \tag{5.175}
\]

and hence

\[
P\left[ \sum_{m=1}^{n_\alpha} J_m^{aa} > \frac{2\tilde{C}}{C\lambda x^2} \right] \leq \exp(-M K^{1/2+2\alpha}). \tag{5.176}
\]

**Part 3** We see that \( n_{aa}(t) \) reaches \( \gamma\Delta x^{2i+2}e^2 \) before \( n_{aA}(t) \) decreases to \( x^{i+1}\epsilon \). Similar to Step 2 we can show that once \( n_{aa}(t) \) hits \( \gamma\Delta x^{2i+2}e^2 \) it will stay close to it an exponentially long time and will not exceed \( \gamma\Delta x^{2i+2}e^2 + M_{aa}(x^{2i+2}e^2)^{1+\alpha} \) again. Thus we can ensure that \( n_{aa}(t) \) is below \( \gamma\Delta x^{2i+2}e^2 + M_{aa}(x^{2i+2}e^2)^{1+\alpha} \) when \( n_{aA}(t) \) reaches \( x^{i+1}\epsilon \).

**Step 8** For the iteration we have to ensure that the sum process increases on the given time interval from the value \( \bar{n}A \) to \( \bar{n}A \) once it hits \( \bar{n}A \). Thus, with the knowledge of Step 3, we show that, for \( \Sigma(0) \geq \bar{n}A - \frac{\Delta+\theta}{cn_A} \gamma\Delta x^{2i}e^2 - M_{\Sigma}(x^{2i}e^2)^{1+\alpha} \). As in the proof of Proposition 3.9 we divide the proof into three parts.

**Part 1** Similarly to Part 1 in Step 7, we show that with high probability \( \Sigma(t) \) increases to \( \bar{n}A - \frac{\Delta+\theta}{cn_A} \gamma\Delta x^{2i+2}e^2 \) before going back to \( \bar{n}A - \frac{\Delta+\theta}{cn_A} \gamma\Delta x^{2i+2}e^2 - 3M_{\Sigma}(x^{2i}e^2)^{1+\alpha} \). We define stopping times on \( \Sigma(t) \):

\[
\tau_{\Sigma^-}^i \equiv \inf\{t \geq 0 : \Sigma(t) \leq \bar{n}A - \frac{\Delta+\theta}{cn_A} \gamma\Delta x^{2i}e^2 - 3M_{\Sigma}(x^{2i}e^2)^{1+\alpha} \equiv v(\Sigma_-) - 3M_{\Sigma}(x^{2i}e^2)^{1+\alpha} \}, \tag{5.177}
\]

\[
\tau_{\Sigma^+}^i \equiv \inf\{t \geq 0 : \Sigma(t) \geq \bar{n}A - \frac{\Delta+\theta}{cn_A} \gamma\Delta x^{2i+2}e^2 \equiv v(\Sigma_+) \}, \tag{5.178}
\]

where \( M_{\Sigma} > 0 \).

**Proposition 5.4** There are constants \( \tilde{C}, \tilde{C} > 0 \) such that

\[
P[\tau_{\Sigma^-}^i < \tau_{\Sigma^+}^i | \Sigma(0) \geq \bar{n}A - \frac{\Delta+\theta}{cn_A} \gamma\Delta x^{2i}e^2 + M_{\Sigma}(x^{2i}e^2)^{1+\alpha}] \leq \tilde{C} K^{\alpha/2} \exp(-\tilde{C} K^{7/2(\alpha+2\alpha^2)}). \tag{5.179}
\]

**Proof** From the step before we know that \( n_{aa}(t) \) decreases under the value \( \gamma\Delta x^{2i+2}e^2 + M_{aa}(x^{2i+2}e^2)^{1+\alpha} \) in a time of order 1 and does not exceed this bound once it hits it. Thus, with the knowledge of Step 3, we show that, for \( t \in [\tau_{\Sigma^-}^i + \theta_i(aa), \tau_{\Sigma^+}^i \wedge \tau_{\Sigma^+}^{i+1}-] \), the sum process has the tendency to increase and
exceed the lower bound $\tilde{n}_A - \frac{\Delta + \vartheta}{c_{\alpha}} \gamma_{\Delta} x^{2i+2} \epsilon^2$ before $n_{aA}(t)$ hits $x^{i+1} \epsilon$. As before, let $Y^\Sigma_n$ be the associated discrete process to $\Sigma(t)$.

**Lemma 5.10** For $n \in \mathbb{N}$ such that $\vartheta_n \in [\tau_{aA}^{-i} + \vartheta_1(aa), \tau_{aA}^{i+1} - \vartheta_{aA} \wedge e^Y K^a]$ there exists a constant $C_0 > 0$ such that

$$\mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k] \leq \frac{1}{2} - C_0 x^{2i+2} \epsilon^2 \equiv p_0(\Sigma).$$  \hspace{1cm} (5.180)

**Proof** To show the lemma we use the results of the steps before. It holds:

$$\mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k] \leq \frac{1}{2} + \frac{-\frac{1}{2}(f - D) + \frac{1}{2} c k K^{-1} + \frac{\Delta}{ck} N_{aa}(t)}{f + D + ck K^{-1} + \frac{\Delta}{ck} N_{aa}(t)}. \hspace{1cm} (5.181)$$

We estimate the nominator

$$\leq -\frac{f - D}{2} + \frac{1}{2} c k K^{-1} + \frac{\Delta}{2} \gamma_{\Delta} x^{2i+2} \epsilon^2 k^{-1} K + \mathcal{O}((x^{2i+2} \epsilon^2)^{1+\alpha}).$$  \hspace{1cm} (5.182)

This term assumes its maximum at $k = v(\Sigma_n)$. If we insert this bound, $\tilde{n}_A - \frac{\Delta + \vartheta}{c_{\alpha}} \gamma_{\Delta} x^{2i+2} \epsilon^2$, we can estimate

$$\leq -\frac{f - D}{2} + \frac{1}{2} c \tilde{n}_A - \frac{\Delta}{2 c \tilde{n}_A} \gamma_{\Delta} x^{2i+2} \epsilon^2 + \frac{\Delta}{2 \tilde{n}_A} \gamma_{\Delta} x^{2i+2} \epsilon^2 + \mathcal{O}((x^{2i+2} \epsilon^2)^{1+\alpha})$$

$$\leq -\frac{\vartheta}{2 \tilde{n}_A} x^{2i+2} \epsilon^2 + \mathcal{O}((x^{2i+2} \epsilon^2)^{1+\alpha}).$$  \hspace{1cm} (5.183)

Thus, there exists a constant $C_0 > 0$ such that

$$\mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k] \leq \frac{1}{2} - C_0 x^{2i+2} \epsilon^2. \hspace{1cm} (5.184)$$

Again we couple $Y^\Sigma_n$ with $Z^l_n$ via:

1. $Z^l_0 = Y^\Sigma_0$,
2. $\mathbb{P}[Z^l_{n+1} = k + 1 | Y^\Sigma_n > Z^l_n = k] = 1 - p_0(\Sigma)$,
3. $\mathbb{P}[Z^l_{n+1} = k - 1 | Y^\Sigma_n > Z^l_n = k] = p_0(\Sigma)$,
4. $\mathbb{P}[Z^l_{n+1} = k - 1 | Y^\Sigma_n = Z^l_n = k] = 1$,
5. $\mathbb{P}[Z^l_{n+1} = k - 1 | Y^\Sigma_{n+1} = k + 1, Y^\Sigma_n = Z^l_n = k] = \frac{p_0(\Sigma) - \mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k]}{1 - \mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k]}$,
6. $\mathbb{P}[Z^l_{n+1} = k + 1 | Y^\Sigma_{n+1} = k + 1, Y^\Sigma_n = Z^l_n = k] = 1 - \frac{p_0(\Sigma) - \mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k]}{1 - \mathbb{P}[Y^\Sigma_{n+1} = k - 1 | Y^\Sigma_n = k]}$.  \hspace{1cm} (5.189)
Observe that by construction $Z_n^l \preceq Y_n^\Sigma$, a.s.. The marginal distribution of $Z_n^l$ is the desired Markov chain with transition probabilities

\begin{align}
\mathbb{P}[Z_{n+1}^l = k + 1 | Z_n^l = k] &= 1 - p_0(\Sigma), \\
\mathbb{P}[Z_{n+1}^l = k - 1 | Z_n^l = k] &= p_0(\Sigma),
\end{align}

and invariant measure

\begin{equation}
\mu(n) = \frac{\prod_{k=1}^{n-1} \left( \frac{1}{2} + C_0 x^{2i+2} \epsilon^2 \right)}{\prod_{k=1}^{n} \left( \frac{1}{2} - C_0 x^{2i+2} \epsilon^2 \right)} = \left( \frac{1}{2} + C_0 x^{2i+2} \epsilon^2 \right)^{n-1} \left( \frac{1}{2} - C_0 x^{2i+2} \epsilon^2 \right)^n.
\end{equation}

Again we get a bound on the harmonic function, for $\bar{n}_A - \frac{\Delta + \theta}{\epsilon n A} \gamma \Delta x^{2i} \epsilon^2 - M_{\Sigma}(x^{2i} \epsilon^2)^{1+\alpha} \leq z < \bar{n}_A - \frac{\Delta + \theta}{\epsilon n A} \gamma \Delta x^{2i} \epsilon^2$,

\begin{align}
\mathbb{P}_{\Sigma} \left[ T_{\Sigma}^{l_{-}} < T_{\Sigma}^{l_{+}} \right] &= \frac{\sum_{n=\epsilon}^{K_{\Sigma}} (1 - 2C_0 x^{2i+2} \epsilon^2) (1 + 2C_0 x^{2i+2} \epsilon^2)}{\sum_{n=\epsilon}^{K_{\Sigma}} (1 - 2C_0 x^{2i+2} \epsilon^2) (1 + 2C_0 x^{2i+2} \epsilon^2)} \leq \tilde{C} K^{\alpha/2} \exp(-\tilde{C} K^{7/2\alpha + 2\alpha^2}).
\end{align}

\textbf{Part 2} As in Step 6, to calculate an upper bound on the time which $\Sigma(t)$ need to increase from $\bar{n}_A - \frac{\Delta + \theta}{\epsilon n A} \gamma \Delta x^{2i} \epsilon^2 - M_{\Sigma}(x^{2i} \epsilon^2)^{1+\alpha}$ to $\bar{n}_A - \frac{\Delta + \theta}{\epsilon n A} \gamma \Delta x^{2i} \epsilon^2$, we estimate the number of jumps of the sum process and the duration of one single jump from above on a given time interval.

\textbf{Proposition 5.5} Let

\begin{equation}
\theta_i(\Sigma) \equiv \inf \{ t \geq 0 : \Sigma(t) \geq \bar{n}_A - \frac{\Delta + \theta}{\epsilon n A} \gamma \Delta x^{2i+2} \epsilon^2 \} \Sigma(0) = \bar{n}_A - \frac{\Delta + \theta}{\epsilon n A} \gamma \Delta x^{2i} \epsilon^2 - M_{\Sigma}(x^{2i} \epsilon^2)^{1+\alpha},
\end{equation}

be the growth time of $\Sigma(t)$ on the time interval $t \in [\tau_{aA}^j - \theta_i(aa), \tau_{aA}^j + \tau_{aA}^{(i+1)-}]$. Then there exist finite, positive constants, $C_i^\Sigma$ and $M$, such that for all $0 \leq i \leq \frac{-\ln(K^{1/4-\alpha})}{\ln(x)}$

\begin{equation}
\mathbb{P}[\theta_i(\Sigma) > C_i^\Sigma] \leq \exp(-MK^{4\alpha}).
\end{equation}

\textbf{Proof} Let $Y_n^\Sigma$ be defined as in the step before and let $W_k$ be i.i.d. random variables with

\begin{align}
\mathbb{P}[W_k = 1] &= \frac{1}{2} + C_0 x^{2i+2} \epsilon^2, \\
\mathbb{P}[W_k = -1] &= \frac{1}{2} - C_0 x^{2i+2} \epsilon^2 \text{ and } \mathbb{E}[W_1] = 2C_0 x^{2i+2} \epsilon^2.
\end{align}
which record a birth or a death event of the lower process $Z^l_n$. We choose $N = \frac{\Delta + \theta}{4C_0cn_A}\chi(1 - x^2)K =: \tilde{C}K$ and show that $P\left[n_\ast \leq \frac{\Delta + \theta}{2C_0cn_A}\chi(1 - x^2)K\right] \geq 1 - \exp(-K^{4\alpha}C_0\chi(\Delta + \theta)x^2(1 - x^2)/2)$. We estimate from above the time the process $\Sigma(t)$ needs to make one jump:

$$\lambda_\Sigma(t) = f \Sigma(t)K + \Sigma(t)K[D + \Delta + c\Sigma(t)] + \Delta N_{aa}(t) \geq C_\lambda K \equiv \lambda_\Sigma^l.$$ \hfill (5.198)

As before let $\tau_m - \tau_{m-1}$ the time between two jumps of $\Sigma(t)$ and let $J_m^\Sigma$ are i.i.d. exponential random variables with parameter $\lambda_\Sigma^l$. As in Step 6, by applying the exponential Chebyshev inequality, we get that $P\left[\sum_{m=1}^{C_K} J_m^\Sigma > \frac{2\tilde{C}}{C_\lambda}\right] \leq \exp(-\tilde{K}\tilde{C}/2)$ and hence

$$P\left[\sum_{m=1}^{n_\ast} J_m^\Sigma > \frac{2\tilde{C}}{C_\lambda}\right] \leq \exp(-MK^{4\alpha}).$$ \hfill (5.199)

\hspace{1cm} \Box

**Part 3** For the iteration we have to ensure that once $\Sigma(t)$ hits the upper bound $\tilde{n}_A - \frac{\theta + \Delta}{c\tilde{N}_A} \chi x^{2i+2}\epsilon^2$ it will stay close to it for an exponential time on the given time interval.

This ensures that the sum process stays in a small enough neighbourhood of $\tilde{n}_A - \frac{\theta + \Delta}{c\tilde{N}_A} \chi x^{2i+2}\epsilon^2$ when $n_{aA}(t)$ hits $x^{i+1}\epsilon$ and the next iteration step can start.

**Proposition 5.6** Assume that $\Sigma(0) = \tilde{n}_A - \frac{\theta + \Delta}{c\tilde{N}_A} \chi x^{2i+2}\epsilon^2$. For all $M > 0$, let

$$\hat{\tau}_\Sigma^\alpha \equiv \inf \left\{ t > 0 : \Sigma(t) - \tilde{n}_A + \frac{\theta + \Delta}{c\tilde{N}_A} \chi x^{2i+2}\epsilon^2 \leq -M(x^{2i+2}\epsilon^2)^{1+\alpha} \right\}.$$ \hfill (5.200)

*Then there exists a constant $M_\Sigma > 0$ such that

$$P[\hat{\tau}_\Sigma^\alpha < \tau_{aA}^{i+1} \land \tau_{aA}^{(i+1)-} \land e^VK^\alpha] = o(K^{-1}).$$ \hfill (5.201)

**Proof** The proof is similar to Step 3. Again we define the difference process $\hat{X}_t$ between $\Sigma(t)K$ and $\tilde{n}_A K - \frac{\theta + \Delta}{c\tilde{N}_A} \chi x^{2i+2}\epsilon^2 K$, which is a branching process with the same rates as $\Sigma(t)$:

$$\hat{X}_t = \Sigma(t)K - \tilde{n}_A K + \frac{\theta + \Delta}{c\tilde{N}_A} \chi x^{2i+2}\epsilon^2 K$$ \hfill (5.202)

$$\hat{X}_0^X \equiv \inf\{t \geq 0 : \hat{X}_t = 0\}$$ \hfill (5.203)

$$\hat{X}_{a,M_\Sigma} \equiv \{ t \geq 0 : \hat{X}_t \leq -M_\Sigma(x^{2i}\epsilon^2)^{1+\alpha} \}.$$ \hfill (5.204)

Let $\hat{Y}_n$ be the discrete process associated to $\hat{X}_t$, obtained as described in Step 1.
Lemma 5.11 For \( n \in \mathbb{N} \) such that \( \hat{\theta}_n \in [\tau_{aA}^{-i-} + \theta_i(aa), \tau_{aA}^{-(i+1)-} e^{VK_a}] \), there exists a constant \( \hat{C}_0 > 0 \) such that

\[
\mathbb{P}[\hat{Y}_{n+1} = -k - 1|\hat{Y}_n = -k] \leq \frac{1}{2} - \hat{C}_0 x^{2i+2} e^2 \equiv \hat{p}_\Sigma. \tag{5.205}
\]

The proof is a re-run of Step 3 by using the rates (5.58). The rest of the proof of Proposition 5.6 is similar to Step 3. \( \square \)

**Final step**

**Calculation of the decay time of \( n_{aA} \)** The following proves Theorem 3.2 (ii). Set

\[
\sigma \equiv -\frac{\ln(\epsilon K^{1/4-\alpha})}{\ln(x)}. \tag{5.206}
\]

Observe that \( x^\sigma \epsilon \geq K^{-1/4+\alpha} \) is just the value until which we can control the decay of \( n_{aA}(t) \). Thus, to calculate the time of the controlled decay of the \( aA \)-population we iterate the system, described above, until \( i = \sigma \). Observe that

\[
\sum_{j=0}^{\sigma-1} \frac{C_u}{x^\epsilon} = \frac{C_u}{1-x} (K^{1/4-\alpha} - \epsilon^{-1}) \geq \frac{C_l x}{1-x} (K^{1/4-\alpha} - \epsilon^{-1}) = \sum_{j=0}^{\sigma-1} \frac{C_l}{x^\epsilon} \quad \text{and that}
\]

the \( \theta_j(aA) = \tau_{aA}^{-(j+1)-} - \tau_{aA}^{j+} \) are independent random variables. Thus, for some constant \( \tilde{M} > M > 0 \), we get

\[
\mathbb{P} \left[ \frac{C_u x}{1-x} (K^{1/4-\alpha} - \epsilon^{-1}) \geq \tau_{aA}^{j+} - \tau_{aA}^{j-} \geq \frac{C_l x}{1-x} (K^{1/4-\alpha} - \epsilon^{-1}) \right] \geq \mathbb{P} \left[ \frac{C_u x}{x^\epsilon} \geq \theta_1(aa) \geq \frac{C_l}{x^\epsilon} \right. \quad \left. \ldots \right. \quad \left. \frac{C_u x}{x^\epsilon} \geq \theta_\sigma(aa) \geq \frac{C_l}{x^\epsilon} \right]
\]

\[
\geq 1 - \sigma \exp(-M K^{1/2+2\alpha}) \geq 1 - \exp(-\tilde{M} K^{1/2+2\alpha}). \tag{5.208}
\]

**Remark 5** A function \( f(t) = \frac{1}{4} \) needs a time \( t_1 - t_0 = K^{1/4-\alpha} - \frac{1}{\epsilon} \) to decrease from \( f(t_0) = \epsilon \) to \( f(t_1) = K^{-1/4+\alpha} \). Compared to the decay-time of \( n_{aA} \) we see that it is of the same order. Thus the stochastic process behaves as the dynamical system.

**Survival until mutation** Now we prove Theorem 3.3. We already know that there is no mutation before a time of order \( \frac{1}{K^{1/4-\alpha}} \) (cf. Lemma 3.1 and (3.9)). Since we have seen that the duration of the first step is \( O(\ln K) \) and the time needed for the second step is bounded, the left hand side of (3.1) ensures that the first two phases are ended before the occurrence of a new mutation. Thus we get the first statement of (3.7).

To justify the second statement of (3.7), we have to calculate an upper bound on the mutation time such that there are still enough \( aA \)-individuals in the population, when
the next mutation to a new allele occurs. We saw that we can control the process only until the $aA$-population has decreased to $K^{-1/4+\alpha}$. Thus we have to verify that the mutation time is smaller than $O \left( K^{1/4-\alpha} \right)$, the time the process $n_{aA}$ needs to decrease to $K^{-1/4+\alpha}$.

The mutation rate of the whole population is the sum of the mutation rates of each subpopulation. For $t \in [\tau_{aA}^{0+}, \tau_{aA}^{\sigma-}]$ with high probability, the new mutation occurs in the $AA$-population, because $n_{aa}(t)$ and $n_{aA}(t)$ are very small. Let

$$p_A(t) = \frac{n_{AA}(t) + \frac{1}{2}n_{aA}(t)}{n_{aa}(t) + n_{aA}(t) + n_{AA}(t)}$$

be the relative frequency of $A$-alleles in the population at time $t$. The mutation rate of the $AA$-population is given by

$$\mu_{AA}^K = \mu_K f p_A(t) n_{AA}(t) K.$$  \hspace{1cm} \text{(5.212)}$$

For $t \in [\tau_{aA}^{0+}, \tau_{aA}^{\sigma-}]$ we know from the results before that $n_{AA}$ is in an $\epsilon$-neighbourhood of its equilibrium, $\bar{n}_A$, and $n_{aA}, n_{aa} \leq \epsilon$. We can estimate $\mu_{AA}^K$:

$$\mu_{AA}^K = \mu_K f \bar{n}_A K + O(\epsilon).$$ \hspace{1cm} \text{(5.213)}$$

From (5.212) we get that the time of a new mutation is smaller or equal to $\frac{1}{f \bar{n}_A K \mu_K} + O(\epsilon)$. Thus to ensure that we still have $a$-alleles in the population, we have to choose $\mu_K$ in such a way that

$$\frac{1}{f \bar{n}_a K \mu_K} \ll \frac{C_{1x}}{1-x} K^{1/4-\alpha},$$ \hspace{1cm} \text{(5.214)}$$

since we can ensure the survival of $n_{aA}(t)$ until the time $\frac{C_{1x}}{1-x} K^{1/4-\alpha}$ (cf. (5.207)). Hence, the right hand side of (3.1) gives us that a new mutation occurs before the $aA$-population died out. This finishes the proof of Theorem 3.3.

References

Athreya KB, Ney PE (2011) T. E. Harris and branching processes. Ann Probab 39(2):429–434
Baar M, Bovier A, Champagnat N (2016) From stochastic individual-based models to the canonical equation of adaptive dynamics—in one step. Ann Appl Probab \textbf{(online first)}
Billiard S, Smadi C (2016) The interplay of two mutations in a population of varying size: a stochastic eco-evolutionary model for clonal interference. Stoch Process Appl \textbf{(in press)}
Bovier A (2006) Metastability: a potential theoretic approach. International Congress of Mathematicians, vol III. Eur Math Soc, Zürich, pp 499–518
Bovier A, den Hollander F (2015) Metastability: a potential-theoretic approach, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol 351. Springer, Cham
Bürger R (2000) The mathematical theory of selection, recombination, and mutation. In: Wiley series in mathematical computational biology. Wiley, Chichester
Champagnat N (2006) A microscopic interpretation for adaptive dynamics trait substitution sequence models. Stoch Process Appl 116(8):1127–1160
Champagnat N, Méleard S (2011) Polymorphic evolution sequence and evolutionary branching. Probab Theory Relat Fields 151(1–2):45–94
Champagnat N, Ferrière R, Ben Arous G (2002) The canonical equation of adaptive dynamics: a mathematical view. Selection 2(1–2):73–83
Champagnat N, Ferrière R, Méleard S (2008) From individual stochastic processes to macroscopic models in adaptive evolution. Stoch Models 24(suppl. 1):2–44
Collet P, Méleard S, Metz JA (2013) A rigorous model study of the adaptive dynamics of mendelian diploids. J Math Biol 67(3):569–607
Coron C (2014) Stochastic modeling of density-dependent populations and the extinction vortex. Adv Appl Probab 46(2):446–477
Coron C (2016) Slow-fast stochastic diffusion dynamics and quasi-stationarity for diploid populations with varying size. J Math Biol 72(1–2):171–202
Coron C, Méleard S, Porcher E, Robert A (2013) Quantifying the mutational meltdown in diploid populations. Am Nat 181(5):623–636
Crow JF, Kimura M (1970) An introduction to population genetics theory. Harper & Row Publishers, New York
Dieckmann U, Law R (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes. J Math Biol 34(5–6):579–612
Ewens WJ (2004) Mathematical population genetics. I. In: Interdisciplinary applied mathematics, vol 27, 2nd edn. Springer, New York
Fisher R (1918) The correlation between relatives on the supposition of Mendelian inheritance. Trans R Soc Edinb 42:399–433
Fournier N, Méleard S (2004) A microscopic probabilistic description of a locally regulated population and macroscopic approximations. Ann Appl Probab 14(4):1880–1919
Freidlin MI, Wentzell AD (1984) Random perturbations of dynamical systems, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol 260. Springer, New York
Golubitsky M, Guillemin V (1973) Stable mappings and their singularities. In: Graduate texts in mathematics, vol 14. Springer, New York
Haldane J (1924a) A mathematical theory of natural and artificial selection. Part I. Trans Camb Phil Soc 23:19–41
Haldane J (1924b) A mathematical theory of natural and artificial selection. Part II. Trans Camb Phil Soc Biol Sci 1:158–163
Hirsch MW, Pugh CC, Shub M (1977) Invariant manifolds. In: Lecture notes in mathematics, vol 583. Springer, Berlin
Hofbauer J, Sigmund K (1990) Adaptive dynamics and evolutionary stability. Appl Math Lett 3(4):75–79
Kisdi E, Geritz SA (1999) Adaptive dynamics in allele space: evolution of genetic polymorphism by small mutations in a heterogeneous environment. Evolution 60:993–1008
Marrow P, Law R, Cannings C (1992) The coevolution of predator–prey interactions: Esess and red queen dynamics. Proc R Soc Lond B Biol Sci 250(1328):133–141
Metz J, Nisbet R, Geritz S (1992) How should we define ‘fitness’ for general ecological scenarios? Trends Ecol Evol 7(6):198–202
Metz JAJ, Geritz SAH, Meszéna G, Jacobs FJA, van Heerwaarden JS (1996) Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In: Stochastic and spatial structures of dynamical systems (Amsterdam, 1995), Konink. Nederl. Akad. Wetensch. Verh. Afd. Natuurk. Eerste Reeks, vol 45. North-Holland, Amsterdam, pp 183–231
Nagylaki T (1992) Introduction to theoretical population genetics. In: Biomathematics, vol 21. Springer, Berlin
Perko L (2001) Differential equations and dynamical systems. In: Texts in applied mathematics, vol 7, 3rd edn. Springer, New York
Rouhani S, Barton N (1987) The probability of peak shifts in a founder population. J Theor Biol 126(1):51–62
Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–157
Yule G (1907) On the theory of inheritance of quantitative compound characters on the basis of Mendel’s laws—a preliminary note. In: Reports of the 3rd international congress on genetics. Spottiswoode, London, pp 140–142