THE FATE OF A RECESSIVE ALLELE IN A MENDELIAN DIPLOID MODEL

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Rebecca Sarah Ströfer

aus Leverkusen

Bonn, Mai 2017

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Anton Bovier
2. Gutachter: Prof. Dr. Patrik Ferrari

Tag der Promotion: 25.07.2017
Erscheinungsjahr: 2017

Abstract

The omnipresence of sexual reproduction in a highly competitive world is a fascinating phenomenon. Its evolution and maintenance is up to now not completely understood since it is known that sexual reproduction require from organisms a lot more costs and energy than asexual reproduction would do.

In this thesis we show mathematically that sexual reproduction provides populations an evolutionary advantage because they can better adapt to a changing ecological system. To this end, we study a stochastic individual based model which describes the genetic evolution of a diploid hermaphroditic population reproducing sexually according to Mendelian laws. This single locus model describes a population of interacting individuals that incorporate the canonical genetic mechanisms of birth, death, mutation, and competition.

In the first part of this thesis the genetic evolution of the population with two alleles is studied under the assumption that a dominant allele is also the fittest one. It is shown that after the invasion of a dominant allele in a resident population of homozygous recessive genotypes, the recessive allele survives in heterozygous individuals for a time of order at least $K^{1/2-\alpha}$, where K is the carrying capacity and $\alpha > 0$. This time of survival of the unfit allele is much longer than it would be in a population reproducing asexually. Therefore, a suitable rescaling of the mutation rate made the appearance of a new advantageous mutation possible before the extinction of the recessive allele.

In the second part of this thesis, we study the fate of the recessive allele after the occurrence of a further mutation to a more dominant allele. It is proven that resulting changes in the composition of the population indeed opens the possibility that individuals of homozygous recessive genotype can reinvade and that coexistence of different genotypes is possible. This leads to genetic variability and can be seen as a statement of genetic robustness exhibited by diploid populations performing sexual reproduction as well as an indicator for the overwhelming biodiversity in nature.

Acknowledgement

Without the support of many people I would not have been able to write this thesis. I would like to use the opportunity to express my gratitude.

First of all I want to thank my supervisor Anton Bovier: Thank you for always taking time for me and the continuous encouragement when I miss the forest for the trees. I benefited a lot from our discussions, your enormous mathematical knowledge and especially your excellent intuition which impresses me time after time. You support me in many ways, not only mathematically (and on your pannier reck). Thank you for being my "Doktorvater".

The second person who took an important role during my PhD studies is Loren Coquille: It was a pleasure for me to work together with you at so many different places. Thank you for being such a good friend (and for being my witness) and for welcoming me in Grenoble.

Special thanks go to Nikolaus Schweizer for being a friend and a mentor and to Raphael Zimmer the best room mate I can imagine.

It was a pleasure to be a member of the stochastic group and I thank all former and current colleagues for creating a friendly atmosphere, especially our kind soul Mei-Ling.

I want to express my gratitude to my husband Henrik and my family. Thank you for your patience, understanding and encouragement.

I am grateful for 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*. It was a pleasure to participate in those programs and the Bonn International Graduate School in Mathematics (BIGS).

Finally, I thank the members of my examination committee.

Contents

List of Figures

I. Introduction

In nature, sexual reproduction is the favoured kind of propagation in higher plants and animals. Most organisms belong to diploid populations which reproduce sexually. However, a comparison of the reproduction types shows that asexual reproduction is more effective for the survival of a population and less costly as the energy-sapping and inefficient partner-dependent sexual reproduction. Up to now, many more or less convincing hypotheses exist why sexual reproduction nevertheless evolved and is maintained in nature and evolutionary advantageous. Yet for a universally valid explanation which is able to outweigh the many costs of sex, biologists are still searching.

In this context, the mathematical endeavour is to construct models which mirror the real life processes or the actual facts in demand as well as possible to get some interesting conclusions on how evolution acts. The unpredictability and complexity of the biological world turns this into a big challenge. The task is to reduce the high amount of natural acting mechanisms and phenomena to the essential ones such that it is possible to build up a mathematical tractable model which gives relevant and realistic implications for the biological questions in demand. For this new methods had to and have to be developed. Such models of populations give insight into the evolutionary population process and an understanding of the acting mechanisms and give biologists a tool to find and to prove new hypothesis.

In this thesis, we use mathematical modelling to show that sexual reproduction enables competitive populations to react and adapt better to changes in the ecological system. To be more precisely, we consider a stochastic population model consisting of diploid, hermaphroditic individuals which reproduce sexually according to Mendelian laws and prove that a subpopulation, carrying a recessive allele, which is forced to extinction by the whole population can resurrect when the genetic composition of the population changes.

The study of Mendelian diploid models started at the beginning of 1900 by Yule, Fisher, Wright and Haldane, the founding fathers of population genetics. But the models in this context typically deal with finite population sizes and models inheritance at the expense of competitive interaction of populations among themselves and their environment. In our approach, we take these factors into account and consider a competitive population model, described by a stochastic system of interacting individuals, which include the genetic results of reproduction, as well as ecological dynamics. Since individuals are competing for resources, the environment of an individual it lives in and interacts with, is defined by the composition of the population. This kind of model belongs to the theory of adaptive dynamics, developed in the 90'ies, which is a variant of population genetics and models the phenotypic evolution of a population in a varying environment. It allows variable population sizes controlled by the competitive interaction of the population. However, most of

the results in the context of adaptive dynamics are based on haploid asexual reproducing population since, unfortunately, especially the Mendelian reproduction and the interaction of genotypically and phenotypically different individuals makes the modelling of sexual reproducing populations quite complicated. Therefore, the consideration of stochastic models on the individual level in a genetic setting just started in 2012.

In this thesis we pick up this line of research and show that in a Mendelian diploid model, under a dominance-recessivity assumption, there can appear a hitchhiking phenomenon of recessive alleles in heterozygous individuals which leads to their prolonged survival in the population and, moreover, after changes in the population composition and under fairly natural assumptions, to genetic variability as well as biodiversity.

The remainder of the introduction is organised as follows. First a simplified overview of the genetics needed to understand the biological background is presented followed by an introduction of the Mendelian laws. We then discuss some historical facts which lead to the theories of population genetics and adaptive dynamics and review the already existing research of diploid models. The first section ceases with the many costs of sexual reproduction and theories why it is maintained in nature nonetheless. The second section concerns the mathematical modelling approaches of the evolution of sexually reproducing populations in the context of population genetics and adaptive dynamics. In Section 3 the stochastic individual-based Mendelian diploid model studied in Chapter II and III is introduced. The contributions of this thesis, presented in detail in Chapter II and III, are shortly summarised in the last section of the introduction.

In Chapter II we examine the genetic evolution of the three genotypes aa, aA and AA of the Mendelian diploid model, introduced in Section 2 of this introduction, in the large population and rare mutation limit. The main results are, that under a dominant-recessivity assumption on the alleles, the recessive a allele can survive in the population long enough such that under a suitable rescaling of the mutation rate a new advantageous mutation can appear before its extinction. This chapter is published in the Journal of Mathematical Biology [84]. The content of Chapter III is the preprint [10] available on arXiv. It builds on Chapter II and studies the population in the large population limit, when still enough recessive a alleles are present and a mutation to a more beneficial B allele has appeared. Therein, a six dimensional deterministic system is analytically studied and under rather natural assumptions, the coexistence of the two homozygous subpopulations of genotype aa and BB is proven. Both chapters together show that sexual reproduction gives to populations an advantage because they can better adapt to a changing ecological system. It could be an indication for the overwhelming biodiversity in nature.

1. Biological background and motivation

The simplest reason why it is important to study sexual reproduction is certainly its overwhelming occurrence in nature. The vast majority of organisms in nature reproduce this way. In this section we take a closer look at this kind of reproduction from a biological viewpoint.

1.1. Genetic background

We start by introducing the minimum genetics that is needed to understand the basic mechanisms of reproduction and Mendelian inheritance. For further readings and more details we refer to Lodish [70] and Barton et al. [4].

An organism or individual is built up by one to many cells. Each of these cells has a cell membrane which separate it from the others and encloses the cytoplasm and other cell organelles. The genetic material, called *genome*, is contained in the cytoplasm and consist of *Deoxyribonucleic acid (DNA)*. In procaryotes (cells without nucleus) the DNA is located free in the cytoplasm whereas most of the genome of eukaryotes (cells with nucleus) is located in its nucleus and is organised linearly in *chromosomes* whose number is characteristic of each specie. A particular region of the DNA is called a *gene*. Each chromosome consists of many genes and the position of a gene is called a *locus*. For example the human genome is composed of 23 chromosomes and 20000-25000 genes (Human Genome Project, 2001 [65]). The set of genes defines the *genotype* of the organisms where the variant forms of a given gene are called *alleles*. Different alleles result in different *phenotypes* which are the expressions of the genotype. Consequently, the phenotype of an individual is the sum of all morphological and physiological traits plus behavioural features which are determined by the alleles (e.g. body size, hair colour, intelligence). Individuals with a single set of chromosomes are called *haploid*, and with more than two copies *polyploid*. In this thesis we consider sexually reproducing *diploid* individuals, meaning organisms whose cells have two homologous copies of each chromosome, usually one of the mother and one of the father. Thus, for each gene the individual can carry two variant forms. If the alleles of the two genes at a particular locus in a diploid individual are the same it is called *homozygous*, otherwise *heterozygous*. An allele is called *recessive* if it is only expressed in the homozygotes otherwise it is called *dominant*. As a consequence heterozygotes which consist of one dominant and one recessive alleles always express the dominant character in their phenotypes. If an organism reproduces asexually, like every prokaryote, then the offspring's genome is an exact copy of the individual's one unless *mutations* alter alleles. This is why it is also termed *clonal* reproduction. In contrast, during the sexual reproduction the diploid organism first creates gametes (sperm or egg) by a mechanism of cell devision, called *meiosis*, which are haploid cells and contain a single set of the individual chromosomes. During fertilisation two gametes of two different individuals fuse together to a *zygote*, from which the offspring develops. The zygote is diploid again and contains one set of chromosomes of each parent.
Therefore, the genome is the base for heredity and the basic laws how genes from the parents are transmitted to the offsprings were first studied by the Austrian monk Gregor Mendel during seven years (1857-1863) of hybridisation experiments with peas plants.

1.2. The contribution of Mendel to evolution

Gregor Mendel published his work, *Versuche über Pflanzen-Hybride* [73], 1866 just seven years after Charles Darwin's publication of *The origin of species* [25] where he argued that *natural selection* would be the driving force of evolution. Darwin defined natural

selection as the process of unavoidable selection favouring individuals best adapted to their environment. His work was the starting point of a big, controversially discussion about evolution under biologists (see [32]). In contrast, Mendel's results found less attention and its importance for inheritance was unregarded and forgotten [77]. Only after his death in January 1884, Mendel's work was rediscovered in 1900 by Hugo de Vries [26], Carl Correns [22] and Erich von Tschermak [94]. The Mendelian laws known today were not originally stated by Mendel but named in his honour:

1. **Law of independent segregation**
   During meiosis, the two alleles for each gene segregate from each other and are distributed to different gametes. Therefore, the gametes only contain one of the two alleles for each trait. At fertilisation, two gametes of two different individuals fuse together and define the genotype of the offspring. Consequently, the offspring is diploid again and receives a pair of alleles for each trait one from each parent (see Figure I.1 (left)).

2. **Law of independent assortment**
   During meiosis, the segregation of the two alleles of one allelic pair is independent of the segregation of the two alleles of another allelic pair. (see Figure I.1 (right)).

3. **Law of dominance**
   The phenotypic trait is always defined by the dominant allele (in Figure I.1 (left) the allele "g" and alleles "C" and "T" (right)).

These laws are only presented as hypotheses in Mendel's paper [73]. De Vries was the first who mentioned the expression "Mendel's law" and it was rigorously defined by Correns [22] in acknowledgment of Mendel's work. But both did not distinguish between different laws. Morgan first explicitly talked about the law of independent segregation and the law of independent assortment [78] and integrated it with the Boveri-Sutton chromosome theory of 1915 in the book "The Mechanism of Mendelian Heredity" [79]. It was Sir Ronald Aylmer Fisher, in 1930, who worked out that Darwin's theory of natural selection does not stand in contradiction with Mendelian inheritance and how the two theories needs to be combined [34, 35]. He considered evolution in a mathematical framework and is one of the founding-fathers of population genetics.

1.3. Sexual reproduction - a biological puzzle

The evolution of sexual reproduction, precisely its origin and its maintenance in a highly competitive world, is still a major problem in biology. Sexual reproduction compared to asexual reproduction seems to be very costly (e.g. [23, 69, 93]). For a start, there are the *cellular-mechanical costs* [68] which correspond to the large amount of cellular mechanisms (meiosis, fusion of gametes *(syngamy)* and of nuclei *(karyogamy)*) which sexual reproduction requires and which take a long time and energy. Furthermore, during meiosis, genetic recombination can destroy advantageous allele combinations which can result in a decrease of the individual's fitness and selection can act again [83]. Also noteworthy

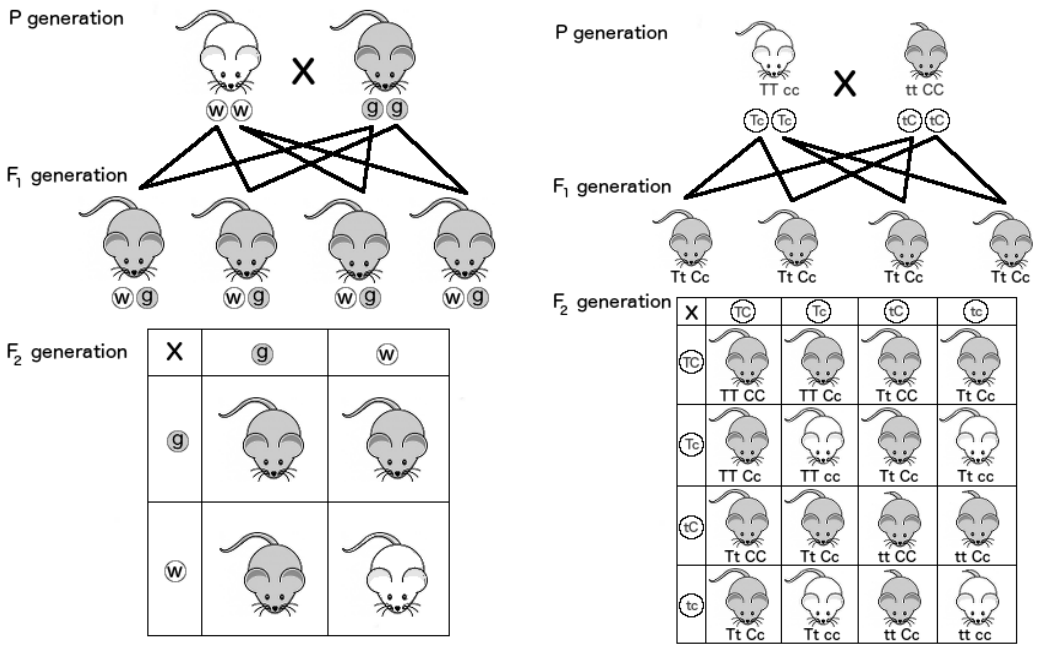

Figure I.1.: Left: Law of independent segregation (grey colour is dominant, white colour is recessive).
Right: Law of independent assortment (grey colour and long tail are dominant, white colour and short tail are recessive).

are the costs for searching for potential partners, and the possibility of spreading diseases between the partners [86]. Additionally, in sexually reproducing populations there is the famous *twofold cost* of sex ([93], see Figure I.2) for producing males. To make this concrete, in asexually reproducing population there is only one sex, precisely each individual owns the female function, i.e. it gives birth to offsprings which again own the female function. By contrast, in sexually reproducing populations there are both sexes present, females and males, at the same fraction. But only females give birth to offsprings of these 50 percent are males and 50 percent are females. Therefore, asexual females invest their full energy in producing offsprings owning the female function whereas sexual females invest 50 percent of their resources in producing males which for their part cannot give birth to offsprings. To be more precisely, assuming two offsprings per female by reproduction, an asexual population double the size of individuals owning the female function whereas in a sexual population the amount of females stay constant (see Figure I.2). In this thesis, we consider a *hermaphrodite* sexual population (organisms with both sexes), where each individual can undertake the male or the female function by reproduction. Nevertheless, also such population grow more slowly compared to asexual ones due to the needed partner for reproduction. To be more precisely, assuming an asexual and a sexual population of the same size and that one reproduction event results in the same amount of offsprings in both populations. Then, the asexual one has double the amount of offsprings compared to the

sexual one because therein each individual produces descendants whereas in the sexually reproducing population two individuals are needed for the same amount of descendants.

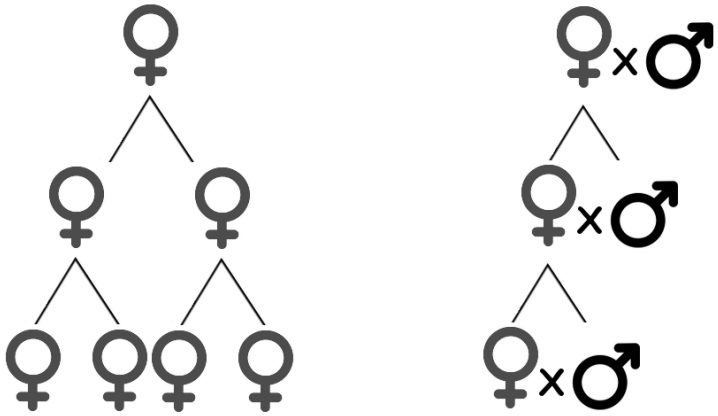

Figure I.2.: Twofold cost of sex: asexual reproduction (left), sexual reproduction (right)

Since sexual reproduction is such a costly endeavour one would suggest that it is rare in nature. But quite to the contrary, the vast majority of eukaryotic organisms reproduce sexually and only 0.1% reproduce exclusively asexually [99]. Consequently, there must be advantages for sexual populations which can outweigh the costs. Many models and theories were developed to ensure the evolution and maintenance of sexual reproduction (e.g. [6, 13, 87]). We will summarise briefly the most popular and respected ones.

1. **Fisher-Muller hypothesis** [35, 80]
   Sexual reproduction accelerates the process of fixation of advantageous mutations by combining beneficial mutations which initially appeared in different individuals into the same genome. Comparatively, in asexual populations a second beneficial mutation will only fixate if it appears by reproduction of an individual carrying already the first beneficial mutation.

2. **Muller's ratchet** [33, 81]
   This hypothesis states that there is an accumulation of deleterious mutations in small asexual populations. Ignoring backward or compensatory mutations, in each generation there is a probability that the class of individuals with the least number of deleterious mutation will disappear due to mutation at the same time or due to the stop of reproduction. Thus, in future generations there will be no individuals with fewer deleterious mutations. In this way, the number of mutations in the population increases. Comparatively, in sexual population, due to recombination, these classes of individuals with the least number of deleterious mutations can be recreated.

3. **Red-Queen hypothesis** [5, 98]
   This is an environmental hypothesis describing that sexual reproduction enables individuals to adapt and evolve constantly to survive in coevolved interaction with

other organisms in a deterministically changing environment. Thereby, recombination can create genotypes which are better adapted and have a competition advance. In coevolution of established populations, this leads to sustained oscillations between genotype frequencies. The best-known application is the host-parasite, coevolution, showing that sexual reproduction implies the ability for hosts to resist parasitic infection. Parasites may thus be the driving force for the maintenance of sexual reproduction in their host [47, 48, 53].

4. **Mutational deterministic hypothesis** [62, 64]
   Under the assumption that each additional deleterious mutation decreases the fitness, sexually reproducing individuals have a short-term advantage because the amount of mutations passed to the offsprings has a big variance. Since offsprings with a majority number of deleterious mutations have also low fitness they will die out and with them a vast number of mutations is eliminated.

Under strict assumptions all these theories give reasons for the evolution and maintenance of sexual reproduction but one generally valid theory is not found so far. It is reasonable to take a pluralistic approach [50, 102] to explain the omnipresence of sexual reproduction in nature. But up to now it remains a big puzzle and further work and rigorous studies of population models are needed.

2. Modelling of sexual reproduction

Our contribution in this direction is to show in a rigorous mathematical framework advantages of sexual compared to asexual reproduction by incorporating ecological dynamics and the random elements of evolution in a population model with Mendelian reproduction. To be more precise, we show that in such a population, the time until a recessive allele at one locus becomes extinct is prolonged since it survives in the fitter heterozygous individuals. If the genetic composition of the population changes, the population of homozygous individuals carrying this recessive allele can recover and can probably live in coexistence with other genetic different subpopulations. In this way, sexually reproducing populations can better adapt to changes in the ecological system because recessive alleles, which can be more advantageous in the changed environment, survive in heterozygotes. Under fairly natural assumptions, this can also lead to genetic diversity, as we will show.

2.1. The evolutionary process

To understand the evolution of a population it is necessary to take a closer look at this process of high complexity. The main acting mechanisms driving the evolutionary process are:

- **heredity** - transmits the parental characteristics (morphological and physiological properties) to the offspring,

- **variation** - changes the characteristics of an individual (e.g. due to mutations, genetic recombinations),
- **selection** - acts onto the different reproduction and survival ability (fitness) of the individuals, a consequence of competition between individuals.

As already said, each individual is characterised by its genotype which is expressed in its phenotype. The different phenotypes of a population compete for resources (e.g. water, space, food), or with other species (e.g. predator, prey, parasites). This competition process changes the mortality and fertility of each individual and affects the selection process acting onto the viability of the different phenotypes. Consequently, the adaptation of a phenotype depends on its environment which in turn is influenced by the phenotypes of the present population. Therefore, the population dynamics is in interaction with the environment and also influences the genotypes through mutations in the genome and additionally in sexually reproducing populations through meioses and recombination. Moreover, the environment has an effect on the translation of the genotype into the phenotype (e.g during its lifetime). This mapping is up to now not completely understood [89].

Consequently, the variable environment of a population plays a crucial role for its evolution and it is unavoidable to include it in a mathematical population model to be as realistic as possible. Unfortunately, the whole complexity of the evolutionary process is not mathematically feasible and for modelling there are some simplifications needed at the expense of realistic factors. The rest of this section is concerned with approaches in this direction with a focus on Mendelian diploid populations.

2.2. Population genetics

A first attempt for modelling Mendelian inheritance was done by Fisher, Haldane and Wright who are the founding fathers of the population genetics theory developed in the 1920's and 1930's [24, 32]. The theory integrates Mendel's inheritance theory into Darwin's theory of natural selection for modelling evolution. Population genetics does not focus on individuals but rather deals with the changes of allele frequencies in the whole population over time [82]. In this approach, the selection advantage of an individual, called *fitness*, depend on its genotype. It is defined *a priori* and is equal to the expected number of progeny [82]. Consequently, selection is directly acting on genotypes instead of phenotypes and changes allele frequencies in accordance with their different fitnesses. The interaction of individuals with their environment is mostly omitted. Therefore, in these models the *adaptive landscape*, first introduced by Wright in 1932 [104], is fixed and selection drives the population in this fitness landscape upwards to a local maximum [103, 105, 106]. In this way, population genetics reduces the complexity of the real world and gives insight into the complex patterns of genetic variation. It includes observational, experimental and theoretical components [30, 32] and studies the dynamics and statics of evolution: first arrived by changes in gene frequency under selection, latter by indicating factors which maintain the population in an approximate equilibrium [24].

One milestone in this context and the traditional starting point of population genetics is the *Hardy-Weinberg-law*, published in 1908 by Hardy [49] and Weinberg [101]. It states

that, in the absence of gene frequency changing factors (mutation, selection, migration, random drift), in an infinite population which has non-overlapping generations and mates randomly, the genetic variation is conserved. To be more precisely, the gene frequencies stay constant from one generation to the next from the second generation on [24, 32].
With the newly discovered molecular understanding of inheritance and the availability of genomic data in the 60'ies, a new theory based on population genetics to treat evolution, especially on the molecular level, arose. In 1968 Kimura [54,55], independently from King and Jukes (1969, see [57]), developed the neutral theory of molecular evolution which states that most of the variation within and between populations on a molecular level is achieved by genetic drift (change in allele frequency due to random sampling) of *neutral mutations* which do not affect the fitness of the organism such that natural selection is not acting on them. It was extended to nearly neutral mutations which are only slightly deleterious in 1973 by Ohta [85]. The theory is not in contradiction with Darwins theory of natural selection as driving force for adaption, as Kimura pointed out [54, 82]. It only considers evolution on the molecular level and states that phenotypic evolution is still controlled by natural selection [54–56].

Despite this extensive knowledge about the genome structure and its evolution, achieved by population genetics, the formation of new species in this context is not well understood [75].

Nevertheless, the genetic aspect of the models in population genetics provides a good tool to give insights for the overwhelming occurrence of sexual reproduction in nature. These models are used, among others, to examine the genetic processes of segregation and recombination, arrived by meiotic crossover, which are only present in sexually reproducing populations and therefore have to be advantageous. For example Kirkpatrick and Jenkins [58] use this theory to show that segregation could be one reason for the maintenance of sex. They argue that in sexually reproducing diploid populations the substitution of an advantageous mutation happen faster since it needs only the beneficial mutation of one allele in individuals whereas in an asexual diploid population it needs this mutation two times in the same lineage (the second mutation in a descendant of the particular individual where the first took place). Further examples like a critical analysis of the contribution of recombination to the maintenance of sex and of Mueller's ratchet as well as the maintenance of sex due to synergistic epistasis (the interaction of alleles at different loci), can be found in [24, 42, 63]. For more insights on population dynamics, we refer to [24, 32, 82].

2.3. Adaptive dynamics

One weakness of the models of population genetics is that they mostly neglect the important aspect of ecology for evolution and the coevolution and interaction of the individuals with their environment. Moreover, mostly constant population sizes are assumed but a real biological population should be able to regulate its size dependent on the environment it lives in.

At the beginning of the 90ies, the theory of adaptive dynamics was developed (e.g. [52, 72, 76]). This theory is a stochastic approach and studies the effects of ecological

aspects of populations on evolution and mostly omit the genetics. It highlights the coevolution of the environment with the population due to ecological interaction by a *density-dependent* selection, that models the survival and reproduction ability of one individual in relation to the whole population. Since genetics is omitted the only source of variation are mutations acting directly on the phenotypes. Only phenotypic evolution is considered whereas the genotypic background of the individuals is ignored. The main difference to earlier approaches of population genetics is that the adaptive landscape is replaced by an *invasion fitness landscape* which is not fixed anymore and measures the selective advantage of an appearing mutant depending on the environment.

Starting points for the theory of adaptive dynamics are mostly asexually reproducing populations consisting in the same phenotypic trait. These populations are called *resident, monomorphic population*. Furthermore, it is assumed that the population size is large and mutations are rare [29, 75] resulting in a separation of the evolutionary timescale from the ecological one in such a way that the phase of competitive interaction is small compared to the evolutionary one driven by mutations. This allows to define the *invasion fitness* of an appearing mutant which is defined by the initial growth of the mutant under the environmental conditions set by the resident populations. Invasion fitness is related to the probability of whether the mutant gets extinct or can fixate, i.e. grows from one individual to a notable size, in the resident population [75, 76]. Thus compared to the fixed fitness landscape, the invasion fitness landscape of adaptive dynamics coevolves with the population since it defines the environment and describes the evolution by successive mutation invasions [74,75,100]. Since the trait space is assumed to be continuous these models have the advantage to represent the whole evolutionary process compared to earlier modelling approaches [3]. Thus each new mutation has never appeared before and evolution does not have to stop at the asymptotic evolutionary state [28, 39, 59].

The strength of this approach is that the techniques can be used to understand evolutionary phenomena in various ecological system especially the possibility of phenotypic diversification. This interesting phenomenon of evolutionary branching where a phenotypic trait splits into two coexisting lineages which evolve in different directions was first investigated by Metz et al. in [75] and further studied in [28, 39, 40]. For example, in [75] and [39] it is shown that the occurrence of evolutionary branching depends on the derivative of the fitness function at so called *evolutionary singularities*. However, outside a neighbourhood of such an evolutionary singularity and under the additional assumption that the invasion of a mutant implies the extinction of the resident phenotype the population stays monomorphic over time. This major concept is the *trait substitution sequence (TSS)* modelling evolution as a continuous time Markov process jumping from one monomorphic population to another according to higher fitness. It was introduced by Metz et al. [76] (see also [28, 75]) and mathematically studied in [14, 15, 17, 18]. It can be extended to the polymorphic case, *polymorphic evolution sequence (PSE)*, where the mutant invades a polymorphic population at the expense of one or more resident populations [28, 75]. The *canonical equation of adaptive dynamics (CEAD)*, introduced by Dieckmann and Law [28], is a deterministic approximation of the monomorphic TSS, under the assumption that the difference between the mutant and resident phenotypes is small. It models evolution as a gradual process due

to small phenotypic changes. In other words it is a sequence of successfully established mutations.

Adaptive dynamics and sexual reproduction

The theory of adaptive dynamics deals with the phenotypic long-term evolution of an initial haploid and asexual monomorphic population. The techniques can be applied to many interesting evolutionary issues like virulence, seed size and cyclic evolution, parasite coexistence and predator-prey systems [41, 61, 71, 91].

The phenomenon of evolutionary branching is one of the most important results of these techniques since in biological implications it can be interpreted in the asexual case as morphological speciation and as a driving force for the biodiversity. But as mentioned above most individuals in nature reproduce sexually and, accordingly, the most interesting phenomenon is the *speciation* in sexual populations, where one population splits into two different lineages which are reproductively isolated. Consequently, to get a realistic insight into evolution, including ecological dynamics and genetics, it is necessary to apply the techniques of adaptive dynamics to sexually reproducing populations. If one wants to adapt the concept of evolutionary branching to sexually reproducing populations one is confronted with several problems. The first is to model the reproduction event. It is now not only a cloning anymore but rather depends on mating success and fecundity of two individuals. Moreover, it is subject to the Mendelian rules and mutations change alleles. Consequently, the genotype of an individual is to take into account and a rule for mapping the genotype to the phenotype has to be assumed since selection acts on phenotypes. Recall that in an asexual model, the genotype is identified with the phenotype and is thus omitted in the evolutionary process. This way, mutations alter the phenotype instead of the genotype. One solution to model the long-term evolution is to go back and consider evolution in the allele space rather than in the phenotypic trait space [27, 38, 60, 95]. But even if we get the branching phenomenon in the allele space, when translating back to the phenotype level we have to deal with the presence of the heterozygotes. Thus, for speciation there are additional assumptions needed such as assortative mating (i.e. like individuals mate more preferentially) [27, 38] or spatial segregation, called *allopatric speciation* [60].

The consideration of Mendelian diploid models just started in 1999 with a paper of Kisdi and Geritz [59] followed by a number of works [38, 60, 90, 95–97]. All this papers consider special models with Mendelian reproduction on a heuristic level. Most of them, [38, 60, 95, 97], as also the one of Kisdi and Geritz [59], use a continuous diploid version of Levene's soft selection model ([67], see Figure I.3) because of its relative simplicity and its well-known population genetics. A haploid clonal counterpart of this model was already studied in 1998 by Kisdi and Geritz [39] which makes a comparison of this model with the Mendelian diploid model versions possible.

In [59], the authors consider one locus with a continuum of possible alleles and an environment which consists of two habitats, 1 and 2, each having an optimal phenotype, m_1 and m_2. They assume that the population size is constant in each generation and that the habitats are of relative size $c_1 = c$ and $c_2 = 1 - c$ (see Figure I.3). In these ecological settings they model evolution of the population in allele space under frequency-dependent

selection.

They assume rare mutations with small phenotypic effects, such that there is enough time for a mutant to invade or extinct and for the population to reach its genetic equilibrium before a further mutation occurs. In this way and since a continuum of possible alleles is assumed in contrast to earlier aporaches, long-term evolution is studied as a sequence of mutation invasions.

In the model, all individuals have the same fecundity and discrete, non-overlapping generations. At the start of each generation, individuals are distributed randomly in the two habitats. There is first a phase of viability selection within each habitat, where the survival probability of an individual depends on its phenotype, x, and is given by the Gaussian function,

$$f_i(x) = \alpha_i \exp\left(-\tfrac{(x-m_i)^2}{2\sigma^2}\right), \tag{I.1}$$

for $i \in \{1, 2\}$, with variance σ^2 and α_i the maximal survival probability. Then follows a second phase of nonselective contest competition during which the available living space is allocated at random among the survivors. In this way, a fixed number of individuals is recruited in each habitat (a fraction c in the first habitat und the remaining fraction $1 - c$ in the second habitat, *soft selection*, see Figure I.3). These individuals form a population where mating occurs randomly and offsprings are produced accordingly to Mendelian rules at the end of each generation. It is assumed that alleles, x and y, act additively on the phenotype, precisely, the heterozygote is exactly in between the two homozygotes (the phenotype of an xy individual is given by $\frac{x+y}{2}$). The invasion fitness of a rare mutant allele y in a resident monomorphic population of allele x is given by

$$S_x(y) = c\frac{f_1\left(\frac{x+y}{2}\right)}{f_1(x)} + (1-c)\frac{f_2\left(\frac{x+y}{2}\right)}{f_2(x)}. \tag{I.2}$$

Kisdi and Geritz [59] show that in the diploid model evolutionary branching in a monomorphic population occurs under exactly the same ecological circumstances as in the haploid model but the further evolution is completely different. In the diploid sexual model, resp. haploid clonal model, the invasion of a mutant allele, resp. trait, implies substitution of the resident allele, resp. trait, as long as the population is away from an evolutionary singular point. If the haploid clonal population reaches an evolutionary singular point it converges to an evolutionary stable dimorphism. Thus there is only one evolutionary outcome. On the contrary, in the diploid sexual model there are up to three different possible outcomes depending on the difference between the optimal phenotypes in the habitats and the habitats' sizes:

(i) a single evolutionary stable (which cannot be invaded by any further mutant allele) genetic dimorphism (both homozygotes are habitat specialists),

(ii) two convergence and evolutionary stable genetic dimorphism (at each one the heterozygote and one of the homozygote are habitat specialists),

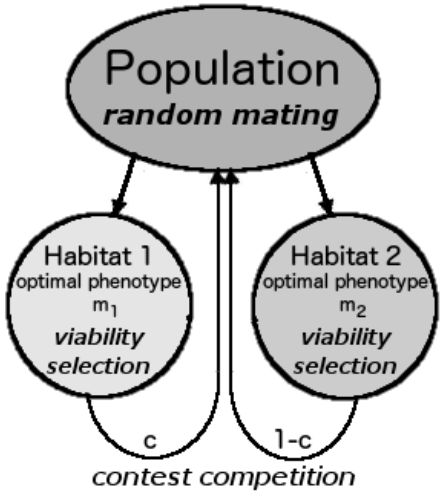

Figure I.3.: Levene's two habitat soft-selection model

(iii) three convergence and evolutionary stable genetic dimorphism (the ones of (i) and (ii)).

In [59], Kisdi and Geritz also discuss the probability for reaching these polymorphisms. However, observe that this branching phenomenon is completely different from the asexual case since we achieve only genetic variability in the allele space which does not leads automatically to speciation. To be more precisely, we end up in a protected genetic polymorphism but do not automatically get phenotypically distinct lineages. In their following work [38], Kisdi and Geritz show that there could be indeed *sympatric speciation* (speciation in the same geographic region). To be more precisely, they consider the polymorphic population after the branching and show that there would be evolution of assortative mating and hence partial reproductive isolation. For this, they extend the single locus model by a second locus with two possible alleles (one dominant over the other) which decodes the mating group the individual belongs to and does not affect the individual fitness. Individuals mate with probability p within its mating group and with $1 - p$ mating is random. In [38], it is discussed that the first locus (which undergoes branching) and the second (controlling mating) have to be in linkage disequilibrium (the nonrandom association of the alleles at the two different loci). The development of such a linkage disequilibrium is only possible if the difference between the two habitats is large enough, such that we get the possible evolutionary outcomes (ii) and (iii) of the one locus model, and selection against heterozygotes is strong. Then there is evolution of assortative mating and there always exists only one evolutionary outcome where the two homozygotes are habitat specialists as in the clonal model. Nevertheless, the evolution of assortative mating is not the only possibility for speciation as van Dooren [95] shows. For a similar version of Levene's soft selection model, as studied by Kisdi and Geritz [38, 59], he demonstrates that the evolution of dominance can lead to elimination of the phenotypic intermediate heterozygote. In [97], van Doorn and Dieckmann study the long-term evolution in a multi-locus version

of Levene's soft selection model under the ecological settings in [59]. By simulations they get that frequency-dependent selection does not maintained genetic polymorphism at a large number of loci. Only one locus with large phenotypic effect retained genetic variation.

2.4. The stochastic individual based model of adaptive dynamics

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., [14–17,28,36] and [3]). To be more precise, in these models finite populations of non-constant size are considered consisting of individuals each owning a birth rate, a natural death rate and an additional death rate due to competition with the other present individuals. Moreover, there is a probability for mutation at each birth event. The population dynamics in these approaches are density-dependent which keeps the population from exploding and allows for variable population size. The aim of these models, first introduced by [8] and [66], is to study the macroscopic phenotypic long-term evolution of a population by tracing it back to the microscopic individual level [14].

In the context of stochastic individual-based models, so far, the biological approach of adaptive dynamics has been put on a rigorous mathematical footing almost exclusively for haploid and asexually reproducing populations. In this framework an important and interesting feature of these models is that the TSS, PES and the CEAD appear as limiting processes on different time-scales as the population size tends to infinity while mutation rates and mutation step-sizes tend to zero. In [14], Champagnat proves convergence to the trait substitution sequence (TSS) in the simultaneous limit of large population and small mutation and in [15], Champagnat, Ferrière and Ben Arous show that this process converges in the limit of small mutation steps to the canonical equation of adaptive dynamics (CEAD). Recently, Baar, Bovier, and Champagnat [2] prove the convergence to the CEAD in the simultaneously combined limits of large population, rare mutations and small mutation steps. The phenomenon of evolutionary branching was rigorously derived by Champagnat and Méléard. In [17], they obtained the convergence to the polymorphic evolution sequence (PES), where jumps occur between equilibria that may include populations that have multiple coexisting phenotypes.

Mendelian diploid model

In the context of individual-based models, the study of diploid sexual reproducing population started recently. The first work in this direction is the one of Collet, Méléard and Metz [18]. Therein they consider a *Mendelian diploid model*, a single locus model of a finite, diploid population with sexual reproduction following the Mendelian rules, under the assumption that alleles act additively on the phenotype, they are *co-dominant*. For this

model they derived, in the limit of large population and rare mutations, the convergence of the suitably time-rescaled process to the TSS model of adaptive dynamics, essentially as shown by Champagnat [14] in the haploid case. This paper was the starting point followed by a series of works by Coron [19,20] and Coron, Méléard, Porcher, and Robert [21] which all consider single locus models. In the first work finite populations are considered. Coron studies the fixation probability of one allele in the absence of mutations in the neutral case (all individuals have the same birth, natural death and competition death rates) and in the non-neutral case (the death rate of one allele slightly deviate from the neutral case). She derived the TSS on the genotype space for the successive fixations of mutations by adding rare mutations and rescaling properly the time. Moreover, if all mutations are deleterious, in [19] it is shown that, after each fixation of a deleterious mutation the natural death rate of the individuals increases. In this way, the existence of an extinction vortex is shown where the increasing frequency of fixation of deleterious mutations is observed and the extinction of the population is inevitable. This mutational meltdown for a model of a finite population similar to the one of [19] is numerically studied in [21]. Therein, the impact of deleterious mutation accumulations on the population size is analysed. Furthermore, the dependence of the strength of the mutational meltdown on demographic parameters is considered. In the last cited work on a Mendelian diploid model [20], Coron proves the convergence to a slow-fast stochastic diffusion dynamic in the large population limit where mutations are accelerated as a consequence of acceleration of the birth and death events and under a fine-tuning of competition. In a second step she shows that this diffusion, conditioned on non-extinction, permits a unique quasi-stationary distribution. Finally, the long-time coexistence of two alleles in three cases (pure neutral competition, over-dominance and separate niches) are numerically studied.

The results so far do not give indications for genetic variation and speciation such that one can suggest an advantage of sexual reproduction in adaptive dynamics (except under some special fine-tuning of parameters). As in the haploid case, also in the diploid case [18], normally the time for fixation of mutant traits in a monomorphic resident population is of the same order as the time for extinction of the residual trait. Thus as soon as a new mutation appears the resident population is already monomorphic again.

One exception in the framework of individual based models is the work of Dieckmann and Doebli [27]. They show that in a multi locus model with only two possible alleles per locus ($+$ and $-$) speciation is possible. They assume that mating probabilities of individuals are expressed in an additional quantitative trait depending on the ecological trait or a marker trait which is ecologically neutral. In the first case they show that evolutionary branching happens under slightly more restrictive conditions as in the corresponding clonal model. However, in the latter case, the development of a linkage disequilibrium between the ecological and the marker trait is needed which can be obtained by genetic drift due to stochastic demographics effects. But in this case the parameter requirements for evolutionary branching are more restrictive than those in the asexual case. However, all these results based only on simulations.

In this thesis, we discuss the dramatic impact on genetic evolution in a Mendelian diploid single locus model if we replace the assumption of co-dominance in [18] by assuming that

the mutant allele is dominant. We will see that the time for extinction of the residual recessive allele is extremely enlarged such that a further more advantageous mutation can appear which paves the way for the appearance of a richer limiting process.

3. The Mendelian diploid model on base

In this section we present the Mendelian diploid model, which is studied in this thesis. It is an extended version of the single locus model introduced in [18] and includes mating groups.

At any time $t \geq 0$, the population under consideration consists of a finite number of individuals N_t on which the three basic mechanisms of evolution, Mendelian heredity, mutation and selection, are acting. The genotype of each individual i is determined by two alleles $u_1^i u_2^i$ at a single locus, taken from some *allele space* $\mathcal{U} \subset \mathbb{R}$. We suppress parental effects, which means that we identify individuals with genotypes u_1u_2 and u_2u_1. Each individual can act as father or mother, i.e. it is hermaphroditic. As explained above the genotype defines the phenotype. In Chapters II and III we make a *complete dominant-recessivity assumption* on the phenotype, i.e. the dominant allele defines the phenotype. Heterozygous individuals thus exhibit the phenotype of the dominant allele which is the same as the phenotype of the individuals which are homozygous for this allele. We construct a Markov process modelling the Mendelian reproduction and the death of each individual without any assumption on the genotype-phenotype mapping. To this end, we introduce the following parameters where we omit the dependence on the phenotype to shorten the notation:

$f_{u_1u_2} \in \mathbb{R}_+$	the per capita birth rate (fertility) of an individual with genotype u_1u_2,
$D_{u_1u_2} \in \mathbb{R}_+$	the per capita natural death rate of an individual with genotype u_1u_2,
$K \in \mathbb{N}$	the carrying capacity, a parameter which scales the population size,
$\frac{c_{u_1u_2,v_1v_2}}{K} \in \mathbb{R}_+$	the competition effect felt by an individual with genotype u_1u_2 from an individual of genotype v_1v_2,
$R_{u_1u_2}(v_1v_2) \in \{0,1\}$	the reproductive compatibility of the genotype v_1v_2 with genotype u_1u_2,
$\mu_K \in \mathbb{R}_+$	the mutation probability per birth event. It is independent of the genotype,
$m(u, dh)$	the mutation law of a mutant allelic trait $u + h \in \mathcal{U}$, born from an individual with allelic trait u.

Individuals are living in an environment which provides them an amount of resources, called the *carrying capacity* of the environment and denoted by the scaling parameter $K \in \mathbb{N}$. Thus the competition is rescaled by $\frac{1}{K}$ which amounts to scaling the population

size to order K. This way the population size is controlled by the amount of resources the environment offers and we count individuals weighted with $\frac{1}{K}$. We are interested in asymptotic results when K is large and mutations are rare.

We denote by $u_1^1(t)u_2^1(t), ..., u_1^{N_t}(t)u_2^{N_t}(t)$ the genotypes present at time t and the whole population, ν_t^K, at time t is represented by the rescaled sum of Dirac measures on $\mathcal{U}^2$,

$$\nu_t^K = \frac{1}{K}\sum_{i=1}^{N_t} \delta_{u_1^i(t)u_2^i(t)}. \tag{I.3}$$

Formally, ν_t^K takes values in the set of re-scaled point measures

$$\mathcal{M}^K = \left\{ \frac{1}{K}\sum_{i=1}^{n} \delta_{u_1^i u_2^i} \;\middle|\; n \geq 0, u_1^1 u_2^1, ..., u_1^n u_2^n \in \mathcal{U}^2 \right\}, \tag{I.4}$$

on $\mathcal{U}^2$, equipped with the vague topology. Define $\langle \nu, g \rangle$ as the integral of the measurable function $g : \mathcal{U}^2 \to \mathbb{R}$ with respect to the measure $\nu \in \mathcal{M}^K$. Then $\langle \nu_t^K, \mathbb{1} \rangle = \frac{N_t}{K} = n_t$ and for any $u_1u_2 \in \mathcal{U}^2$ the positive number $\langle \nu_t, \mathbb{1}_{u_1u_2} \rangle$ is called the *density* at time t of the genotype u_1u_2.

The dynamics of the process are as follows (see [36], [18]): We start at time $t = 0$ with a (possibly random) distribution $\nu_0 \in \mathcal{M}^K$. Each individual with genotype u_1u_2 has three independent exponential clocks:

- a reproduction without mutation $Exp\,(f_{u_1u_2}(1-\mu))$-clock: When it rings then the individual chooses at random an individual v_1v_2 as partner and reproduces with it at rate $f_{u_1u_1}\frac{f_{v_1v_2}R_{u_1u_2}(v_1v_2)}{K\langle \nu R_{u_1u_2}, f\rangle}$. The offspring's genotype is a pair of two alleles, each one chosen randomly of each parent.

- a reproduction with mutation $Exp\,(f_{u_1u_2}\mu)$-clock: When it rings then the individual chooses at random an individual v_1v_2 as partner and reproduces with it at the same rate as before but the offspring gets a genotype where one of the parental alleles changes from u to $u + h$ with h chosen randomly according to the mutation law $m(u, dh)$. Since we assume rare mutations, i.e. $\mu_K \ll 1$, only one parental allele changes.

- a death $Exp(D_{u_1u_2} + \frac{1}{K}\sum_{j=1}^{N_t} c_{u_1u_2,u_1^ju_2^j})$-clock: When it rings then the individual dies. Observe that this parameter depends on the natural death rate and on an additional death rate due to ecological competition with the other individuals present in the population.

Once one of the clocks rings all clocks are reset to zero.

Let us now construct the generator of this process $(\nu_t^K)_{t\geq 0}$. As in [18] we first define,

for the genotypes u_1u_2, v_1v_2 and a point measure ν, the Mendelian reproduction operator

$$(A_{u_1u_2,v_1v_2}F)(\nu) = \frac{1}{4}\left[F\left(\nu+\frac{\delta_{u_1v_1}}{K}\right)+F\left(\nu+\frac{\delta_{u_1v_2}}{K}\right)+F\left(\nu+\frac{\delta_{u_2v_1}}{K}\right)+F\left(\nu+\frac{\delta_{u_2v_2}}{K}\right)\right]-F(\nu), \tag{I.5}$$

and the Mendelian reproduction-cum-mutation operator

$$\begin{aligned}(M_{u_1u_2,v_1v_2}F)(\nu) =& \frac{1}{8}\int_{\mathbb{R}}\Bigg[\left(F\left(\nu+\frac{\delta_{u_1+h,v_1}}{K}\right)+F\left(\nu+\frac{\delta_{u_1+h,v_2}}{K}\right)\right)m(u_1,h)\\ &+\left(F\left(\nu+\frac{\delta_{u_2+h,v_1}}{K}\right)+F\left(\nu+\frac{\delta_{u_2+h,v_2}}{K}\right)\right)m(u_2,h)\\ &+\left(F\left(\nu+\frac{\delta_{u_1,v_1+h}}{K}\right)+F\left(\nu+\frac{\delta_{u_2,v_1+h}}{K}\right)\right)m(v_1,h)\\ &+\left(F\left(\nu+\frac{\delta_{u_1,v_2+h}}{K}\right)+F\left(\nu+\frac{\delta_{u_2,v_2+h}}{K}\right)\right)m(v_2,h)\Bigg]dh-F(\nu).\end{aligned} \tag{I.6}$$

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

$$\begin{aligned}(L^KF)(\nu) =& \int_{\mathcal{U}^2}\left(D_{u_1u_2}+\int_{\mathcal{U}^2}c_{u_1u_2,v_1v_2}\nu(d(v_1v_2))\right)\left(F\left(\nu-\frac{\delta_{u_1u_2}}{K}\right)-F(\nu)\right)K\nu(d(u_1u_2))\\ &+\int_{\mathcal{U}^2}(1-\mu_K)f_{u_1u_2}\left(\int_{\mathcal{U}^2}\frac{f_{v_1v_2}R_{u_1u_2}(v_1v_2)}{\langle\nu R_{u_1u_2},f\rangle}(A_{u_1u_2,v_1v_2}F)(\nu)\nu(d(v_1v_2))\right)K\nu(d(u_1u_2))\\ &+\int_{\mathcal{U}^2}\mu_Kf_{u_1u_2}\left(\int_{\mathcal{U}^2}\frac{f_{v_1v_2}R_{u_1u_2}(v_1v_2)}{\langle\nu R_{u_1u_2},f\rangle}(M_{u_1u_2,v_1v_2}F)(\nu)\nu(d(v_1v_2))\right)K\nu(d(u_1u_2)).\end{aligned} \tag{I.7}$$

The first non-linear term is density-dependent and describes the competition between individuals. It makes selection *frequency-dependent*, i.e. the fitness of an individual depends on the frequencies of the different individuals present in the population. The second and last non-linear terms describe birth with and without mutation. Note that $R_{u_1u_2}(v_1v_2)$ can be interpreted as a mating ability and models whether an individual with genotype u_1u_2 can reproduce with an individual with genotype v_1v_2. Thus $\nu R_{u_1u_2}$ is the population restricted to the pool of potential partners of an individual of genotype u_1u_2.

For fixed K and all $u_1u_2, v_1v_2\in\mathcal{U}^2$, under the following assumptions

- $\mathbb{E}(\langle\nu_0^K,\infty\rangle)<\infty$,

- The functions f, D and c are measurable and bounded, which means that there exists $\bar{f}, \bar{D}, \bar{c} < \infty$ such that

$$0 \leq f_{u_1u_2} \leq \bar{f}, \quad 0 \leq D_{u_1u_2} \leq \bar{D} \quad \text{and} \quad 0 \leq c_{u_1u_2,v_1v_2} \leq \bar{c}, \tag{I.8}$$

- There exists a function, $\bar{m} : \mathbb{R} \to \mathbb{R}_+$, such that $\int \bar{m}(h)dh < \infty$ and $m(u,h) \leq \bar{m}(h)$ for any $u \in \mathcal{U}$ and $h \in \mathbb{R}$,

the existence and uniqueness in law of such a process, in the space $\mathbb{D}(\mathbb{R}_+, \mathcal{M}^K)$ of càdlàg functions from $\mathbb{R}^+$ to $\mathcal{M}_K$, with infinitesimal generator L^K can be derived from Fournier and Méléard [36]. The construction of the process given in [36] as solution of a stochastic differential equation driven by Poisson point measures describing each jump can be adapted to our settings.

We are interested in the large population limit. In this case, under mild restrictive assumptions, the process converges to a deterministic process, which is the solution to a non-linear integro-differential equation. The proof of this can be deduced from Fournier and Méléard [36] as well. We state here only the case where the mutation rate is zero which can be interpreted as the short time evolution of an initial population.

Theorem I.1 (Theorem 3.1 in [18]). *When K tends to infinity in law and if ν_0^K converges in law to a deterministic measure ν_0, then, for any measurable, symmetric function $g : \mathcal{U}^2 \to \mathbb{R}$, the process (ν_t^K) converges in law to the deterministic continuous measure-valued function $(\nu_t)_{t\geq 0}$ solving*

$$\begin{aligned}\langle \nu_t, g\rangle =& \langle \nu_0, g\rangle - \int_0^t \left\langle \nu_s, g(u_1u_2)\left(D_{u_1u_2} + \frac{1}{K}\sum_{j=1}^{N_s} c_{u_1u_2,v_1^jv_2^j}\right)\right\rangle ds \\ &+ \int_0^t \left\langle \nu_s \otimes \nu_s, \frac{f_{u_1u_2}f_{v_1v_2}}{4\langle \nu_s, f\rangle}\left(g(u_1v_1) + g(u_1v_2) + g(u_2v_1) + g(u_2v_2)\right)\right\rangle ds.\end{aligned} \tag{I.9}$$

We take a closer look at the three cases of initial populations which are of greatest interest for this thesis:

(1) the one allele case where $\mathcal{U} = \{A\}$ and hence all individuals are homozygotes with genotype AA,

(2) the two allele case where $\mathcal{U} = \{a, A\}$ and the population consists of the three genotypes aa, aA and AA,

(3) the three allele case where $\mathcal{U} = \{a, A, B\}$ and the population consists of the six genotypes aa, aA, AA, aB, AB and BB.

Already on these three simplest cases the problems of modelling a sexual diploid population become clear. An allele space $\mathcal{U}$ of n alleles provides a *genotype space* $\mathcal{G}$ of

$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ genotypes. One new appearing allele, due to mutation, in the allele space $\mathcal{U}$ blows up the genotype space $\mathcal{G}$ by $n+1$ genotypes through the inevitable occurrence of heterozygotes, whereas in a haploid clonal population it increases only by one. Moreover, the birth rates of the different genotypes are completely different and rather more complicated compared to the clonal case. In a sexual population the birth rate of an individual carrying a certain genotype depends on the whole state of the population or of several subpopulations and not only on the population carrying this genotype as in the clonal case. In the following we make this more precise in the three special cases.

(1) In the first case let us assume that the initial population consists of $n_0^K \delta_{AA}$ individuals with $n_0^K \to n_0$, for $K \to \infty$. By Theorem I.1 the process $(n_t)_{t\geq 0}$ converges in law, when $K \to \infty$, to the solution of the classical logistic equation

$$\dot{n}(t) = n(t)(f_{AA} - D_{AA} - c_{AA,AA} n(t)) \quad \text{and} \quad n(0) = n_0. \tag{I.10}$$

The unique stable fixed point of this equation is the *equilibrium size* of a monomorphic AA population and is equal to the carrying capacity:

$$\bar{n}_{AA} = \frac{f_{AA} - D_{AA}}{c_{AA,AA}}. \tag{I.11}$$

The birth rates $b_{AA}(n(t))$, can be derived by computing the reproduction rates with the Mendelian rules as described in (III.5). In this case it is simply $b_{AA}(n(t)) = f_{AA} n_{AA}(t)$ and comparable to birth rates in the clonal case.

(2)+(3) In the last two cases (compare Proposition 3.2 in [18]) let us assume that the initial condition $n_0^K = (n_{aa}(0), n_{aA}(0), n_{AA}(0))$, resp.
$n_0^K = (n_{aa}(0), n_{aA}(0), n_{AA}(0), n_{aB}(0), n_{AB}(0), n_{BB}(0))$ converges to a deterministic vector (x_0, y_0, z_0), resp. $(u_0, v_0, w_0, x_0, y_0, z_0)$, for $K \to \infty$. Then the process $(n_t)_{t\geq 0}$ converges in law, for $K \to \infty$, to the solution of

$$\dot{n}(t) = b(n(t)) - d(n(t)), \tag{I.12}$$

$$\text{with} \quad \dot{n}_i(t) = b_i(n(t)) - n_i(t) \left(D_i + \sum_{i \in \mathcal{G}} c_{i,j} n_j(t) \right), \quad \forall i \in \mathcal{G}. \tag{I.13}$$

The calculation of the birth rates $b_i(n(t))$, $i \in \mathcal{G}$, in these cases becomes quite complicated. We will illustrate it by computing the birth rate of an aa individual.

In the thesis we consider the case where the fertility is neutral, that means that $f_i = f$, for all $i \in \mathcal{G}$, and that the B allele is the most dominant and the a allele the least dominant one. Consequently, the ascending order of dominance is given by $a < A < B$. Furthermore, we make the dominant-recessivity assumption, that the most dominant allele defines the phenotype and consider the case where the a phenotype (consequently only expressed by individuals of genotype aa) is not

capable of reproducing with individuals of B phenotype (which are the ones with genotypes aB, AB and BB). In mathematical terms,

$$R_i(j) = \begin{cases} 1, & \text{for } i,j \in \{aa, aA, AA\}, \\ 1, & \text{for } i,j \in \{aA, AA, aB, AB, BB\}, \\ 0, & \text{for } i = aa \text{ and } j \in \{aB, AB, BB\}, \\ 0, & \text{for } i \in \{aB, AB, BB\} \text{ and } j = aa. \end{cases} \tag{I.14}$$

In case (2) possible matings which results in an aa individual are (see Figure I.4 (left)):

- $aa \times aa$, with probability 1,
- $aa \times aA$, with probability $\frac{1}{2}$,
- $aA \times aA$, with probability $\frac{1}{4}$.

Since all individuals can reproduce with each other in this case, the whole population acts as pool of possible partners when an individual chooses a partner. In this way the birth rate of an individual of genotype aa is given by:

$$b_{aa}(n(t)) = \frac{f(n_{aa}(t) + \frac{1}{2}n_{aA}(t))^2}{n_{aa}(t) + n_{aA}(t) + n_{AA}(t)}. \tag{I.15}$$

In case (3) we have to bear in mind that aa individuals are not capable of reproducing with individuals of phenotype B. Consequently, the possible matings which result in an aa individual in this case are (see Figure I.4 (right)):

- $aa \times aa$, with probability 1,
- $aa \times aA$, with probability $\frac{1}{2}$,
- $aA \times aA$, with probability $\frac{1}{4}$,
- $aA \times aB$, with probability $\frac{1}{4}$,
- $aB \times aB$, with probability $\frac{1}{4}$.

The pool of possible partners for an aa individual consists, as in the case before, of all individuals with genotypes aa, aA, and AA. In contrast, the whole population acts as pool of possible partners for phenotypic A individuals and for phenotypic B individuals the pool consist only of aA, AA, aB, AB and BB individuals. Therefore, in this case the birth rate of an individual of genotype aa is given by:

$$\begin{aligned} b_{aa}(n(t)) =& f\frac{n_{aa}(t)\left(n_{aa}(t) + \frac{1}{2}n_{aA}(t)\right)}{n_{aa}(t) + n_{aA}(t) + n_{AA}(t)} \\ &+ f\frac{\frac{1}{2}n_{aA}(t)\left(n_{aa}(t) + \frac{1}{2}n_{aA}(t) + \frac{1}{2}n_{aB}(t)\right)}{n_{aa}(t) + n_{aA}(t) + n_{AA}(t) + n_{aB}(t) + n_{AB}(t) + n_{BB}(t)} \\ &+ f\frac{\frac{1}{2}n_{aB}(t)\left(\frac{1}{2}n_{aA}(t) + \frac{1}{2}n_{aB}(t)\right)}{n_{aA}(t) + n_{AA}(t) + n_{aB}(t) + n_{AB}(t) + n_{BB}(t)}. \end{aligned} \tag{I.16}$$

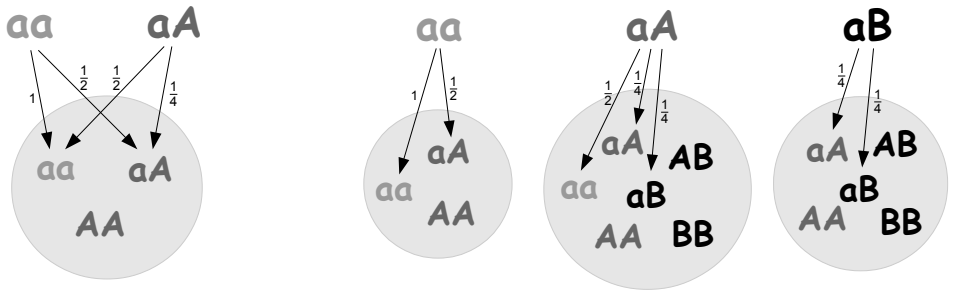

Figure I.4.: Birth rate of an aa individual in Case (2) (left) and Case (3) (right).

One important result (see [18, 36]) is that on finite time intervals the behaviour of a large population can be approximated by the solution of the deterministic system (I.12). Consequently, the knowledge about the deterministic system is a good starting point for understanding the behaviour of the stochastic system. However, this is not as easy as it seems since the result holds only on finite time intervals and, as we will see in Chapter II, we have to control the behaviour of a stochastic process 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 which justifies that the stochastic system behaves like the deterministic system [31, 36]. Moreover, analysing the behaviour of high dimensional systems like (I.12) is a considerable challenge. For the 3-dimensional system this is done by Collet, Méléard and Metz in [18]. In Chapter III a rigorous analytic study of the 6-dimensional deterministic system is provided.

4. Outline of Chapter II and III

Chapters II and III present the main results of this thesis. In Chapter II we study the genetic time evolution of the stochastic individual-based model introduced in Section 3 which describes a diploid hermaphroditic population reproducing according to Mendelian laws. The aim is to show that under a complete dominance-recessivity assumption the recessive allele has a prolonged survival time compared to the previous result in [18] under the co-dominance assumption. Chapter III picks up the result of Chapter II and considers the fate of the recessive allele in the population after a further mutation to a more dominant allele in the large population limit. Although both chapters are related they can be read independently from each other. The following two subsections outline these chapters. Therein we will call an u_1u_2 population a population containing only individuals with genotype u_1u_2 which we name u_1u_2 individuals or more briefly u_1u_2.

4.1. Outline: Survival of a recessive allele in a Mendelian diploid model

In the second chapter the evolution of a recessive allele in a Mendelian diploid model under a complete dominance-recessivity assumption is studied. The results appeared as a joint work with Prof. Dr. Anton Bovier in the *Journal of Mathematical Biology* [84]:

> R. Neukirch and A. Bovier, *Survival of a recessive allele in a Mendelian diploid model*, Journal of Mathematical Biology (2016), pp. 1-54.

Chapter II contains this article mostly as it has been published, only a straightforward improvement of the lower bound on the survival time of the recessive allele to $K^{1/2-\alpha}$ and modifications in formatting are made.

We start with a monomorphic population of aa individuals with one additional mutant with genotype aA. That means our allelic trait space consists of the recessive a allele and the dominant A allele, $\mathcal{U} = \{a, A\}$, which generates a genotype space $\mathcal{G} = \{aa, aA, AA\}$. In the model, an individual chooses a partner uniformly at random for reproduction, thus $R_i(j) = 1$, for all $i, j \in \mathcal{G}$, and all individuals can reproduce with each other. For the mutant to be able to invade the resident aa population, it needs to have a higher fitness, which we obtain by assuming that its natural death rate is slightly reduced, namely for some $\Delta > 0$:

$$D_{aa} = D + \Delta \quad \text{and} \quad D_{aA} = D. \tag{I.17}$$

We assume that all the other parameters, $f_{u_1u_2}$ and $c_{u_1u_2,v_1v_2}$, are neutral which means that they are the same for each phenotype.
In the article [18], Collet, Méléard and Metz studied this Mendelian diploid model under the assumption that the two alleles are co-dominant and that the allele A is slightly fitter than the allele a, namely:

$$D_{aa} = D + \Delta, \quad D_{aA} = D + \tfrac{1}{2}\Delta \quad \text{and} \quad D_{AA} = D. \tag{I.18}$$

In these settings they show that, after the invasion of the mutant AA (which takes time of order $\ln K$), the genotypes containing the a allele, aa and aA, die out exponentially fast in time of order $\ln K$, as already derived in [14] for the haploid asexual model. Therefore, the time scales for the fixation of a new trait and the extinction of the resident trait are the same and the population is monomorphic again before a new mutation occurs.

Thus, as known from the haploid asexual case [14], the suitably time-rescaled process converges to the TSS model of adaptive dynamics in the large population and rare mutation limit. In other words, the population jumps from one homozygote to another homozygote population according to higher fitness and no genetic variability is obtained. Therefore, to get genetic variability it is necessary to ensure, that the recessive allele survives in the population long enough such that a new advantageous mutation can appear before its extinction. For this reason we make the complete dominant-recessivity assumption, precisely we assume that the a allele is recessive and the A allele is dominant. Consequently, individuals with genotype aA and genotype AA have both the same phenotype A, according

to the third Mendelian law and since the parameters, introduced in Section 3, depend on the phenotype we get for the natural death rates:

$$D_{aa} = D + \Delta \quad \text{and} \quad D_{aA} = D_{AA} = D. \tag{I.19}$$

Through the dependence of the parameters on the phenotype, this yields that the AA individuals are as fit as the aA individuals and both are fitter than the aa individuals. This assumption has a dramatic effect on the evolution of the population. The resulting different behaviour under the two assumptions can be traced back to the deterministic system that arises in the large population limit. The linear stability analysis of the unique stable fixed point $\bar{\mathfrak{n}}_{AA}$, corresponding to a monomorphic AA population, yields that it is degenerated i.e. has a zero eigenvalue under the dominant-recessivity assumption. This leads to a different long-term behaviour towards the stable fixed point $\bar{\mathfrak{n}}_{AA}$ compared to the model with the co-dominance assumption. To be more precisely, it implies that in the deterministic system, the aa and aA populations decay in time only polynomially fast to zero, namely like $\frac{1}{t^2}$ and $\frac{1}{t}$, respectively, in contrast to the exponential decay in the co-dominant scenario, corresponding to only strictly negative eigenvalues (see Figure I.5). This type of decay of a recessive allele has been observed earlier in the context of population genetic models (see, e.g., [82], Chapter 4). The main result of Chapter II is that this behaviour of

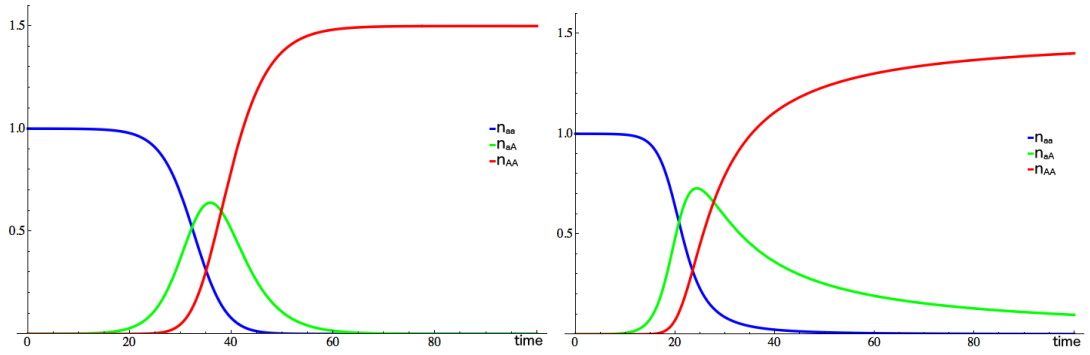

Figure I.5.: Evolution of the model from a resident aa population at equilibrium with a small amount of mutant aA, and when the alleles a and A are co-dominant (left) or when the mutant phenotype A is dominant (right).

the deterministic system translates in the stochastic model into survival of the genotypes carrying the recessive allele for a time of order almost $K^{1/2-\alpha}$, for $\alpha > 0$. This allows for a reasonable scaling of the mutation rate μ_K such that a new mutant will occur with high probability in the AA population before the a allele is extinct. To be more precisely, we scale the mutation rate by

$$\ln(K) \ll \frac{1}{K\mu_K} \ll K^{1/2-\alpha}, \quad \text{as } K \to \infty. \tag{I.20}$$

This scaling can be motivated as follows: The mutation probability for an individual with genotype u_1u_2 is given by μ_K. Hence, the time until the next mutation in the whole population is of order $\frac{1}{K\mu_K}$. Since the time a mutant population needs to invade a resident

population is of order $\ln(K)$ (which can be adapted from [14, 18]) we set the mutation time $\frac{1}{K\mu_K} \gg \ln(K)$ to ensure that the mutant invades before a new mutation appears. The right hand side of (I.20), $K^{1/2-\alpha}$, is the time until which the recessive a allele survives in the population, shown in Chapter II, and corresponds to the decay time of a function of $\frac{1}{t}$. The exponent can be explained as follows: since the population carrying the recessive a allele can get directly extinct when it decreases to the size of order of the natural fluctuation, we can ensure its survival only until a slightly higher size, namely $K^{-1/2+\alpha}$. Notice, that the extinction of the a allele depends on the aA population which decays although it is as fit as the AA population but it is disadvantaged in reproduction. The aa population behaves like the square of the aA population and thus can already die out, due to natural fluctuations, but will always be reproduced by the aA population.

Under this scaling it is shown in Chapter II that the survival time of the a population, consisting of all aa- and aA individuals, is of order $K^{1/2-\alpha}$ and that there will be a further mutation before the a allele get extinct. The main step is to show that the behaviour of the deterministic system translate into the stochastic model. The difficulty thereby is, compared to earlier works [14,18], to control the behaviour of the stochastic system, since now time diverges like a power of K after the invasion of the AA population. This precludes to adapt the techniques used in [14] and [18], 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, Bovier and Champagnat [2]. This scheme combines coupling methods with discrete time Markov chains and standard potential theoretic approaches for the exit from an attractive domain and uses results from the theory of branching processes. Since we consider a diploid sexual model instead of a haploid asexual one as in [2] we have to handle three subpopulations rather than only two which all influences each other. Thus to apply the Euler scheme to the present model a main task is to find the right order to control the subpopulations over small steps such that we can show that the stochastic system follows the deterministic system. The scaling of the mutation rate (I.20) allows the fitter AA population to invade the resident population but ensures that there will be a further mutation before the a population is extinct. Consequently, unfit alleles can survive in heterozygotes and there could appear a new mutant allele, call it B, which has strong competition with the AA population but weak competition with the aa population and can coexists with the recessive aa population. In this way a resurgence of the aa population at the expense of the AA population and coexistence of the types aa and BB may be observed. This would increase the genetic variability of the population and could be a first step in the direction of speciation. These suggestions are the motivation for the following work of Chapter III.

4.2. Outline: The recovery of a recessive allele in a Mendelian diploid model

In the third chapter we study the further evolution of the stochastic individual based model of Chapter II in the large population limit. We show that after environmental changes the aa population can recover and that coexistence of homozygous genotypes is possible. The results are a joint work with Dr. Loren Coquille from the Institute Fourier (University of

Grenoble Alpes) and Prof. Dr. Anton Bovier and are available as research paper on the online-portal arXiv:

> A. Bovier, L. Coquille and R. Neukirch, *The recovery of a recessive allele in a Mendelian diploid model*, arXiv e-print 1703.02459, March 2017

Chapter III contains this article mostly as it has been published subject to minor modifications in formatting. The starting point is the scenario of Chapter II after the invasion of the mutant AA. To understand the behaviour of the stochastic system, it is important to study first the deterministic system, corresponding to the large population limit of the stochastic counterpart, which should be a good approximation if the a population is not killed by the typical fluctuations (see Figure I.6). To be more concrete, we start with the

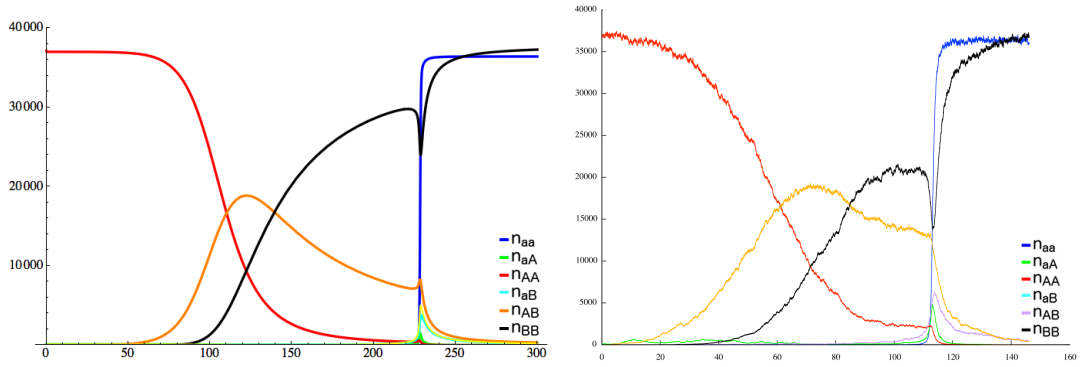

Figure I.6.: time vs. individuals plot: deterministic model (left) stochastic model (right)

following initial conditions for the deterministic system which correspond to the situation of the stochastic model after invasion of AA: the AA population is large and close to its equilibrium (see (I.11)), the aA population is of size ϵ and consequently the aa population is of order ϵ^2. Moreover we assume that there occurs a further mutation to a most dominant B allele such that individuals exhibiting the B phenotype are the fittest in the population and we start with AB individuals of size ε^3. Notice, that after this next mutation, we have six possible genotypes, aa, aA, AA, aB, AB and BB present in the population, three of phenotype B (aB, AB and BB), two of phenotype A (aA and AA) and only one of phenotype a (aa). Therefore, we have to deal with a six-dimensional deterministic system. The high complexity of this system, composed of six interacting subpopulations, makes the analysis intricate. To be able to identify fixed points of the system, we assume that there is no reproduction and no competition between individuals of phenotype a and phenotype B. Observe that this assumption lets aa individuals and B individuals belong to different species since they are reproductively isolated. In this way, we can show analytically the existence of a fixed point, p_{aB}, where the two subpopulations aa and BB can coexist. However, the linear stability analysis fails since the system has at p_{aB} two zero eigenvalues. Nevertheless, using the Center Manifold Theorem (asserting that the qualitative behaviour of the dynamical system in a neighbourhood of the non-hyperbolic critical point p_{aB} is determined by its behaviour on the center manifold near p_{aB}, [44,51]), we can

indeed prove the stability of this fixed point and the convergence of the system towards it. This can be translated in an evolution of biodiversity. Notice, that in the stochastic system the population bearing the a allele can die out due to the natural fluctuations if it is too small. Therefore, we prove that also in the deterministic model the a population stays above a certain level which corresponds to the order of natural fluctuation in the stochastic model. Another possibility to avoid the extinction of the a allele is to ensure that the aa population only grows after the mutation. For that reason we introduce an additional parameter η which lowers the competition of BB with aA compared to the one of BB with AA. This way, competition does not depend only on phenotypes anymore. Instead it can be interpreted as a refinement of a phenotypic competition for resources with genotypic influences: the strength (or ability to get resources) of an individual not only depends on its phenotype but also on the degree of dominance of its genotype. To be more precise, the AA individuals have one more dominant allele compared to the aA individuals, namely a second A allele instead of the a allele. Consequently, they compete more strongly for resources with the BB individuals than the aA individuals do. This lack in competition accelerates the evolution of the system since η allows aA to decrease more slowly in such a way that the AA and aA population get faster to the same order which is important for the aa population to get exponential growth and thus to recover. In particular, the birth rate of an aa individual (see (I.16)) has a term

$$f n_{aa} \frac{(n_{aa} + \frac{1}{2} n_{aA})}{n_{aa} + n_{aA} + n_{AA}}, \tag{I.21}$$

describing the reproduction of aa individuals with aa- and aA individuals out of the pool of potential partners, which yields the exponential growth as soon as aA and AA are on the same order.

An interesting feature of adding the system controlling parameter η is that there arises a bifurcation phenomenon for η larger than some threshold. In particular, for these values of η the coexistence fixed point p_{aB} becomes unstable and the system converges to another fixed point where all six subpopulations coexist. We discussed this by numerical simulations (see Figure I.7).

The reason for this different limiting behaviour lies in the fact that if η is chosen large enough the competition felt by aA from BB is not strong enough to force its extinction. Therefore, a fraction of aA individuals survives in the population and through Mendelian reproduction all other subpopulations are reproduced. To sum up, the main result of Chapter III is:

Theorem I.2. *Consider the dynamical system* (I.12) *started with the initial conditions mentioned above. Suppose the following assumptions on the parameters hold:*

(C1) Δ *sufficiently small,*

(C2) f *sufficiently large,*

(C3) $0 \leq \eta < c/2$.

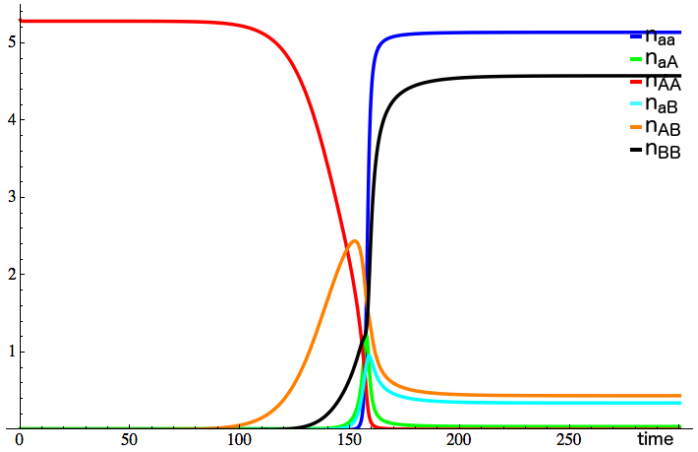

Figure I.7.: 6-population fixed point: $\eta = 0.6$

Then the system converges to the fixed point p_{aB}. More precisely, for any fixed $\delta > 0$, as $\epsilon \to 0$, it reaches a δ-neighbourhood of p_{aB} in a time of order $\Theta(\epsilon^{-1/(1+\eta\bar{n}_B-\Delta)})$. Moreover, it holds:

1. *for $\eta = 0$, the amount of allele a in the population decays to $\Theta(\epsilon^{1+\Delta/(1+\Delta)})$ before reaching $\Theta(1)$,*
2. *for $\eta > \frac{4\Delta}{\bar{n}_B}$, the amount of allele a in the population is bounded below by $\Theta(\epsilon)$ for all $t > 0$,*

where $x = \Theta(y)$ whenever $x = \mathcal{O}(y)$ and $y = \mathcal{O}(x)$ as $\epsilon \to 0$. The proof of this theorem is divided into four phases.

In the first phase we show that the mutant B population grows up to a level ϵ_0 exponentially fast, with a rate corresponding to the invasion fitness of an AB individual in a resident AA population, without perturbing the 3-system (aa, aA, AA).

The second phase ends when the aA population and the AA population are of the same order. By a comparison result we show that in this phase the effective 3-system (AA, AB, BB) is almost unperturbed and behaves like the 3-system (aa, aA, AA) analysed in Chapter II. We show that in this time BB approaches its equilibrium $\bar{n}_{BB}$.

In Phase 3 we ensure the exponential growth of the aa population until an ϵ_0-neighbourhood of its equilibrium $\bar{n}_{aa}$. Notice that the growth rate of aa, which corresponds to the invasion fitness, $S_{aa,BB}$, of an aa individual in a resident BB population is given by

$$S_{aa,BB} = f - D - \Delta, \tag{I.22}$$

is much larger than the one of the B population in a resident AA population in Phase 1, given by

$$S_{AB,AA} = f - D + \Delta - c\bar{n}_{AA} = \Delta. \tag{I.23}$$

This follows from Assumption **(C2)** and that there is no competition between aa and BB individuals.

Using the Center Manifold Theorem [44, 51] we prove in the last phase that the stable fixed point is approached with speed $\frac{1}{t}$.

Finally we relax the assumption on the parameters and discuss the different limits of the resulting models by numerical simulations. A rigorous analytic analysis is not possible here since we are already unable to calculate all the fixed points.

The main requirement for the recovery of the aa population is that it has a positive invasion fitness in a resident BB population. Consequently, we can relax the no competition assumption between aa and the B phenotype and add small competition between them, denoted c_{aB}. Therefore, also under this assumptions, we can end up with two or six coexisting populations depending on the parameters choices (see Figure I.8).

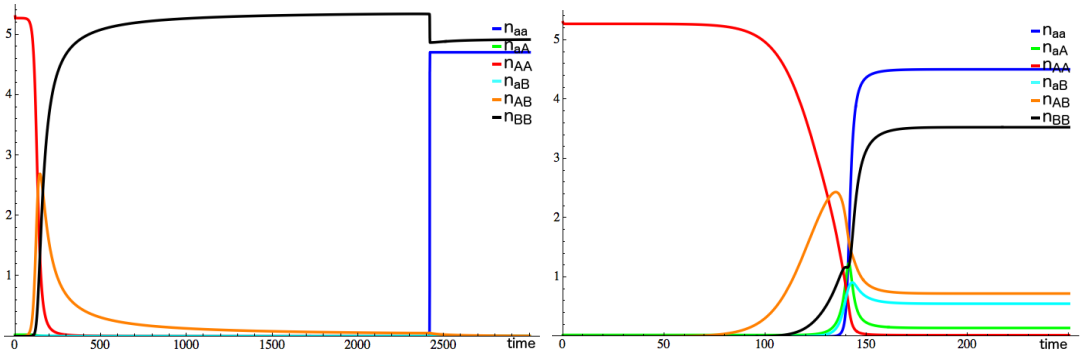

Figure I.8.: Model without reproduction between a and B: $c_{aB} = 0.1$ and $\eta = 0$ (left), $c_{aB} = 0.1$ and $\eta = 0.6$ (right)

If we instead remove the no reproduction assumption, that means we set $R_i(j) \equiv 1$, for all $i, j \in \mathcal{G}$, such that aa also reproduces with phenotype B, the 2- population fixed point, p_{aB}, cannot be observed anymore. The reason for this is that as soon as the aa population could recover the coexisting BB- and aa population will instantly give birth to aB individuals. Consequently, we observe a 3-population fixed point (see Figure I.9). The BB population cannot increase to its monomorphic equilibrium, $\bar{n}_{BB}$, anymore due to competition with aB. Adding competition c_{aB} and the factor η yields similar results as before, this time with three or six coexisting populations (see Figure I.9). Based on reproduction among all phenotypes this model is named *all-with-all model*. Notice, compared to the previous models in this model the fecundity has to be much bigger to observe the recovery of the aa population since all individuals act as possible partners and thus the birth rate of aa individuals is smaller.

Finally, let us mention that if we set in both models (in the model without reproduction between a and B and the all-with-all model) the competition between each phenotype neutral, that means $c_{aB} = c$, and add a certain value η, we naturally end up in a 6-population fixed point, since as argued above the reduced competition ensures the survival of a fraction of aA individuals such that all subpopulations will be produced.

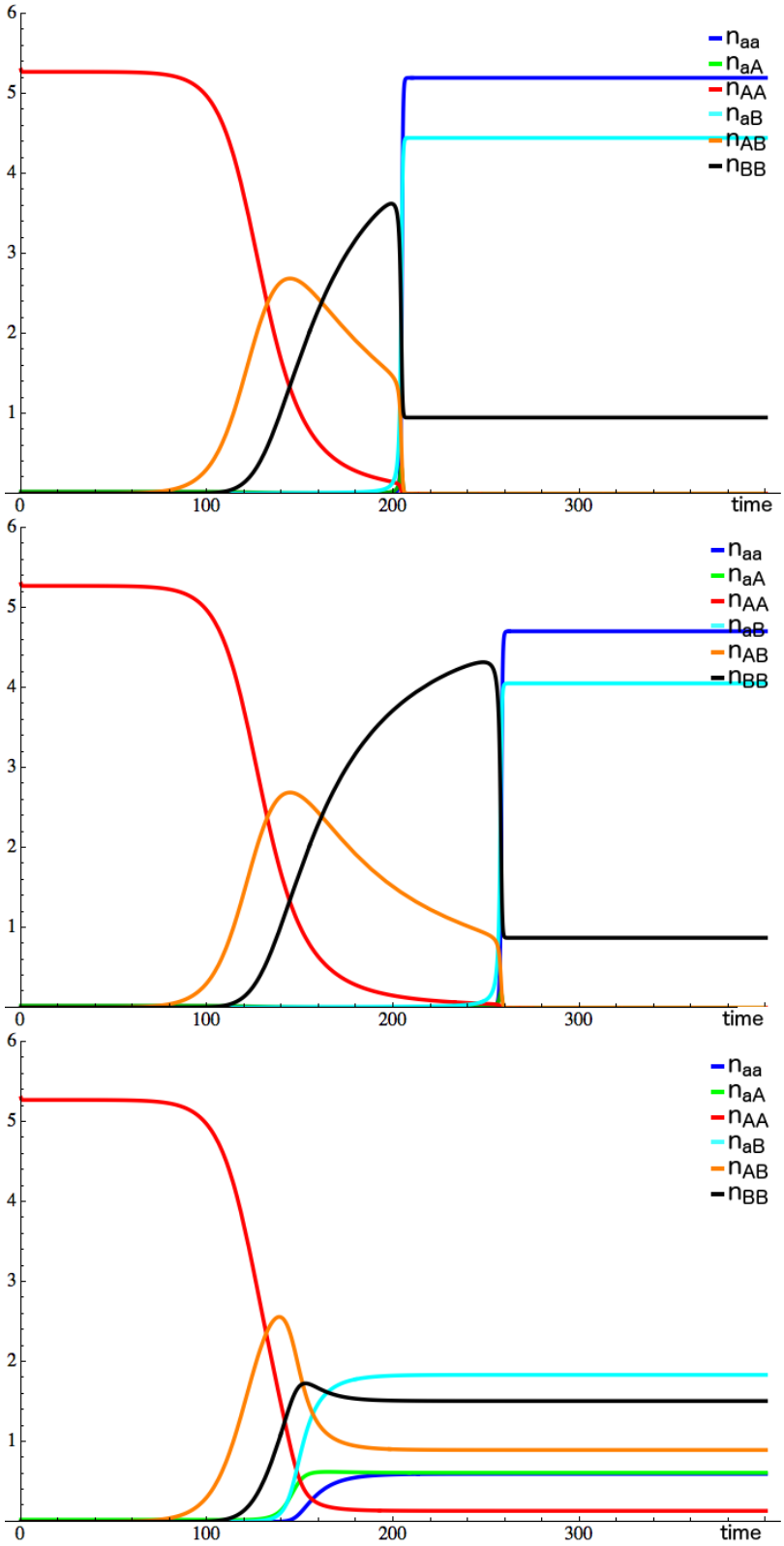

Figure I.9.: All-with-all model: $c_{aB} = 0$ and $\eta = 0$ (top), $c_{aB} = 0.1$ and $\eta = 0$ (center), $c_{aB} = 0.9$ and $\eta = 0.4$ (bottom)

II. Chapter

Survival of a recessive allele in a Mendelian diploid model

Anton Bovier and Rebecca Neukirch

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, Méléard and Metz [18] 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 $K^{1/2-\alpha}$, where K is the size of the population and $\alpha > 0$.

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.

1. Introduction

Mendelian diploid models have been studied for over a century in the context of *population genetics* (see, e.g., [107], [34], [103], [45], and [46]). Text book expositions of population genetics are given in, e.g., [24], [82], [32], and [12]. While population genetics typically deals with models of fixed population size, adaptive dynamics, a variant that has been developed in the 90ies (e.g., [52], [72], and [76]), 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 [59].

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., [28], [15], [14], [36], [16], and [17]). 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 [14] proves convergence to the *Trait Substitution Sequence* (TSS) (see, e.g., [28], and [75]) and to the *Canonical Equation of Adaptive Dynamics* (CEAD). Champagnat and Méléard [17] also rigorously derive the phenomenon of *evolutionary branching* under the assumption of coexistence. In a recent paper [2] 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, Méléard and Metz [18] 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, Méléard, Porcher, and Robert [21] and Coron [19, 20]. 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, Méléard and Metz [18], 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 offsprings 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, Méléard and Metz [18] 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 [14] in the haploid case. We now reverse the assumption made by Collet, Méléard and Metz [18] 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, Méléard and Metz [18]. This type of behaviour has been observed earlier in the context of population genetic models (see, e.g., [82], 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 a allele for a time of order almost $K^{1/2}$, 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, Bovier and Champagnat [2]. 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., [32]), 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 [7].

2. Model setup and goals

2.1. Introduction of the model

We consider a Mendelian diploid model introduced by Collet, Méléard and Metz [18]. It models a population of a finite number of individuals with sexual reproduction, where each individual i is characterised by two alleles, $u_1^i u_2^i$, from some allele space $\mathcal{U} \in \mathbb{R}$. These two alleles define the genotype of individual i, which in turn defines its phenotype, $\phi(u_1^i u_2^i)$, through a function $\phi : \mathcal{U}^2 \to \mathbb{R}$. We suppress parental effects, thus $\phi(u_1^i u_2^i) = \phi(u_2^i u_1^i)$. 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,
$\frac{c_{u_1u_2,v_1v_2}}{K} \in \mathbb{R}_+$	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 $u_1^1 u_2^1, ..., u_1^{N_t} u_2^{N_t}$ the genotypes of the population at time t. The population, ν_t, at time t is represented by the rescaled sum of Dirac measures on $\mathcal{U}^2$,

$$\nu_t = \frac{1}{K} \sum_{i=1}^{N_t} \delta_{u_1^i u_2^i}. \tag{II.1}$$

ν_t takes values in

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

where $\mathcal{M}$ denotes the set of finite, nonnegative measures on $\mathcal{U}^2$ equipped with the vague topology. Define $\langle \nu, g \rangle$ as the integral of the measurable function $g : \mathcal{U}^2 \to \mathbb{R}$ with respect to the measure $\nu \in \mathcal{M}^K$. Then $\langle \nu_t, \mathbb{1} \rangle = \frac{N_t}{K}$ and for any $u_1 u_2 \in \mathcal{U}^2$, the positive number $\langle \nu_t, \mathbb{1}_{u_1 u_2} \rangle$ is called the *density* at time t of the genotype $u_1 u_2$. The generator of the process is defined as by Collet, Méléard and Metz [18]: First we define, for the genotypes $u_1 u_2, v_1 v_2$ and a point measure ν, the Mendelian reproduction operator:

$$\begin{aligned} &(A_{u_1 u_2, v_1 v_2} F)(\nu) \\ &= \frac{1}{4} \left[F\left(\nu + \frac{\delta_{u_1 v_1}}{K}\right) + F\left(\nu + \frac{\delta_{u_1 v_2}}{K}\right) + F\left(\nu + \frac{\delta_{u_2 v_1}}{K}\right) + F\left(\nu + \frac{\delta_{u_2 v_2}}{K}\right) \right] - F(\nu), \end{aligned} \tag{II.3}$$

and the Mendelian reproduction-cum-mutation operator:

$$\begin{aligned} (M_{u_1 u_2, v_1 v_2} F)(\nu) =& \frac{1}{8} \int_{\mathbb{R}} \Bigg[\left(F\left(\nu + \frac{\delta_{u_1 + h v_1}}{K}\right) + F\left(\nu + \frac{\delta_{u_1 + h v_2}}{K}\right) \right) m_\sigma(u_1, h) \\ &+ \left(F\left(\nu + \frac{\delta_{u_2 + h v_1}}{K}\right) + F\left(\nu + \frac{\delta_{u_2 + h v_2}}{K}\right) \right) m_\sigma(u_2, h) \\ &+ \left(F\left(\nu + \frac{\delta_{u_1 v_1 + h}}{K}\right) + F\left(\nu + \frac{\delta_{u_2 v_1 + h}}{K}\right) \right) m_\sigma(v_1, h) \\ &+ \left(F\left(\nu + \frac{\delta_{u_1 v_2 + h}}{K}\right) + F\left(\nu + \frac{\delta_{u_2 v_2 + h}}{K}\right) \right) m_\sigma(v_2, h) \Bigg] dh - F(\nu). \end{aligned} \tag{II.4}$$

The process $(\nu_t)_{t \geq 0}$ 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 $\nu \in \mathcal{M}^K$ by:

$$\begin{aligned} &(L^K F)(\nu) \\ &= \int_{\mathcal{U}^2} \left(D_{u_1 u_2} + \int_{\mathcal{U}^2} c_{u_1 u_2, v_1 v_2} \nu(d(v_1 v_2)) \right) \left(F\left(\nu - \frac{\delta_{u_1 u_2}}{K}\right) - F(\nu) \right) K \nu(d(u_1 u_2)) \\ &+ \int_{\mathcal{U}^2} (1 - \mu_K) f_{u_1 u_2} \left(\int_{\mathcal{U}^2} \frac{f_{v_1 v_2}}{\langle \nu, f \rangle} (A_{u_1 u_2, v_1 v_2} F)(\nu) \nu(d(v_1 v_2)) \right) K \nu(d(u_1 u_2)) \\ &+ \int_{\mathcal{U}^2} \mu_K f_{u_1 u_2} \left(\int_{\mathcal{U}^2} \frac{f_{v_1 v_2}}{\langle \nu, f \rangle} (M_{u_1 u_2, v_1 v_2} F)(\nu) \nu(d(v_1 v_2)) \right) K \nu(d(u_1 u_2)). \end{aligned} \tag{II.5}$$

The first non-linear term describes the competition between individuals. The second and last non-linear terms describe the birth without and with mutation. There, $f_{u_1 u_2} \frac{f_{v_1 v_2}}{K \langle \nu, f \rangle}$ is the reproduction rate of an individual with genotype $u_1 u_2$ with an individual with genotype $v_1 v_2$. 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 $u_1 u_2, v_1 v_2 \in \mathcal{U}^2$, 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

$$0 \leq f_{u_1u_2} \leq \bar{f}, \quad 0 \leq D_{u_1u_2} \leq \bar{D} \quad \text{and} \quad 0 \leq c_{u_1u_2,v_1v_2} \leq \bar{c}. \tag{II.6}$$

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

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

For fixed K, under the Assumptions **(A1) + (A3)** and assuming that $\mathbb{E}(\langle \nu_0, \mathbb{1} \rangle) < \infty$, Fournier and Méléard [36] 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 $\mathbb{D}(\mathbb{R}_+, \mathcal{M}^K)$ of the process ν^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 μ_K. Hence, the time until the next mutation in the whole population is of order $\frac{1}{K\mu_K}$. Since the time a mutant population needs to invade a resident population is of order $\ln K$ (see, e.g., [14]), we set the mutation rate $\frac{1}{K\mu_K} \gg \ln(K)$ 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 $\frac{1}{KK^{1/2-\alpha}}$, $\alpha > 0$, then a mutation will occur while the resident phenotypic a population is small, but still alive, in contrast to the setting of Collet, Méléard and Metz [18], 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 II.1 (A allele fittest and dominant) and Figure II.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, Méléard and Metz [18] 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^β. We show that this is indeed the case, with $\beta = 1/2 - \alpha$. 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 $n_{uv}(t) \equiv \frac{1}{K} N_{uv}(t)$.

Definition II.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*

$$\bar{n}_u = \frac{f_{uu} - D_{uu}}{c_{uu,uu}}. \tag{II.7}$$

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

$$S_{uv,uu} = f_{uv} - D_{uv} - c_{uv,uu}\bar{n}_u, \tag{II.8}$$

is called the invasion fitness *of a mutant uv in a resident uu population, where $\bar{n}_u$ is given by (III.11).*

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) $c_{u_1u_2,v_1v_2} =: c, \quad \forall u_1u_2, v_1v_2 \in \{aa, aA, AA\}$.

Remark II.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)**,

$$\begin{aligned} S_{aA,aa} &= S_{AA,aa} = f - D - c\bar{n}_{aa} = f - D - c\frac{f - D - \Delta}{c} = \Delta, \\ S_{aa,aA} &= S_{aa,AA} = f - D - \Delta - c\bar{n}_{AA} = f - D - \Delta - c\frac{f - D}{c} = -\Delta. \end{aligned} \tag{II.9}$$

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 1, the second and third one with probability $\frac{1}{2}$ and the last one with probability $\frac{1}{4}$. Therefore we have the following birth rates:

$$\begin{aligned} b_{aa}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) &= \frac{f\left(N_{aa}(t) + \frac{1}{2}N_{aA}(t)\right)^2}{N_{aa}(t) + N_{aA}(t) + N_{AA}(t)}, \\ b_{aA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) &= \frac{2f\left(N_{aa}(t) + \frac{1}{2}N_{aA}(t)\right)\left(N_{AA}(t) + \frac{1}{2}N_{aA}(t)\right)}{N_{aa}(t) + N_{aA}(t) + N_{AA}(t)}, \\ b_{AA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) &= \frac{f\left(N_{AA}(t) + \frac{1}{2}N_{aA}(t)\right)^2}{N_{aa}(t) + N_{aA}(t) + N_{AA}(t)}. \end{aligned} \tag{II.10}$$

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

$$\begin{aligned} d_{aa}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) &= N_{aa}(t)\left(D + \Delta + c(n_{aa}(t) + n_{aA}(t) + n_{AA}(t))\right), \\ d_{aA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) &= N_{aA}(t)\left(D + c(n_{aa}(t) + n_{aA}(t) + n_{AA}(t))\right), \\ d_{AA}(N_{aa}(t), N_{aA}(t), N_{AA}(t)) &= N_{AA}(t)\left(D + c(n_{aa}(t) + n_{aA}(t) + n_{AA}(t))\right). \end{aligned} \tag{II.11}$$

In the sequel, the *sum process*, $\Sigma(t)$, defined by

$$\Sigma(t) = n_{aa}(t) + n_{aA}(t) + n_{AA}(t), \tag{II.12}$$

plays an important role. A simple calculation shows that the sum process jumps up to $1/K$ (resp. down) with rate b_Σ (resp. d_Σ) given by

$$b_\Sigma(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = fK\Sigma(t), \tag{II.13}$$

$$d_\Sigma(N_{aa}(t), N_{aA}(t), N_{AA}(t)) = (D + c\Sigma(t))K\Sigma(t) + \Delta N_{aa}(t). \tag{II.14}$$

3. Main theorems

In the sequel we denote by τ_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 μ_K:

$$\textbf{(C)} \qquad \ln(K) \ll \frac{1}{K\mu_K} \ll K^{1/2-\alpha}. \tag{II.15}$$

We recall one result from Collet, Méléard and Metz [18] (Proposition D.2) which carries over to our setting: it is shown that if the resident population $n_{aa}(t)$ is in a δ-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 δ. The proof of this result is based on large deviation estimates (see [37]).

Proposition II.1 (Proposition D.2 in [18]). *Let* $\mathrm{supp}(\nu_0^K) = \{aa\}$ *and let* τ_1 *denote the first mutation time. For any sufficiently small* $\delta > 0$*, if* $\langle \nu_0^K, \mathbb{1}_{aa}\rangle$ *belongs to the* $\delta/2$*-neighbourhood of* $\bar{n}_a$ *then the time of exit of* $\langle \nu_t^K, \mathbb{1}_{aa}\rangle$ *from the* δ*-neighbourhood of* $\bar{n}_a$ *is bigger than* $\mathrm{e}^{VK} \wedge \tau_1$*, 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*,*

$$D + cK\langle \nu_t^K, \mathbb{1}_{aa}\rangle, \tag{II.16}$$

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

$$\tau_\delta^{mut} \equiv \inf\{t \geq 0 : 2n_{AA}(t) + n_{aA}(t) \geq \delta\}, \tag{II.17}$$

$$\tau_0^{mut} \equiv \inf\{t \geq 0 : 2n_{AA}(t) + n_{aA}(t) = 0\}. \tag{II.18}$$

Theorem II.1 (Proposition D.4 in [18]). *Let (z_K) be a sequence of integers such that $\frac{z_K}{K}$ converges to $\bar{n}_a$, for $K \to \infty$. Then*

$$\lim_{\delta\to 0}\lim_{K\to\infty} \mathbb{P}_{\frac{z_K}{K}\delta_{aa}+\frac{1}{K}\delta_{aA}}(\tau_\delta^{mut} < \tau_0^{mut}) = \frac{\Delta}{f}, \tag{II.19}$$

where we recall that Δ is the invasion fitness of a mutant aA in a resident aa population (cf. III.17).

We now state the main results of this paper:

Theorem II.2. *Consider the model verifying Assumptions **A** and **B**. Let $\delta > \varepsilon$, $\alpha > 0$ and*

$$\tau_\eta^{hit} \equiv \inf\{t \geq \tau_\delta^{mut} : n_{aA}(t) \leq \eta\}. \tag{II.20}$$

Define $\tau_{sur} \equiv \tau^{hit}_{K^{-1/2+\alpha}} - \tau_\varepsilon^{hit}$. Then, conditional on survival of the mutant, i.e., on the event $\{\tau_\delta^{mut} < \tau_0^{mut}\}$, with probability converging to one as $K \uparrow \infty$, the following statements hold:

(i) $\tau_\varepsilon^{hit} = \mathcal{O}(\ln K)$, *and*

(ii) $\tau_{sur} = \mathcal{O}(K^{1/2-\alpha})$.

Remark II.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 $\frac{1}{2} - \alpha$ 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 II.3. *Assume that Assumption **C** is satisfied. Then, with probability converging to one,*

$$\tau_\varepsilon^{hit} \wedge \tau_1 = \tau_\varepsilon^{hit} \quad \text{and} \quad \tau_1 \wedge \tau_0^{hit} = \tau_1. \tag{II.21}$$

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 (Theorem II.1, II.2 and II.3) have several parts and are quite technical, we first give an outline of them before we turn to the details (Section 5.2).

3.1. Outline of the proofs

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 [18]. More precisely, they have dissimilar long term behaviour (cf. Figure II.1 and II.2).
By analysing the systems one gets that both have the same fixed points $\mathfrak{n}_{aa} \equiv (\bar{n}_a, 0, 0)$ and $\mathfrak{n}_{AA} \equiv (0, 0, \bar{n}_A)$. The computation of the eigenvalues of the Jacobian matrix at the fixed point $\mathfrak{n}_{aa}$ yields in both models two negatives and one positive eigenvalues. Hence in both systems the fixed point $\mathfrak{n}_{aa}$ is unstable. In contrast, the eigenvalues at the fixed point $\mathfrak{n}_{AA}$ are all negative in the system of Collet, Méléard and Metz [18] whereas in this model there is one zero eigenvalue. This leads to the different long term behaviour towards the stable fixed point $\mathfrak{n}_{AA}$. In Collet, Méléard and Metz's [18] 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.

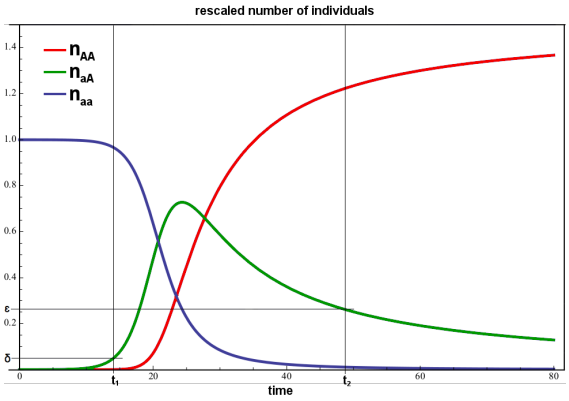

Figure II.1.: Our Model: A fittest and dominant

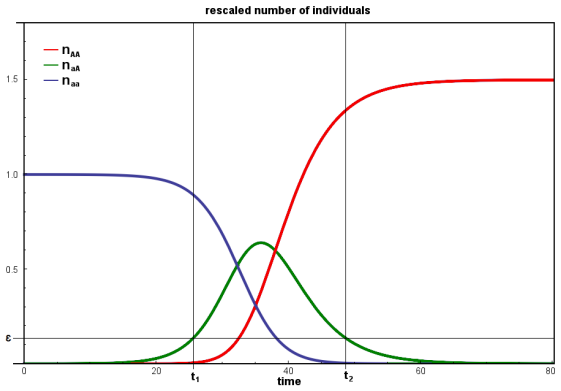

Figure II.2.: Collet et al. Model: a and A co-dominant

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 [14], [18], or [2] (cf. Figure II.3) (the general idea that an invasion can be divided in these phases is much older, see, e.g., [92]):

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, Méléard and Metz [18], whereas the last phase will be analysed in eight steps. Technically, the analysis uses tools developed by Baar, Bovier and

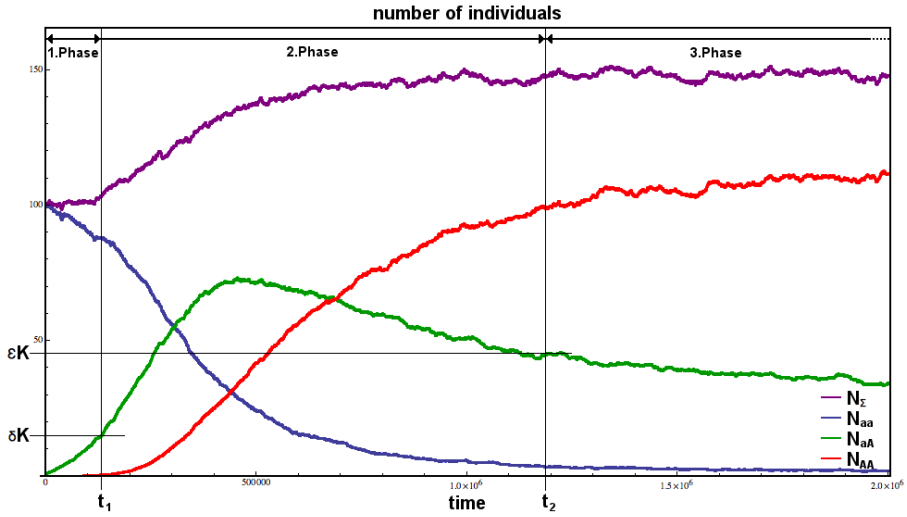

Figure II.3.: The three phases of the proof

Champagnat [2] and classical potential methods (see, e.g., [11]).

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 δ, the resident aa population stays close to its equilibrium $\bar{n}_a$ (Proposition II.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 δ-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 δ. This is the content of Theorem II.1.

Phase 2: Invasion of the mutant population. The fixation (Phase 1) ends with a macroscopic mutant population of size δ. In the second phase the mutant population invades the resident population and suppresses it. By the Large Population Approximation (Theorem II.11, [36]) the behaviour of the process is now close to the solution of the deterministic system (II.43) 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 Section 4 that any solution starting in a δ-neighbourhood of $(\bar{n}_a, 0, 0)$ converges to an ε-neighbourhood of $(0, 0, \bar{n}_A)$ in finite time (t_2 in Figure II.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 ε-level to ensure that the duration of this phase is still finite.

In [14], it is shown that the duration of the first phase is of order $\mathcal{O}(\ln K)$ and that the time for the second phase is bounded. Thus the time needed by the aA population to reach again the ε-level after the fixation is of order $\mathcal{O}(\ln K)$ (cf. Theorem II.2 (i)). From Proposition II.1 we get that the resident aa population stays in a δ-neighbourhood of its equilibrium $\bar{n}_a$ an exponentially long time $\exp(VK)$ as long as the mutant population is smaller than δ. 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 [14] 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 II.1 (Lemma 2 (a) in [14]). *Assume that the initial condition of ν_t satisfies $\sup_K \mathbb{E}(\langle \nu_0, \mathbb{1} \rangle) < \infty$. Then, for any $\eta > 0$, there exists an $\varepsilon > 0$ such that, for any $t > 0$,*

$$\limsup_{K\to\infty} \mathbb{P}^K_{\nu_0}\left(\exists n \geq 0 : \frac{t}{K\mu_K} \leq \tau_n \leq \frac{t+\varepsilon}{K\mu_K}\right) < \eta, \tag{II.22}$$

where τ_n are the ordered sequence of times when a mutation occurs, defined in the beginning of Section 3.

Using Lemma II.1, we get that, for fixed $\eta > 0$, there exists a constant, $\eta > \rho > 0$, such that, for sufficiently large K,

$$\mathbb{P}_{\frac{z_K}{K}\delta_{aa} + \frac{1}{K}\delta_{aA}}\left(\tau_1 < \frac{\rho}{K\mu_K}\right) < \delta, \tag{II.23}$$

where τ_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

$$\ln(K) \ll \frac{1}{K\mu_K}, \tag{II.24}$$

(cf. left inequality of (II.15)) 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 ε-level. This phase is different from the one in [14] and [18], since the analysis of the deterministic system in Section 4 reveals that $n_{aA}(t)$ decreases only like a function $f(t) = \frac{1}{t}$, in contrast to the exponential decay in [18]. 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, Bovier and Champagnat [2], where it was necessary to show that the stochastic system remains close to a deterministic one over times of order K^α 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 [2] to show that the behaviour of the deterministic and the stochastic systems remain close for a time of order $K^{1/2-\alpha}$.

Let us put this scheme on a mathematically footing:

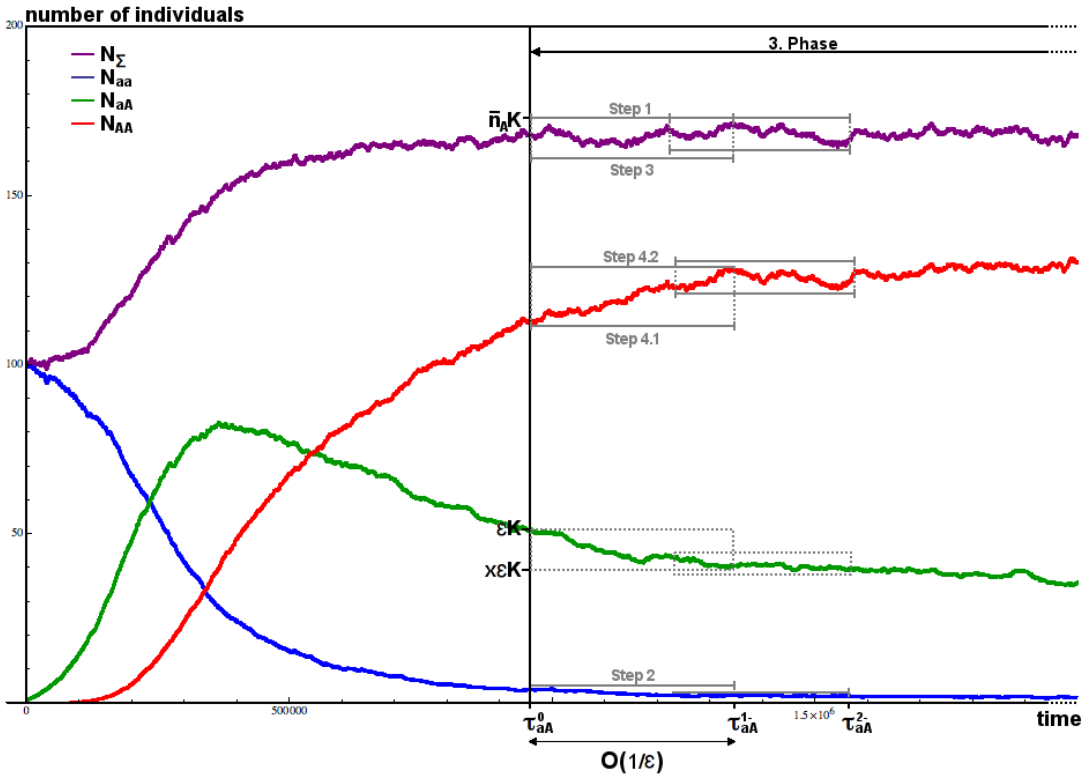

Figure II.4.: Steps of the proof

Settings for the steps. Let

$$\gamma \equiv \frac{f}{4\bar{n}_A(f+\Delta)}, \quad \gamma_{\Delta/2} \equiv \frac{f+\frac{\Delta}{2}(1-\vartheta)}{4\bar{n}_A(f+\Delta)} \quad \text{and} \quad \gamma_\Delta \equiv \frac{f+\frac{\Delta}{2}}{4\bar{n}_A(f+\Delta)}. \tag{II.25}$$

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 $\varepsilon > 0$ and $\vartheta > 0$ such that $\varepsilon < \frac{\Delta}{2} < \vartheta < \Delta$. We set

$$x = \left(\frac{f+\vartheta}{f+\Delta}\right)^{\frac{1}{2}}, \tag{II.26}$$

and, for $0 \leq i \leq \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1 =: \sigma - 1$, with $\alpha > 0$, we define the stopping times on $n_{aA}(t)$ by

$$\tau_{aA}^{i-} \equiv \begin{cases} \tau_{\varepsilon}^{hit}, & \text{for } i = 0, \\ \inf\left\{ t \geq \tau_{aA}^{(i-1)-} : n_{aA}(t) \leq x^i \epsilon \right\}, & \text{else}, \end{cases} \tag{II.27}$$

$$\tau_{aA}^{i+} \equiv \inf\left\{ t \geq \tau_{aA}^{i-} : n_{aA}(t) \geq x^i \epsilon + K^{-1/2+3/4\alpha} \right\}. \tag{II.28}$$

During the time intervals $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-}\right]$, $n_{aA}(t) \in \left[x^{i+1}\varepsilon, x^i\varepsilon + K^{-1/2+3/4\alpha}\right]$. The upper bound on i is chosen in such a way that

$$x^{i+1}\varepsilon \geq K^{-1/2+\alpha}. \tag{II.29}$$

The following eight steps will be iterated from $i = 0$ to $i = \sigma - 1 = \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$.

Remark II.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 $K^{-1/2+\alpha}$. Afterwards it 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 its death.*

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

Proposition II.2. *For all $M > 0$ and $0 \leq i \leq \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$, let*

$$\tau_{\Sigma,M}^{\alpha} \equiv \inf\left\{ t > \tau_{aA}^{i-} : \Sigma(t) - \bar{n}_A \geq MK^{-1/2+\alpha/2} \right\}. \tag{II.30}$$

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

$$\mathbb{P}\left[\tau_{\Sigma,M_\Sigma}^{\alpha} < \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}} \right] = o(K^{-1}). \tag{II.31}$$

To prove Proposition II.2, we define the difference process between $\Sigma(t)K$ and $\bar{n}_A K$ and couple it with a birth-death 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 $M_\Sigma K^{1/2+\alpha/2}$ before going to zero, is very small. Then we show that the process returns many times to zero until it reaches the level $M_\Sigma K^{1/2+\alpha/2}$ and calculate the time for one such return.

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

Proposition II.3. *For all $M > 0$ and $0 \le i \le \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$, let*

$$\tau^{\alpha}_{aa,M} \equiv \inf\left\{ t > \tau^{i-}_{aA} : n_{aa}(t) - \gamma_\Delta x^{2i}\varepsilon^2 \ge M(x^{2i}\varepsilon^2)^{1+\alpha} \right\}. \tag{II.32}$$

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

$$\mathbb{P}\left[\tau^{\alpha}_{aa,M_{aa}} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}} \right] = o(K^{-1}). \tag{II.33}$$

The proof is similar to the one of Proposition II.2.

Step 3: Lower bound on $\Sigma(t)$. With Proposition II.2 and II.3, we can bound $\Sigma(t)$ from below. We show that on the time interval $t \in \left[\tau^{i-}_{aA}, \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}} \right]$ the sum process does not drop below $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}$, with high probability:

Proposition II.4. *For all $M > 0$ and $0 \le i \le \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$, let*

$$\tau^{\alpha}_{\Sigma,M} \equiv \inf\left\{ t > \tau^{i-}_{aA} : \Sigma(t) - \left(\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 \right) \le -MK^{-1/2+\alpha/2} \right\}. \tag{II.34}$$

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

$$\mathbb{P}\left[\tau^{\alpha}_{\Sigma,M_\Sigma} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}} \right] = o(K^{-1}). \tag{II.35}$$

The proof is similar to those of Proposition II.2 and II.3.

Step 4: Lower and upper bound on $n_{AA}(t)$. Since we now have bounded the processes $n_{aa}(t), n_{aA}(t)$ 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 $t \in \left[\tau^{i-}_{aA}, \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}} \right]$. Precisely, there exists a constant, $M_{AA} > 0$, such that with high probability $n_{AA}(t)$ does not drop below $\bar{n}_A - x^i\varepsilon - M_{AA}(x^{2i}\varepsilon^2 + K^{-1/2+3/4\alpha})$ (Proposition II.5), and does not exceed the level $\bar{n}_A - x^{(i+1)}\varepsilon + M_{AA}K^{-1/2+\alpha/2}$ (Proposition II.6):

Proposition II.5. *For all $M > 0$ and $0 \le i \le \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$, let*

$$\tau^{2i}_{AA,M} \equiv \inf\left\{ t > \tau^{i-}_{aA} : n_{AA}(t) - \left(\bar{n}_A - x^i\varepsilon \right) \le -M(x^{2i}\varepsilon^2 + K^{-1/2+3/4\alpha}) \right\}. \tag{II.36}$$

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

$$\mathbb{P}\left[\tau^{2i}_{AA,M_{AA}} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}} \right] = o(K^{-1}). \tag{II.37}$$

Proposition II.6. *For all $M > 0$ and $0 \le i \le \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$, let*

$$\tau^{\alpha}_{AA,M} \equiv \inf\left\{ t > \tau^{i-}_{aA} : n_{AA}(t) - \left(\bar{n}_A - x^{i+1}\varepsilon \right) \ge MK^{-1/2+\alpha/2} \right\}. \tag{II.38}$$

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

$$\mathbb{P}\left[\tau^{\alpha}_{AA,M_{AA}} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}} \right] = o(K^{-1}). \tag{II.39}$$

Step 5: Decay of $n_{aA}(t)$. We now have upper and lower bounds for all the single processes, for $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$. 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 $x^i\varepsilon$ (i.e. we set $\tau_{aA}^{i-} = 0$), hits the level $x^{i+1}\varepsilon$ before it reaches the level $x^i\varepsilon + K^{-1/2+3/4\alpha}$.

Proposition II.7. *There exists a constant $C > 0$ such that, for all $0 \leq i \leq \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$*

$$\mathbb{P}\left[\tau_{aA}^{i+} < \tau_{aA}^{(i+1)-} \middle| n_{aA}(0) = x^i\varepsilon\right] \leq K^{1/2-3/4\alpha} \exp\left(-CK^{7/4\alpha}\right). \tag{II.40}$$

For the proof we couple $n_{aA}(t)$ with majorising and minorising birth-death 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 $x^{i+1}\varepsilon$ before going back to $x^i\varepsilon + K^{-1/2+3/4\alpha}$.

Step 6: Decay time 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 $x^i\varepsilon$ to $x^{i+1}\varepsilon$ is of order $\frac{1}{x^i\varepsilon}$:

Proposition II.8. *Let*

$$\theta_i(aA) \equiv \inf\left\{t \geq 0 : n_{aA}(t) \leq x^{i+1}\varepsilon \middle| n_{aA}(0) = x^i\varepsilon\right\}, \tag{II.41}$$

be the decay time of $n_{aA}(t)$ on the time interval $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$. Then there exist positive constants, C_l, C_u, and a constant $M > 0$, such that for all $0 \leq i \leq \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor - 1$

$$\mathbb{P}\left[\frac{C_u}{x^i\varepsilon} \geq \theta_i(aA) \geq \frac{C_l}{x^i\varepsilon}\right] \geq 1 - \exp\left(-MK^{2\alpha}\right). \tag{II.42}$$

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 $x^{i+1}\varepsilon$, 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. Precisely, we show that $n_{aa}(t)$ decreases from $\gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}$ to $\gamma_{\Delta/2}x^{2i+2}\varepsilon^2$, and stays smaller than $\gamma_\Delta x^{2i+2}\varepsilon^2$ when $n_{aA}(t)$ reaches $x^{i+1}\varepsilon$:

Proposition II.9. *For $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$, the process n_{aA} decreases from $x^i\varepsilon$ to $x^{i+1}\varepsilon$ and $n_{aa}(t)$ decreases from $\gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}$ below $\gamma_\Delta x^{2i+2}\varepsilon^2$.*

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 $\gamma_{\Delta/2}x^{2i+2}\varepsilon^2$ before going back to $\gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}$. 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 $\gamma_\Delta x^{2i+2}\varepsilon^2$ until $n_{aA}(t)$ hits the level $x^{i+1}\varepsilon$ 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 $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}$ to $\bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ on a given time interval and is greater than $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ when the aA population reaches the level $x^{i+1}\varepsilon$. Observe that this step is only needed for $i \leq \left\lfloor \frac{-\ln(\varepsilon K^{1/4-\alpha/4})}{\ln(x)} \right\rfloor$, afterwards $x^{2i}\varepsilon^2 \leq K^{-1/2+\alpha/2}$ and the bound $\bar{n}_A - M_\Sigma K^{-1/2+\alpha/2}$ suffices for the iteration.

Proposition II.10. *For $0 \leq i \leq \left\lfloor \frac{-\ln(\varepsilon K^{1/4-\alpha/4})}{\ln(x)} \right\rfloor$, while n_{aA} decreases from $x^i\varepsilon$ to $x^{i+1}\varepsilon$, the sum process $\Sigma(t)$ increases from $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\, x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}$ to $\bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ and stays above $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ until the aA population hits the $x^{i+1}\varepsilon$-level.*

The proof uses the same three parts as described in the proof of Proposition II.9.

Final Step: Total decay time of $n_{aA}(t)$. We iterate Step 1 to 8 until $i = \left\lfloor \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} \right\rfloor$, the value for which $n_{aA}(t)$ is of order $K^{-1/2+\alpha}$. Finally, we sum up the decay time of the aA population in each iteration step and get the desired result (Theorem II.2 (ii)).
Moreover, we ensure that the upper bound on the mutation probability μ_K in Assumption **(C)** (II.15) 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 [36] or [18] 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 II.11 (Proposition 3.2 in [18]). *Let $T > 0$ and $C \subset \mathbb{R}^3_+$ be compact. Assume that the initial condition $\frac{1}{K}(N^0_{aa}, N^0_{aA}, N^0_{AA})$ 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*

$$\dot{\phi}(t;(x_0,y_0,z_0)) = \begin{pmatrix} \tilde{b}_{aa}(x(t),y(t),z(t)) - \tilde{d}_{aa}(x(t),y(t),z(t)) \\ \tilde{b}_{aA}(x(t),y(t),z(t)) - \tilde{d}_{aA}(x(t),y(t),z(t)) \\ \tilde{b}_{AA}(x(t),y(t),z(t)) - \tilde{d}_{AA}(x(t),y(t),z(t)) \end{pmatrix} =: X(x(t),y(t),z(t)), \tag{II.43}$$

where

$$\begin{aligned} \tilde{b}_{aa}(x(t),y(t),z(t)) &= \frac{(f_{aa}x(t) + \frac{1}{2}f_{aA}y(t))^2}{(f_{aa}x(t) + f_{aA}y(t) + f_{AA}z(t))}, \\ \tilde{d}_{aa}(x(t),y(t),z(t)) &= x(t)(D_{aa} + c_{aa,aa}x(t) + c_{aa,aA}y(t) + c_{aa,AA}z(t)), \end{aligned}$$

and similar expression for the aA and AA types. Then, for all $T > 0$,

$$\lim_{K\to\infty} \sup_{t\in[0,T]} |n_{uv}(t) - \phi_{uv}(t;(x_0,y_0,z_0))| = 0 \quad , a.s., \tag{II.44}$$

for all $uv \in \{aa, aA, AA\}$.

Thus, to understand the behaviour of the process we have to analyse the deterministic system (II.43) above. The vector field (II.43) of the model we consider is given by

$$X(x,y,z) = X_\Delta(x,y,z) = \begin{pmatrix} f\frac{(x+\frac{1}{2}y)^2}{x+y+z} - (D+\Delta+c(x+y+z))x \\ 2f\frac{(x+\frac{1}{2}y)(z+\frac{1}{2}y)}{x+y+z} - (D+c(x+y+z))y \\ f\frac{(z+\frac{1}{2}y)^2}{x+y+z} - (D+c(x+y+z))z \end{pmatrix}, \tag{II.45}$$

which has some particular properties:

Theorem II.4. *Assume **(A)**+**(B)** and let $\varepsilon > 0$, then*

(i) *the vector field* (II.45) *has the unstable fixed point $\mathfrak{n}_{aa} \equiv (\bar{n}_a, 0, 0)$ and the stable fixed point $\mathfrak{n}_{AA} \equiv (0, 0, \bar{n}_A)$,*

(ii) *the Jacobian matrix at the unstable fixed point $\mathfrak{n}_{aa}$, $DX_\Delta(\mathfrak{n}_{aa})$, has two negative and one positive eigenvalues,*

(iii) *the Jacobian matrix at the stable fixed point $\mathfrak{n}_{AA}$, $DX_\Delta(\mathfrak{n}_{AA})$, has two negative and one zero eigenvalues,*

(iv) *for $0 < \varrho < f\Delta$, and as soon as the aA population decreased to an ε-level, then*

$$\frac{2\bar{n}_A(f+\Delta)}{(f\Delta+\varrho)t + \frac{2\bar{n}_A(f+\Delta)}{\varepsilon}} \leq n_{aA}(t) \leq \frac{2\bar{n}_A(f+\Delta)}{(f\Delta-\varrho)t + \frac{2\bar{n}_A(f+\Delta)}{\varepsilon}}. \tag{II.46}$$

There is also a biological explanation for the behaviour of $n_{aA}(t)$ described in Theorem II.4 (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 II.4 (i)+(ii) also holds in the model of Collet, Méléard and Metz [18] (cf. Proposition 3.3 therein) but the Jacobian matrix of their fixed point $\mathfrak{n}_{AA}$ 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, Méléard and Metz [18]:

Theorem II.5 (Theorem C.2 in [18])**.** *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 $\mathfrak{n}_{aa}$ is unstable and we have fixation for the macroscopic dynamics.*
More precisely, there exists an invariant stable curve Γ_ζ which joins $\mathfrak{n}_{aa}$ to $\mathfrak{n}_{AA}$. Moreover there exists an invariant tubular neighbourhood $\mathcal{V}$ of Γ_ζ such that the orbit of any initial condition in $\mathcal{V}$ converges to $\mathfrak{n}_{AA}$.
If $\zeta \frac{dS_{aA,aa}}{d\zeta}(0) < 0$ the fixed point $\mathfrak{n}_{aa}$ 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 (II.43) in the case of neutrality between the alleles A and a. That is $f_{u_1u_2} = f$, $D_{u_1u_2} = D$ and $C_{u_1u_2,v_1v_2} = c$, for $u_1u_2, v_1v_2 \in \{aa, aA, AA\}$. They get that this system has a line of fixed points Γ_0 which is transversally hyperbolic. Afterwards they consider the system X_ζ with small

perturbations ζ. From Theorem 4.1 in Hirsch, Pugh and Shub [51] they deduce that there exists an attractive and invariant curve Γ_ζ, converging to Γ_0, as $\zeta \to 0$. Hence, there is a small enough tubular neighbourhood $\mathcal{V}$ of Γ_0 such that Γ_ζ is contained in $\mathcal{V}$ and attracts all orbits with initial conditions in $\mathcal{V}$ (cf. Figure II.5). To show that the orbit of any initial condition on Γ_ζ converges to one of the two fixed points $\mathfrak{n}_{aa}$ or $\mathfrak{n}_{AA}$, one have to ensure that the vector field does not vanish on Γ_ζ, except for these two fixed points. Since the curve Γ_ζ 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, Méléard and Metz [18] use linear combinations of the left eigenvectors of $DX_0(\Gamma_0(v))^t$, with the perturbed vector field X_ζ. 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 [18]). 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_ζ has only two zeros in a tabular neighbourhood of Γ_0 (cf. Theorem B.4 in [18]).
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 (II.77) and (II.78)) we can indeed prove that this point is an isolated zero and we can deduce from Theorem II.5 the following corollary, which is the main result we need about the dynamical system.

Corollary II.1. *Let $\Delta \neq 0$ small enough.*

(i) *The attracting and invariant curve Γ_Δ of the perturbed vector field X_Δ (II.43) contained in the positive quadrant, is the piece of unstable manifold between the equilibrium points $\mathfrak{n}_{aa}$ and $\mathfrak{n}_{AA}$.*

(ii) *There exists an invariant tubular neighbourhood $\mathcal{V}$ of Γ_Δ such that the orbit of any initial condition in $\mathcal{V}$ converges to the equilibrium point $\mathfrak{n}_{AA}$.*

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

5. Proofs of Theorem II.4 and the main theorems

5.1. Analysis of the deterministic system

Because of Proposition II.11 we have to analyse the deterministic system (II.43) (a simulation is shown in Figure II.1).

Proof of Theorem II.4. In the following we consider the differential equations of $n_{aa}(t)$,

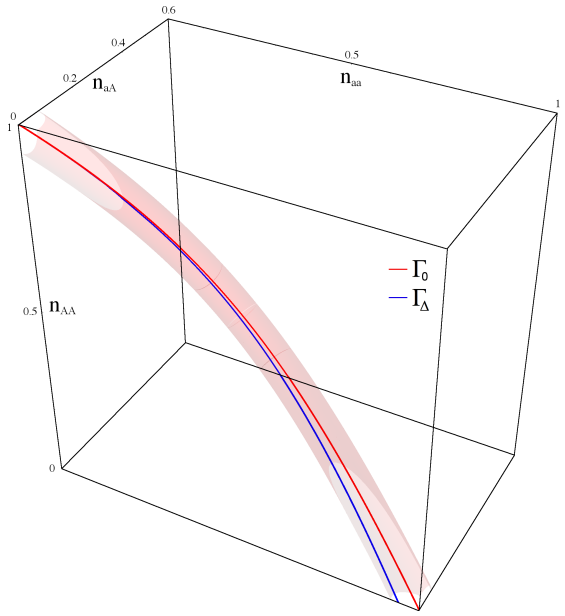

Figure II.5.: The curves Γ_0, Γ_Δ and the tube $\mathcal{V}$ in the perturbed vector field X_Δ

$n_{aA}(t)$ and $n_{AA}(t)$, given by (II.45):

$$\dot{n}_{aa}(t) = f\frac{\left(n_{aa}(t) + \frac{1}{2}n_{aA}(t)\right)^2}{\Sigma(t)} - n_{aa}(t)(D + \Delta + c\Sigma(t)), \tag{II.47}$$

$$\dot{n}_{aA}(t) = 2f\frac{\left(n_{aa}(t) + \frac{1}{2}n_{aA}(t)\right)\left(n_{AA}(t) + \frac{1}{2}n_{aA}(t)\right)}{\Sigma(t)} - n_{aA}(t)(D + c\Sigma(t)), \tag{II.48}$$

$$\dot{n}_{AA}(t) = f\frac{\left(n_{AA}(t) + \frac{1}{2}n_{aA}(t)\right)^2}{\Sigma(t)} - n_{AA}(t)(D + c\Sigma(t)). \tag{II.49}$$

The Fixed Points: By summing (II.47) to (II.49) first, and two times (II.49) and (II.48), we see that the vector field (II.45) vanishes for the points $\mathfrak{n}_{aa}$ and $\mathfrak{n}_{AA}$. The Jacobian matrix at the fixed point $\mathfrak{n}_{aa}$ is given by

$$DX_\Delta((\bar{n}_a, 0, 0)) = \begin{pmatrix} -f + D + \Delta & -f + D + \Delta & -2f + D + \Delta \\ 0 & \Delta & 2f \\ 0 & 0 & -f + \Delta \end{pmatrix}. \tag{II.50}$$

The matrix has the three eigenvalues $\lambda_1 = -(f - D - \Delta)$, $\lambda_2 = \Delta$ and $\lambda_3 = -(f - \Delta)$. For Δ small enough and from Assumption **(A2)** we know that $\lambda_1, \lambda_3 < 0$, whereas $\lambda_2 > 0$. Thus the fixed point $\mathfrak{n}_{aa}$ is unstable.
The Jacobian matrix at the fixed point $\mathfrak{n}_{AA}$ is given by

$$DX_\Delta((0, 0, \bar{n}_A)) = \begin{pmatrix} -f - \Delta & 0 & 0 \\ 2f & 0 & 0 \\ -2f + D & -f + D & -f + D \end{pmatrix}, \tag{II.51}$$

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 [18]. Because of the zero eigenvalue, $\mathfrak{n}_{AA}$ 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 [51,88] that asserts that the qualitative behaviour of the dynamical system in a neighbourhood of the non-hyperbolic critical point $\mathfrak{n}_{AA}$ is determined by its behaviour on the center manifold near $\mathfrak{n}_{AA}$.

Theorem II.6 (The Local Center Manifold Theorem 2.12.1 in [88]). *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*

$$\begin{aligned} \dot{x} &= Cx + F(x, y) \\ \dot{y} &= Py + G(x, y), \end{aligned} \tag{II.52}$$

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 δ-neighbourhood of 0, that defines the local center manifold and satisfies:

$$Dh(x)[Cx + F(x, h(x))] - Ph(x) - G(x, h(x)) = 0, \tag{II.53}$$

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

$$\dot{x} = Cx + F(x, h(x)), \tag{II.54}$$

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

The fact that the center manifold of our system near $\mathfrak{n}_{AA}$ 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 $\mathfrak{n}_{AA}$ is indeed a stable fixed point and that the flow on the center manifold near the critical point behaves like a function $\frac{1}{t}$. This can be seen as follows: Assume that $n_{aA}(t)$ has decreased to a level ε. 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

$$Y(z, y, x) = \begin{pmatrix} f\frac{(z+\bar{n}_A+\frac{1}{2}y)^2}{z+y+x+\bar{n}_A} - (D + c(z + y + x + \bar{n}_A))(z + \bar{n}_A) \\ 2f\frac{(z+\bar{n}_A+\frac{1}{2}y)(x+\frac{1}{2}y)}{z+y+x+\bar{n}_A} - (D + c(z + y + x + \bar{n}_A))y \\ f\frac{(x+\frac{1}{2}y)^2}{z+y+x+\bar{n}_A} - (D + \Delta + c(z + y + x + \bar{n}_A))x \end{pmatrix}, \tag{II.55}$$

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

$$DY((0, 0, 0)) = \begin{pmatrix} -(f - D) & -(f - D) & -(2f - D) \\ 0 & 0 & 2f \\ 0 & 0 & -(f + \Delta) \end{pmatrix}, \tag{II.56}$$

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

$$EV_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad EV_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad EV_3 = \begin{pmatrix} \frac{fD+\Delta(2f-D)}{(f+\Delta)(D+\Delta)} \\ -\frac{2f}{f+\Delta} \\ 1 \end{pmatrix}. \tag{II.57}$$

We perform a new change of variable to work in the basis of eigenvectors of $Y(z,y,x)$

$$T = \begin{pmatrix} 1 & 1 & \frac{D}{D+\Delta} \\ 0 & 1 & \frac{2f}{f+\Delta} \\ 0 & 0 & 1 \end{pmatrix}, \tag{II.58}$$

The change of variables

$$\begin{pmatrix} \tilde{z} \\ \tilde{y} \\ \tilde{x} \end{pmatrix} = T \begin{pmatrix} z \\ y \\ x \end{pmatrix}, \tag{II.59}$$

casts the system into the form

$$\dot{\tilde{z}} = \dot{z} + \dot{y} + \frac{D}{D+\Delta}\dot{x}, \qquad \dot{\tilde{y}} = \dot{y} + \frac{2f}{f+\Delta}\dot{x}, \qquad \dot{\tilde{x}} = \dot{x}. \tag{II.60}$$

Let $h(\tilde{y})$ be the local center manifold. We approximate $h(\tilde{y}) = \binom{h_1}{h_2}\tilde{y}^2 + \mathcal{O}(\tilde{y}^3)$ and substitute the series expansions into the center manifold equation (III.304)

$$\begin{pmatrix} 2h_1\tilde{y} \\ 2h_2\tilde{y} \end{pmatrix} \dot{\tilde{y}}(h_1\tilde{y}^2, \tilde{y}, h_2\tilde{y}^2) = \begin{pmatrix} \dot{\tilde{z}}(h_1\tilde{y}^2, \tilde{y}, h_2\tilde{y}^2) \\ \dot{\tilde{x}}(h_1\tilde{y}^2, \tilde{y}, h_2\tilde{y}^2) \end{pmatrix}. \tag{II.61}$$

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. 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 $-h_2(f+\Delta) + \frac{f}{4\bar{n}_A}$, equals zero and we get $h_2 = \frac{f}{4\bar{n}_A(f+\Delta)}$. 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 $Dh_1 - \frac{f\Delta(2D+\Delta)}{4\bar{n}_A(f+\Delta)(D+\Delta)}$, we get $h_1 = \frac{f\Delta(2D+\Delta)}{4\bar{n}_A D(f+\Delta)(D+\Delta)}$. Thus the local center manifold is given by

$$h(\tilde{y}) = \begin{pmatrix} \frac{f\Delta^2}{4\bar{n}_A D(f+\Delta)(D+\Delta)} \\ \frac{f}{4\bar{n}_A(f+\Delta)} \end{pmatrix} \tilde{y}^2 + \mathcal{O}(\tilde{y}^3). \tag{II.62}$$

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

$$\dot{\tilde{y}} = -\frac{f\Delta}{2\bar{n}_A(f+\Delta)}\tilde{y}^2 + \mathcal{O}(\tilde{y}^3). \tag{II.63}$$

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

$$\frac{2\bar{n}_A(f+\Delta)}{(f\Delta+\varrho)t + \frac{2\bar{n}_A(f+\Delta)}{\varepsilon}} \leq \tilde{y}(t) \leq \frac{2\bar{n}_A(f+\Delta)}{(f\Delta-\varrho)t + \frac{2\bar{n}_A(f+\Delta)}{\varepsilon}}, \tag{II.64}$$

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

In contrast to the model of Collet, Méléard and Metz [18], we see in the proof that $n_{aA}(t)$ does not dies out exponentially fast. Instead it decays like a function $f(t) = \frac{1}{t}$, as soon as $n_{aA} + n_{AA} \approx \bar{n}_A$ and thus survives a longer time in the population (cf. Figure II.1 and II.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 $\mathfrak{n}_{aa}$, it will leave this neighbourhood in finite time. Whereas, if the process is in a neighbourhood of the stable fixed point $\mathfrak{n}_{AA}$, it converges to this point $\mathfrak{n}_{AA}$, but the convergence is slower than in the model of Collet, Méléard and Metz [18].
We now turn to the analysis of the behaviour between these points.

Behaviour between the Fixed Points: We show that the deterministic system (II.43) moves from a neighbourhood of the unstable fixed point $\mathfrak{n}_{aa}$ to a neighbourhood of the stable fixed point $\mathfrak{n}_{AA}$ (Corollary II.1).

Proof of Corollary II.1. The proof is similar to the proof of Theorem C.2 in [18]. 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

$$X_0 = \begin{pmatrix} f\frac{(x+\frac{1}{2}y)^2}{x+y+z} - (D + c(x+y+z))x \\ 2f\frac{(x+\frac{1}{2}y)(z+\frac{1}{2}y)}{x+y+z} - (D + c(x+y+z))y \\ f\frac{(z+\frac{1}{2}y)^2}{x+y+z} - (D + c(x+y+z))z \end{pmatrix}. \tag{II.65}$$

The content of Theorem B.1 in [18] is that X_0 (II.65) has a line of fixed points given by,

$$\Gamma_0(v) = \begin{pmatrix} \frac{(v-\bar{n}_A)^2}{4\bar{n}_A} \\ -\frac{v^2-\bar{n}_A^2}{2\bar{n}_A} \\ \frac{(v+\bar{n}_A)^2}{4\bar{n}_A} \end{pmatrix}, \quad v \in [-\bar{n}_A, \bar{n}_A], \tag{II.66}$$

and that the differential of the vector field X_0 at each point of the curve Γ_0, $DX_0(\Gamma_0(v))$, has the three eigenvectors

$$e_1(v) = \begin{pmatrix} \frac{(v-\bar{n}_A)^2}{4\bar{n}_A} \\ -\frac{v^2-\bar{n}_A^2}{2\bar{n}_A} \\ \frac{(v+\bar{n}_A)^2}{4\bar{n}_A} \end{pmatrix}, \quad e_2(v) = \begin{pmatrix} \frac{v-\bar{n}_A}{2\bar{n}_A} \\ -\frac{v}{\bar{n}_A} \\ \frac{v+\bar{n}_A}{2\bar{n}_A} \end{pmatrix}, \quad e_3(v) = \frac{1}{2\bar{n}_A}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \tag{II.67}$$

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

$$\beta_1(v) = \frac{1}{\bar{n}_A}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \beta_2(v) = \begin{pmatrix} -\frac{v+\bar{n}_A}{\bar{n}_A} \\ -\frac{v}{\bar{n}_A} \\ -\frac{v-\bar{n}_A}{\bar{n}_A} \end{pmatrix}, \quad \beta_3(v) = \begin{pmatrix} \frac{(v+\bar{n}_A)^2}{2\bar{n}_A} \\ \frac{v^2-\bar{n}_A^2}{2\bar{n}_A} \\ \frac{(v-\bar{n}_A)^2}{2\bar{n}_A} \end{pmatrix}, \tag{II.68}$$

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

$$\langle \beta_i(v), e_j(v) \rangle = \delta_{i,j}. \tag{II.69}$$

Next we analyse the asymptotics of the flow associated to the perturbed vector field, X_Δ (II.45), as $t \to \infty$. From Theorem B.1 in [18] we know that the curve of fixed points, Γ_0, is transversally hyperbolic and invariant for the vector field X_0. Thus, Theorem 4.1 in [51] implies that, for small enough Δ, there is an attractive curve, Γ_Δ that is invariant under X_Δ. Moreover, Γ_Δ is regular and converges to Γ_0, as $\Delta \to 0$. Hence, there is a small tubular neighbourhood, $\mathcal{V}$, of Γ_0 such that Γ_Δ is contained in $\mathcal{V}$ and attracts all orbits with initial conditions in $\mathcal{V}$ (cf. Figure (II.5)).

We want to study the flow associated to the vector field X_Δ. From the remarks above we know that the curve Γ_Δ 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 Γ_Δ except for the two fixed points, $\mathfrak{n}_{aa}$ and $\mathfrak{n}_{AA}$. Thus the orbit of any initial condition on Γ_Δ 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 Γ_Δ because of the attractiveness of this curve. Since Hirsch, Pugh and Shub [51] gives us only a local statement, we consider $\mathcal{V}$ in local frames. A point $(x, y, z) \in \mathcal{V}$ is represented by the parametrisation

$$M(v, r, s) = \Gamma_0(v) + re_1(v) + se_3(v) = (1 + r)\Gamma_0(v) + s\frac{d^2\Gamma_0(v)}{dv^2}, \tag{II.70}$$

with $v \in [-\bar{n}_A - \delta, \bar{n}_A + \delta]$ 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 $-\frac{r+1}{2}$ and thus invertible and does not vanish if $0 < \delta < 1$. Moreover, it is a diffeomorphism which maps $[-\bar{n}_A - \delta, \bar{n}_A + \delta] \times [-\delta, \delta]^2$ 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_Δ. First we look for zeros of the two linear combinations of X_Δ with the eigenvectors β_1 and β_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_Δ with β_2 and zeros of this linear combination on the curve above.

Proposition II.12 (Proposition B.2 in [18])**.** *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 Δ and converging to 0 when Δ tends to zero such that, for any $v \in [-\bar{n}_A - \delta, \bar{n}_A + \delta]$, we have*

$$\langle \beta_1(v), X_\Delta(M(v, r_\Delta(v), s_\Delta(v)))\rangle = \langle \beta_3(v), X_\Delta(M(v, r_\Delta(v), s_\Delta(v)))\rangle = 0. \tag{II.71}$$

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

$$\langle \beta_1(v), X_\Delta(M(v, r, s))\rangle = \langle \beta_3(v), X_\Delta(M(v, r, s))\rangle = 0, \tag{II.72}$$

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

Next, we look for the points of the resulting curve (obtained from Proposition II.12) where the third linear combination of the components vanishes. Since

$$\langle \beta_2(v), X(0, M(v, r_0(v), s_0(v)))\rangle = \langle \beta_2(v), X(0, M(v, 0, 0))\rangle = \langle \beta_2(v), X(0, \Gamma_0(v))\rangle = 0, \tag{II.73}$$

this function vanishes for $\Delta = 0$ in v and we apply the Malgrange preparation Theorem [43], which provides a representation for the linear combination near $\Delta = 0$:

$$\langle \beta_2(v), X(\Delta, M(v, r_\Delta(v), s_\Delta(v)))\rangle = \Delta^2 h(\Delta, v) + \Delta g(v), \tag{II.74}$$

where h, g are two smooth functions. To show that the third linear combination indeed vanishes only in small neighbourhoods of the points $\pm\bar{n}_A$, Collet, Méléard and Metz [18] use a representation for the function g which is independent of the perturbation Δ.

Lemma II.2 (Lemma B.3 in [18]). *The function g in* (II.74) *is given by*

$$g(v) = \langle \beta_2(v), \partial_\Delta X_0(\Gamma_0(v))\rangle. \tag{II.75}$$

Then they ensure that this function has only two zeros, which implies, because of the representation (II.74), for $|\Delta| \neq 0$ small enough, that the perturbed vector field X_Δ has only the two known fixed points $\mathfrak{n}_{aa}$ and $\mathfrak{n}_{AA}$.

Theorem II.7 (Theorem B.4 in [18]). *Assume the function*

$$g(v) = \langle \beta_2(v), \partial_\Delta X_0(\Gamma_0(v))\rangle, \tag{II.76}$$

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_Δ has only two zeros in a tubular neighbourhood of Γ_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 II.7 remains true. Namely, we have from (II.45) that

$$\partial_\Delta X_\Delta(x, y, z) = \begin{pmatrix} -x \\ 0 \\ 0 \end{pmatrix}, \tag{II.77}$$

and thus

$$g(v) = \langle \beta_2(v), \partial_\Delta X_0(\Gamma_0(v))\rangle = \left\langle \begin{pmatrix} -\frac{v+\bar{n}_A}{\bar{n}_A} \\ -\frac{v}{\bar{n}_A} \\ -\frac{v-\bar{n}_A}{\bar{n}_A} \end{pmatrix}, \begin{pmatrix} -\frac{(v-\bar{n}_A)^2}{4\bar{n}_A} \\ 0 \\ 0 \end{pmatrix} \right\rangle = \frac{(v+\bar{n}_A)(v-\bar{n}_A)^2}{4\bar{n}_A^2}. \tag{II.78}$$

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 (II.74) that there is a $\delta > 0$ such that, for Δ 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 II.4 we get the existence of these two fixed points, which have to be the points $\mathfrak{n}_{aa}$ and $\mathfrak{n}_{AA}$. Finally, we need the following lemma, which is analogous to Theorem C.1 in Collet, Méléard and Metz [18].

Lemma II.3.

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

(b) The local unstable manifold is contained in the curve Γ_Δ.

Proof. We start by proving (a). From Theorem II.4 we get the hyperbolicity, thus we can apply Theorem 4.1 in Hirsch, Pugh and Shub [51]. The Jacobian matrix $DX_\Delta((\bar{n}_a, 0, 0))$ (cf. (II.50)) has the three eigenvalues $\lambda_1 = -(f - D - \Delta), \lambda_2 = \Delta$ and $\lambda_3 = -(f - \Delta)$ with corresponding eigenvectors

$$EV(\lambda_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_1(-\bar{n}_A) \tag{II.79}$$

$$EV(\lambda_2) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \frac{\Delta}{f - D} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_2(-\bar{n}_A) + \mathcal{O}(\Delta) \tag{II.80}$$

$$EV(\lambda_3) = \frac{1}{2\bar{n}_A} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \frac{\Delta}{2\bar{n}_A(f - D)} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_3(-\bar{n}_A) + \mathcal{O}(\Delta). \tag{II.81}$$

Let $\mathcal{E}^s_{aa}(\Delta)$ the two dimensional affine stable subspace spanned by the eigenvectors $EV(\lambda_1)$ and $EV(\lambda_3)$ with origin in $\mathfrak{n}_{aa}$:

$$\mathcal{E}^s_{aa}(\Delta) = \{x \in \mathbb{R}^3 | x = (\bar{n}_a, 0, 0)^t + sEV(\lambda_1) + tEV(\lambda_3), \forall s, t \in \mathbb{R}\}. \tag{II.82}$$

Again, by the lamination and permanence condition in Theorem 4.1 in [51], we get the existence of a stable manifold $W^{s,loc}_{aa}$ and an unstable manifold $W^{u,loc}_{aa}$ of the fixed point $\mathfrak{n}_{aa}$. The local stable manifold $W^{s,loc}_{aa}$ is a piece of regular manifold tangent in $\mathfrak{n}_{aa}$ to the subspace $\mathcal{E}^s_{aa}(\Delta)$. We see that the x-axis is invariant for X_Δ and is contained in $W^{s,loc}_{aa}$. From (II.82), we get that $\mathcal{E}^s_{aa}(\Delta)$ intersects the closed positive quadrant only along the line $y = z = 0$. Hence, the same is true for $W^{s,loc}_{aa}$ since it is a piece of the subspace $\mathcal{E}^s_{aa}(\Delta)$. This shows (a). To show (b), we show that the local unstable manifold $W^{u,loc}_{aa}$ is contained in the closed positive quadrant. This follows because $W^{u,loc}_{aa}$ is tangent to the linear unstable direction in $EV(\lambda_2)$ in $\mathfrak{n}_{aa}$, which points into the positive quadrant. From Theorem 4.1 in [51] we get that the invariant curve, Γ_Δ, is unique, thus $W^{u,loc}_{aa} \subset \Gamma_\Delta$ and (b) follows by the invariance of the positive quadrant under the flow. □

The preceding steps conclude the proof of Corollary II.1. □

5.2. Proof of the main theorems in Section 3

We carry out the proofs of the main theorems (Section 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:

$$\begin{aligned} b_A &= \frac{2fK}{\Sigma(t)} \left(\left(n_{AA}(t) + \tfrac{1}{2}n_{aA}(t)\right)^2 + \left(n_{AA}(t) + \tfrac{1}{2}n_{aA}(t)\right)\left(n_{aa}(t) + \tfrac{1}{2}n_{aA}(t)\right) \right) \\ &= fK(2n_{AA}(t) + n_{aA}(t)) = A(t)Kf, \end{aligned} \tag{II.83}$$

$$d_A = 2n_{AA}(t)K(D + c\Sigma(t)) + n_{aA}(t)K(D + c\Sigma(t)) = A(t)K(D + c\Sigma(t)). \tag{II.84}$$

Phase 1: Fixation of the mutant population

Recall the stopping times (II.28) and (II.27), when the mutant population $A(t)$ increased to a δ-level and its stopping time of extinction. We show Theorem II.1:

Proof of Theorem II.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 (II.15), there will be no further mutation in the process as we will show.
Proposition II.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 δ-neighbourhood for an exponentially long time as long as the mutant population is less than δ. Hence we get that, as long as the mutant population is smaller than δ, the time the process $n_{aa}(t)$ needs to exit from its domain $\bar{n}_a$ is bigger than $\exp(VK)$ with probability converging to 1, for some $V > 0$ (cf. [14, 18]) and the dynamics of the mutant population are negligible for $n_{aa}(t)$.
With this knowledge we analyse the fate of the mutants for $t < \tau_\delta^{mut} \wedge \tau_0^{mut} \wedge \mathrm{e}^{VK}$. We use the comparison results of birth and death processes (Theorem 2 in [14]) to bound the mutant process from below and above. We denote by $\preccurlyeq$ the following stochastic dominant relation: if $\mathbf{P_1}$ and $\mathbf{P_2}$ are the laws of $\mathbb{R}$-valued processes, we will write $\mathbf{P_1} \preccurlyeq \mathbf{P_2}$ if we can construct on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ two processes X^1 and X^2 such that the law of the processes is $\mathcal{L}(X^i) = \mathbf{P_i}$, $i \in \{1, 2\}$ and for all $t > 0, \omega \in \Omega : X_t^1(\omega) \leq X_t^2(\omega)$.
First we construct a process $A_l(t) \preccurlyeq A(t)$ which is the minorising process of the mutant process. This process has the birth and death rates:

$$b_l(t) = A_l(t)Kf, \qquad d_l(t) = A_l(t)K[D + c(\bar{n}_a + 2\delta)]. \tag{II.85}$$

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

$$b_u(t) = A_u(t)fK, \qquad d_u(t) = A_u(t)K[D + c(\bar{n}_a - \delta)]. \tag{II.86}$$

We define the stopping times

$$T_{n/K}^l \equiv \inf\left\{t \geq 0 : A_l(t) = \frac{n}{K}\right\}, \qquad T_{n/K}^u \equiv \inf\left\{t \geq 0 : A_u(t) = \frac{n}{K}\right\} \tag{II.87}$$

$$\Theta_a \equiv \inf\left\{t \geq 0 : |n_{aa}(t) - \bar{n}_a| > \delta\right\}, \tag{II.88}$$

which are the first times that the processes A_l, resp. A_u reach the level $\frac{n}{K}$ 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 [14]:

Lemma II.4 (Theorem 4 (b) in [14])**.** *Let $b, d > 0$. For any $K \geq 1$ and any $z \in \mathbb{N}/K$, let $\mathbb{P}_z^K$ 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*

$$T_\rho = \inf\{t \geq 0 : \omega_t = \rho\}. \tag{II.89}$$

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

$$(a) \quad \lim_{K \to \infty} \mathbb{P}_{\frac{1}{K}}^K \left(T_0 \leq t_K \wedge T_{\lceil \varepsilon K \rceil / K}\right) = \frac{d}{b}, \tag{II.90}$$

$$(b) \quad \lim_{K \to \infty} \mathbb{P}_{\frac{1}{K}}^K \left(T_{\lceil \varepsilon K \rceil / K} \leq t_K\right) = 1 - \frac{d}{b}. \tag{II.91}$$

With respect to Theorem II.3, we prove the theorem for arbitrary μ_K. From (II.23) we know that the next mutation occurs with high probability not before a time $\frac{\rho}{K\mu_K}$. Then

$$\begin{aligned}\mathbb{P}_{\frac{1}{K}}\Big[T^l_{\lceil \delta K\rceil/K} < T^l_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big] &\leq \mathbb{P}_{\frac{1}{K}}\Big[\tau^{mut}_\delta < \tau^{mut}_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big] \\ &\leq \mathbb{P}_{\frac{1}{K}}\Big[T^u_{\lceil \delta K\rceil/K} < T^u_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big]. \end{aligned} \tag{II.92}$$

Using Proposition II.1, there exists $V > 0$ such that, for sufficiently large K,

$$\mathbb{P}_{\frac{1}{K}}\left[\tau_1 \wedge \mathrm{e}^{VK} < \Theta_a\right] \geq 1 - \delta. \tag{II.93}$$

With (II.93) and for K large enough we can estimate,

$$\begin{aligned}\mathbb{P}_{\frac{1}{K}}\Big[T^l_{\lceil \delta K\rceil/K} < T^l_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big] &\geq \mathbb{P}_{\frac{1}{K}}\Big[T^l_{\lceil \delta K\rceil/K} < T^l_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \mathrm{e}^{VK}\Big] - \delta \\ &= \mathbb{P}_{\frac{1}{K}}\Big[T^l_{\lceil \delta K\rceil/K} < T^l_0 \wedge \tfrac{\rho}{K\mu_K}\Big] - \delta \\ &\geq \mathbb{P}_{\frac{1}{K}}\Big[\Big\{T^l_{\lceil \delta K\rceil/K} < T^l_0 \wedge \tfrac{\rho}{K\mu_K}\Big\} \cap \Big\{T^l_0 \geq \tfrac{\rho}{K\mu_K}\Big\}\Big] - \delta \\ &= \mathbb{P}_{\frac{1}{K}}\Big[\Big\{T^l_{\lceil \delta K\rceil/K} < \tfrac{\rho}{K\mu_K}\Big\} \cap \Big\{T^l_0 \geq \tfrac{\rho}{K\mu_K}\Big\}\Big] - \delta \\ &\geq \mathbb{P}_{\frac{1}{K}}\Big[T^l_{\lceil \delta K\rceil/K} < \tfrac{\rho}{K\mu_K}\Big] - \mathbb{P}_{\frac{1}{K}}\Big[T^l_0 < \tfrac{\rho}{K\mu_K}\Big] - \delta. \end{aligned} \tag{II.94}$$

The extinction time for a binary branching process when $b \neq d$ is given by (cf. page 109 in [1])

$$\mathbb{P}_n(T_0 \leq t) = \left(\frac{d(1-\mathrm{e}^{-(b-d)t})}{b - d\mathrm{e}^{-(b-d)t}}\right)^n, \tag{II.95}$$

for any $t \geq 0$ and $n \in \mathbb{N}$. Under our condition on μ_K (II.15) in Theorem II.3, this implies that

$$\lim_{K\to\infty} \mathbb{P}_{\frac{1}{K}}\Big[T^l_0 \leq \tfrac{\rho}{K\mu_K}\Big] = 0. \tag{II.96}$$

Together with Lemma II.4, we get

$$\lim_{K\to\infty} \mathbb{P}_{\frac{1}{K}}\Big[\tau^{mut}_\delta < \tau^{mut}_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big] \geq 1 - \frac{d_l}{b_l} - \delta = \frac{\Delta}{f} - \left(\frac{c(M+1)}{f} + 1\right)\delta. \tag{II.97}$$

If we next consider the upper bound of (II.92), we see that

$$P_{\frac{1}{K}}\Big[T^u_{\lceil \delta K\rceil/K} < T^u_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big] \leq \mathbb{P}_{\frac{1}{K}}\Big[T^u_{\lceil \delta K\rceil/K} < T^u_0 \wedge \tfrac{\rho}{K\mu_K}\Big]. \tag{II.98}$$

Similar as in (II.97)

$$\lim_{K\to\infty} \mathbb{P}_{\frac{1}{K}}\Big[\tau^{mut}_\delta < \tau^{mut}_0 \wedge \tfrac{\rho}{K\mu_K} \wedge \Theta_a\Big] = \frac{\Delta}{f} + \frac{cM\delta}{f}. \tag{II.99}$$

Now, we let $\delta \to 0$ and get the desired result. □

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 δ due to its advantage in recombination in contrast to the AA population. Applying the Large Population Approximation (Theorem II.11, [36]) with this initial condition, we get that the behaviour of the process is close to the solution of the deterministic system (II.43), when K tends to infinity. We approximate the population process by the solution of the dynamical system (II.43). Result (III.36) in Proposition II.43 is known (see [17] or [18]). We use this result only until the aA population decreases to an ε-level.

Phase 3: Survival of the recessive a allele

This phase starts as soon as the aA population hits the ε 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:

$$n_{aA}(0) = \varepsilon, \tag{II.100}$$

$$n_{aa}(0) \leq \frac{f}{4\bar{n}_A(f+\Delta)}\varepsilon^2 + M_{aa}\varepsilon^{2+\alpha}, \tag{II.101}$$

$$|n_{AA}(0) - (\bar{n}_A - \varepsilon)| \leq M_{AA}\varepsilon^2, \tag{II.102}$$

$$\left|\Sigma(0) - \left(\bar{n}_A - \tfrac{\Delta}{c\bar{n}_A}\gamma\varepsilon^2\right)\right| \leq M_\Sigma\varepsilon^{2+\alpha}, \tag{II.103}$$

where $M_i, i \in \{aa, AA, \Sigma\}$ are constants.
In the following stopping times denoted by τ 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 processes:

$$\begin{aligned} b_\Sigma(t) &= f\Sigma(t)K, \\ d_\Sigma(t) &= D\Sigma(t)K + c\Sigma^2(t)K + \Delta N_{aa}(t), \end{aligned} \tag{II.104}$$

$$\begin{aligned} b_{aa}(t) &= fN_{aa}(t)\left(1 - \frac{n_{AA}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}n_{aA}^2(t), \\ d_{aa}(t) &= N_{aa}(t)(D + \Delta + c\Sigma(t)), \end{aligned} \tag{II.105}$$

$$\begin{aligned} b_{aA}(t) &= fN_{aA}(t)\left(1 - \frac{n_{aA}(t)}{2\Sigma(t)}\right) + 2fN_{aa}(t)\frac{n_{AA}(t)}{\Sigma(t)}, \\ d_{aA}(t) &= N_{aA}(t)(D + c\Sigma(t)), \end{aligned} \tag{II.106}$$

$$\begin{aligned} b_{AA}(t) &= fN_{AA}(t)\left(1 - \frac{n_{aa}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}n_{aA}^2(t), \\ d_{AA}(t) &= N_{AA}(t)(D + c\Sigma(t)). \end{aligned} \tag{II.107}$$

We proceed as described in the outline. Recall the settings for the steps (II.26), (II.27), (II.28) and (II.29).

Step 1: We prove the upper bound of the sum process (Proposition II.2). For this we construct a process, the difference process, which records the drift from the sum process away from the upper bound $\bar{n}_A K$.

Proof of Proposition (II.2). The difference process $X_t^{u\Sigma}$ between $\Sigma(t)K$ and $\bar{n}_A K$ is a branching process with the same rates as $\Sigma(t)K$. Set

$$\begin{aligned} X_t^{u\Sigma} &= \Sigma(t)K - \bar{n}_A K, \\ T_0^{X,u\Sigma} &\equiv \inf\{t \geq 0 : X_t^{u\Sigma} = 0\}, \\ T_{\alpha,M_\Sigma}^{X,u\Sigma} &\equiv \inf\{t \geq 0 : X_t^{u\Sigma} \geq M_\Sigma K^{1/2+\alpha/2}\}. \end{aligned} \tag{II.108}$$

This is a process in continuous time. For the following we need the discrete process $Y_n^{u\Sigma}$ associated to $X_t^{u\Sigma}$. To obtain this process we introduce a sequence of stopping times ϑ_n which records the times, when $X_t^{u\Sigma}$ makes a jump. Formally, ϑ_n is the smallest time t such that $X_t^{u\Sigma} \neq X_s^{u\Sigma}$, for all $\vartheta_{n-1} \leq s < t$. We set $X_{\vartheta_n}^{u\Sigma} = Y_n^{u\Sigma}$. This discretisation has probability less than $\frac{1}{2}$ to make an upward jump:

Lemma II.5. *For $n \in \mathbb{N}$ such that $\vartheta_n \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$ and $1 \leq k \leq M_\Sigma K^{1/2+\alpha/2}$ there exists a constant $C_0 > 0$ such that*

$$\mathbb{P}[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k] \leq \frac{1}{2} - C_0 k K^{-1} \equiv p_\Sigma(k). \tag{II.109}$$

Proof. With the rates (II.104), we get by some straightforward computations

$$\begin{aligned} \mathbb{P}[Y_{n+1} = k+1 | Y_n = k] &= \frac{b_\Sigma}{b_\Sigma + d_\Sigma} \\ &= \frac{(\bar{n}_A K + k) f}{(\bar{n}_A K + k)[f + D + c(\bar{n}_A + kK^{-1})] + \Delta N_{aa}(t)} \\ &\leq \frac{1}{2} + \frac{\frac{1}{2}f - \frac{1}{2}D - \frac{1}{2}c\bar{n}_A - \frac{1}{2}ckK^{-1} - \frac{1}{2}\frac{\Delta N_{aa}(t)}{\bar{n}_A K + k}}{f + D + c(\bar{n}_A + kK^{-1}) + \frac{\Delta N_{aa}(t)}{\bar{n}_A K + k}} \\ &\leq \frac{1}{2} + \frac{-\frac{1}{2}ckK^{-1} - \frac{1}{2}\frac{\Delta n_{aa}(t)}{\bar{n}_A + kK^{-1}}}{2f + ckK^{-1} + \frac{\Delta n_{aa}(t)}{\bar{n}_A} + \mathcal{O}((x^i\varepsilon)^{4+\alpha})} \\ &\leq \frac{1}{2} - \frac{\frac{1}{2}ckK^{-1}}{2f + ckK^{-1} + \frac{\Delta n_{aa}(t)}{\bar{n}_A} + \mathcal{O}((x^i\varepsilon)^{4+\alpha})} \\ &= \frac{1}{2} - C_0 k K^{-1}. \end{aligned} \tag{II.110}$$

□

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

$$(1)\quad Z_0^{u\Sigma} = Y_0^{u\Sigma} \vee 0, \tag{II.111}$$

$$(2)\quad \mathbb{P}[Z_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} < Z_n^{u\Sigma} = k] = p_\Sigma(k), \tag{II.112}$$

$$(3)\quad \mathbb{P}[Z_{n+1}^{u\Sigma} = k-1 | Y_n^{u\Sigma} < Z_n^{u\Sigma} = k] = 1 - p_\Sigma(k), \tag{II.113}$$

$$(4)\quad \mathbb{P}[Z_{n+1}^{u\Sigma} = k+1 | Y_{n+1}^{u\Sigma} = k+1, Y_n^{u\Sigma} = Z_n^{u\Sigma} = k] = 1, \tag{II.114}$$

$$(5)\quad \mathbb{P}[Z_{n+1}^{u\Sigma} = k+1 | Y_{n+1}^{u\Sigma} = k-1, Y_n^{u\Sigma} = Z_n^{u\Sigma} = k] = \tfrac{p_\Sigma(k) - \mathbb{P}[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k]}{1 - \mathbb{P}[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k]}, \tag{II.115}$$

$$(6)\quad \mathbb{P}[Z_{n+1}^{u\Sigma} = k-1 | Y_{n+1}^{u\Sigma} = k-1, Y_n^{u\Sigma} = Z_n^{u\Sigma} = k] = 1 - \tfrac{p_\Sigma(k) - \mathbb{P}[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k]}{1 - \mathbb{P}[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k]}. \tag{II.116}$$

Observe that by construction $Z_n^{u\Sigma} \succcurlyeq Y_n^{u\Sigma}$, a.s. and that $\mathbb{P}\left[Z_{n+1}^{u\Sigma} = 1 | Z_n^{u\Sigma} = 0\right] = 1$. The marginal distribution of $Z_n^{u\Sigma}$ is the desired Markov chain with transition probabilities

$$\begin{aligned}
&\mathbb{P}\left[Z_{n+1}^{u\Sigma} = k+1 | Z_n^{u\Sigma} = k\right] \\
&= p_\Sigma(k)\mathbb{P}\left[Y_n^{u\Sigma} < Z_n^{u\Sigma} | Z_n^{u\Sigma} = k\right] + \mathbb{P}\left[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k\right] \mathbb{P}\left[Y_n^{u\Sigma} = Z_n^{u\Sigma} | Z_n^{u\Sigma} = k\right] \\
&\quad + \tfrac{p_\Sigma(k) - \mathbb{P}\left[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k\right]}{1 - \mathbb{P}\left[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k\right]} \mathbb{P}\left[Y_n^{u\Sigma} = Z_n^{u\Sigma} | Z_n^{u\Sigma} = k\right] \left(1 - \mathbb{P}\left[Y_{n+1}^{u\Sigma} = k+1 | Y_n^{u\Sigma} = k\right]\right) \\
&= p_\Sigma(k) \left(\mathbb{P}\left[Y_n^{u\Sigma} < Z_n^{u\Sigma} | Z_n^{u\Sigma} = k\right] + \mathbb{P}\left[Y_n^{u\Sigma} = Z_n^{u\Sigma} | Z_n^{u\Sigma} = k\right]\right) = p_\Sigma(k),
\end{aligned} \tag{II.117}$$

$$\mathbb{P}\left[Z_{n+1}^{u\Sigma} = k-1 | Z_n^{u\Sigma} = k\right] = 1 - p_\Sigma(k), \tag{II.118}$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1} \frac{1}{2} - C_0 k K^{-1}}{\prod_{k=1}^{n} \frac{1}{2} + C_0 k K^{-1}}, \tag{II.119}$$

with $\mu(0) = 1$ and $\mu(1) = \frac{1}{\frac{1}{2} + C_0 K^{-1}}$. We want to calculate the probability that the Markov chain $Z_n^{u\Sigma}$, starting at a point $0 < zK < \frac{1}{2} M_\Sigma K^{1/2+\alpha/2}$, reaches first $M_\Sigma K^{1/2+\alpha/2}$ before going to zero, which is the equilibrium potential of a one dimensional chain (see [9]). Using Equation 7.1.57 in the book by Bovier and den Hollander [11], Chapter 7.1.4, we get (for K large enough)

$$\begin{aligned}
\mathbb{P}_{zK}[T^Z_{\alpha, M_\Sigma} < T^Z_0] &= \frac{\sum_{n=1}^{zK} \frac{1}{1-p_\Sigma(n)} \frac{1}{\mu(n)}}{\sum_{n=1}^{3M_\Sigma (x^{2i}\varepsilon^2)^{1+\alpha} K} \frac{1}{1-p_\Sigma(n)} \frac{1}{\mu(n)}} \\
&= \frac{\sum_{n=1}^{zK} \prod_{k=1}^{n-1} \frac{1+2C_0 k K^{-1}}{1-2C_0 k K^{-1}}}{\sum_{n=1}^{M_\Sigma K^{1/2+\alpha/2}} \prod_{k=1}^{n-1} \frac{1+2C_0 k K^{-1}}{1-2C_0 k K^{-1}}} \\
&\leq \frac{\sum_{n=1}^{zK} \exp\left(\sum_{k=1}^{n-1} \ln\left(\frac{1+2C_0 k K^{-1}}{1-2C_0 k K^{-1}}\right)\right)}{\sum_{n=1}^{M_\Sigma K^{1/2+\alpha/2}} \exp\left(\sum_{k=1}^{n-1} \ln\left(\frac{1+2C_0 k K^{-1}}{1-2C_0 k K^{-1}}\right)\right)} \\
&\leq \frac{\sum_{n=1}^{zK} \exp\left(\sum_{k=1}^{n-1} 4C_0 k K^{-1}\right)}{\sum_{n=1}^{M_\Sigma K^{1/2+\alpha/2}} \exp\left(\sum_{k=1}^{n-1} 4C_0 k K^{-1} - \mathcal{O}\left((kK^{-1})^2\right)\right)}
\end{aligned}$$

$$\leq \frac{\sum_{n=1}^{zK} \exp\left(2C_0 n^2 K^{-1} - 2C_0 n K^{-1}\right)}{\sum_{n=1}^{M_\Sigma K^{1/2+\alpha/2}} \exp\left(2C_0 n^2 K^{-1} - 2C_0 n K^{-1} - \mathcal{O}\left(n(nK^{-1})^2\right)\right)}$$
$$\leq \frac{zK \exp\left(2C_0 z^2 K\right)}{\sum_{n=1/2M_\Sigma K^{1/2+\alpha/2}}^{M_\Sigma K^{1/2+\alpha/2}} \exp\left(2C_0 n^2 K^{-1} - 2C_0 n K^{-1} - \mathcal{O}\left(n(nK^{-1})^2\right)\right)}$$
$$\leq \frac{2zK^{-1/2-\alpha/2} \exp\left(2C_0 z^2 K\right)}{M_\Sigma \exp\left(\frac{1}{2}C_0 M_\Sigma^2 K^\alpha - 2C_0 M_\Sigma K^{-1/2+\alpha/2} - \mathcal{O}\left(K^{-1/2+3/2\alpha}\right)\right)}$$
$$\leq \exp\left(-2\hat{C}_0 K^\alpha \left(\tfrac{1}{4}M_\Sigma^2 - z^2 K^{1-\alpha}\right)\right). \tag{II.120}$$

We denote by R the number of times that the process $Z_n^{u\Sigma}$ returns to zero before it reaches $M_\Sigma K^{1/2+\alpha/2}$. 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:

$$T_0^1 = \inf\{t > 0 : Z_t^{u\Sigma} = 0\}, \qquad T_0^n = \inf\{t > T_0^{n-1} : Z_t^{u\Sigma} = 0\}. \tag{II.121}$$

We then have

$$R_z^k = \mathbb{P}_{zK}\left[T_0 < T_{\alpha,M_\Sigma}^{Z^{u\Sigma}}\right]\left(1 - \mathbb{P}_0\left[T_{\alpha,M_\Sigma}^{Z^{u\Sigma}} < T_0\right]\right)^{k-1} \mathbb{P}_0\left[T_{\alpha,M_\Sigma}^{Z^{u\Sigma}} < T_0\right] \leq (1-A)^{k-1}A, \tag{II.122}$$

where

$$A \equiv \mathbb{P}_0\left[T_{\alpha,M_\Sigma}^{Z^{u\Sigma}} < T_0\right] = \sum_{i\geq 1} p(0,i)\mathbb{P}_i\left[T_{\alpha,M_\Sigma}^{Z^{u\Sigma}} < T_0\right] = p(0,1)\mathbb{P}_1\left[T_{\alpha,M_\Sigma}^{Z^{u\Sigma}} < T_0\right]$$
$$\leq \exp\left(-\tilde{C}_0 M_\Sigma^2 K^\alpha\right), \tag{II.123}$$

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

$$\mathbb{P}[R \leq N] \leq \sum_{i=1}^{N} R_z^i \leq \sum_{i=1}^{N}(1-A)^{i-1}A = 1 - (1-A)^N. \tag{II.124}$$

We choose, e.g., $N \sim \frac{1}{K^2 A}$, so that $\mathbb{P}\left[R \leq N\right] = o\left(K^{-1}\right)$. Let $I_0 \equiv T_0^1$ and $I_n \equiv T_0^n - T_0^{n-1}$ 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:

$$\sum_{n=1}^{R} I_n \leq T_{\alpha,M_\Sigma}^{Z^{u\Sigma}} \leq \sum_{n=1}^{R+1} I_n. \tag{II.125}$$

The underlying process, the sum process $\Sigma(t)$ (II.12), of $Z_n^{u\Sigma}$ jumps with a rate

$$\lambda_\Sigma = f\Sigma(t)K + D\Sigma(t)K + c\Sigma(t)^2 K + \Delta N_{aa}(t) \leq C_\lambda K. \tag{II.126}$$

Since the Markov chain $Z_n^{u\Sigma}$ 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

$$\mathbb{P}[I_j < y] \leq \mathbb{P}[W < y] = 1 - \exp(-C_\lambda K y). \tag{II.127}$$

We have

$$\begin{aligned}\mathbb{P}\Big[T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}}\Big] &\leq \mathbb{P}\left[T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}}\right] \\ &= \mathbb{P}\left[T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R > N\}\right] + \mathbb{P}\left[T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R \leq N\}\right] \\ &\leq \mathbb{P}\left[T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R > N\}\right] + \mathbb{P}[R \leq N]. \end{aligned} \tag{II.128}$$

First we estimate $\mathbb{P}\left[T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R > N\}\right]$. Since $T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} \geq \sum_{n=1}^{R} I_n$, it holds that if $\frac{n}{2}$ of the I_j are greater than $\frac{2}{n}\mathrm{e}^{VK^{\alpha/2}}$, then $T^{Z^{u\Sigma}}_{\alpha,M_\Sigma} \geq \mathrm{e}^{VK^{\alpha/2}}$.

$$\begin{aligned}\mathbb{P}\left[T^{X,u\Sigma}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R > N\}\right] &\leq \sum_{n=N}^{\infty} \mathbb{P}\left[T^{X,u\Sigma}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R = n\}\right] \\ &\leq \sum_{n=N}^{\infty} \mathbb{P}\left[\sum_{j=1}^{n} I_j < \mathrm{e}^{VK^{\alpha/2}}\right] \\ &\leq \sum_{n=N}^{\infty} \mathbb{P}\left[\sum_{j=1}^{n} \mathbb{1}_{\left\{I_j < \frac{2}{n}\mathrm{e}^{VK^{\alpha/2}}\right\}} > \frac{n}{2}\right]. \end{aligned} \tag{II.129}$$

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

$$\begin{aligned}\sum_{n=N}^{\infty} \mathbb{P}\left[\sum_{j=1}^{n} \mathbb{1}_{\left\{I_j < \frac{2}{n}\mathrm{e}^{VK^{\alpha/2}}\right\}} > \frac{n}{2}\right] &= \sum_{n=N}^{\infty} \sum_{m=n/2}^{n} \binom{n}{m} (1-p_n)^{n-m} p_n^m \\ &\leq \sum_{n=N}^{\infty} 4^n p_n^{\frac{n}{2}} \leq \frac{(16p_N)^{N/2}}{1 - 4p_N^{1/2}}, \end{aligned} \tag{II.130}$$

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

$$\begin{aligned}(16p_N)^{N/2} &\leq \left(16\left(1 - \exp\left(-\tfrac{2C_\lambda}{N}K\mathrm{e}^{VK^{\alpha/2}}\right)\right)\right)^{N/2} \leq \left(16\left(\tfrac{2C_\lambda}{N}K\mathrm{e}^{VK^{\alpha/2}}\right)\right)^{N/2} \\ &= \left(16\left(2C_\lambda A K^3 \mathrm{e}^{VK^{\alpha/2}}\right)\right)^{N/2}. \end{aligned} \tag{II.131}$$

Recalling that $A = \mathrm{e}^{-O(K^\alpha)}$, one sees that (II.130) is bounded by $o\left(\mathrm{e}^{-K^\alpha}\right)$. Hence we get

$$\begin{aligned}\mathbb{P}\left[T^{X,u\Sigma}_{\alpha,M_\Sigma} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}}\right] &\leq \mathbb{P}\left[T^{X,u\Sigma}_{\alpha,M_\Sigma} < \mathrm{e}^{VK^{\alpha/2}} \cap \{R > N\}\right] + \mathbb{P}[R \leq N] \\ &= o\left(\frac{1}{K}\right). \end{aligned} \tag{II.132}$$

This concludes the proof of Proposition II.2. □

Step 2: We derive a rough upper bound on the process n_{aa} (Proposition II.3). Recall that

$$\gamma_\Delta \equiv \frac{f+\frac{\Delta}{2}}{4\bar{n}_A(f+\Delta)}. \tag{II.133}$$

Proof of Proposition II.3. The proof is similar to the one of Proposition II.2. Again, we define the difference process X_t^{aa} between N_{aa} and $\gamma_\Delta x^{2i}\varepsilon^2 K$. This is a branching process with the same rates as n_{aa}. Set

$$\begin{aligned} X_t^{aa} &= N_{aa}(t) - \gamma_\Delta x^{2i}\varepsilon^2 K, && \text{(II.134)}\\ T_0^{X,aa} &\equiv \inf\{t\geq 0 : X_t^{aa}=0\}, && \text{(II.135)}\\ T_{\alpha,M_{aa}}^{X,aa} &\equiv \inf\{t\geq 0 : X_t^{aa}\geq M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}K\}. && \text{(II.136)}\end{aligned}$$

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

Lemma II.6. *For $n\in\mathbb{N}$ such that $\vartheta_n \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+}\wedge\tau_{aA}^{(i+1)-}\wedge \mathrm{e}^{VK^{\alpha/2}}\right]$ and $1\leq k\leq M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}K$, there exists a constant $C_0>0$ such that*

$$\mathbb{P}\left[Y_{n+1}^{aa}=k+1|Y_n^{aa}=k\right]\leq \frac{1}{2}-C_0\equiv p_{aa}. \tag{II.137}$$

Proof. In the following we use Proposition II.1 for the first iteration step, since the mutant population n_{AA} increased to an ε-neighbourhood of its equilibrium $\bar{n}_A$ and the other two populations decreased to an ε order. Thus the influence of the small aA- 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 the i^{th} iteration-step we use the finer bounds of $n_{AA}(t)$ (Proposition II.5 and II.6) and $\Sigma(t)$ (Proposition II.4) estimated in the $(i-1)^{th}$ iteration-step before. By (II.105), we have

$$\begin{aligned}
&\mathbb{P}[Y_{n+1}^{aa}=k+1|Y_n^{aa}=k]\\
&\quad=\frac{(\gamma_\Delta x^{2i}\varepsilon^2K+k)f\left(1-\frac{n_{AA}(t)}{\Sigma(t)}\right)+\frac{fK}{4\Sigma(t)}n_{aA}^2(t)}{(\gamma_\Delta x^{2i}\varepsilon^2K+k)\left[f\left(1-\frac{n_{AA}(t)}{\Sigma(t)}\right)+D+\Delta+c\Sigma(t)\right]+\frac{fK}{4\Sigma(t)}n_{aA}^2(t)}\\
&\quad=\frac{1}{2}+\frac{\frac{1}{2}f-\frac{f}{2\Sigma(t)}n_{AA}(t)+\frac{fK}{8\Sigma(t)}\frac{n_{aA}^2(t)}{(\gamma_\Delta x^{2i}\varepsilon^2K+k)}-\frac{1}{2}D-\frac{1}{2}\Delta-\frac{1}{2}c\Sigma(t)}{f\left(1-\frac{n_{AA}(t)}{\Sigma(t)}\right)+D+\Delta+c\Sigma(t)+\frac{fK}{4\Sigma(t)}\frac{n_{aA}^2(t)}{(\gamma_\Delta x^{2i}\varepsilon^2K+k)}}.
\end{aligned} \tag{II.138}$$

Using Propositions II.6, II.4 , and II.2 and (II.25), one sees that the numerator in the second summand of (II.138) is bounded from above by

$$\begin{aligned}
&\frac{f}{2}-\frac{f(\bar{n}_A-x^{i-1}\varepsilon)}{2(\bar{n}_A+M_\Sigma K^{-1/2+\alpha/2})}+\frac{f(x^i\varepsilon+K^{-1/2+3/4\alpha})^2}{8\gamma_\Delta x^{2i}\varepsilon^2(\bar{n}_A-\frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i-2}\varepsilon^2+M_\Sigma K^{-1/2+\alpha/2})}-\frac{D+\Delta+c\bar{n}_A}{2}+\mathcal{O}(x^i\varepsilon)\\
&\leq -\frac{f+\Delta}{2}+\frac{f}{8\bar{n}_A\gamma_\Delta}+\mathcal{O}(x^i\varepsilon)=-\frac{\Delta(f+\Delta)}{2(2f+\Delta)}+\mathcal{O}(x^i\varepsilon).
\end{aligned} \tag{II.139}$$

For ε small enough, there exists a constant $C_0>0$ such that

$$\mathbb{P}[Y_{n+1}^{aa}=k+1|Y_n^{aa}=k]\leq\frac{1}{2}-C_0\equiv p_{aa}. \tag{II.140}$$

□

As in Step 1 we couple Y_n^{aa} with a process Z_n^{aa} via:

$$(1)\quad Z_0^{aa} = Y_0^{aa} \vee 0, \tag{II.141}$$
$$(2)\quad \mathbb{P}[Z_{n+1}^{aa} = k+1 | Y_n^{aa} < Z_n^{aa} = k] = p_{aa}, \tag{II.142}$$
$$(3)\quad \mathbb{P}[Z_{n+1}^{aa} = k-1 | Y_n^{aa} < Z_n^{aa} = k] = 1 - p_{aa}, \tag{II.143}$$
$$(4)\quad \mathbb{P}[Z_{n+1}^{aa} = k+1 | Y_{n+1}^{aa} = k+1, Y_n^{aa} = Z_n^{aa} = k] = 1, \tag{II.144}$$
$$(5)\quad \mathbb{P}[Z_{n+1}^{aa} = k+1 | Y_{n+1}^{aa} = k-1, Y_n^{aa} = Z_n^{aa} = k] = \tfrac{p_0 - \mathbb{P}[Y_{n+1}^{aa}=k+1 | Y_n^{aa}=k]}{1-\mathbb{P}[Y_{n+1}^{aa}=k+1|Y_n^{aa}=k]}, \tag{II.145}$$
$$(6)\quad \mathbb{P}[Z_{n+1}^{aa} = k-1 | Y_{n+1}^{aa} = k-1, Y_n^{aa} = Z_n^{aa} = k] = 1 - \tfrac{p_0 - \mathbb{P}[Y_{n+1}^{aa}=k+1 | Y_n^{aa}=k]}{1-\mathbb{P}[Y_{n+1}^{aa}=k+1|Y_n^{aa}=k]}. \tag{II.146}$$

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

$$\mathbb{P}[Z_{n+1}^{aa} = k+1 | Z_n^{aa} = k] = p_{aa} = 1 - \mathbb{P}[Z_{n+1}^{aa} = k-1 | Z_n^{aa} = k], \tag{II.147}$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1}\left(\frac{1}{2} - C_0\right)}{\prod_{k=1}^{n}\left(\frac{1}{2} + C_0\right)} = \frac{\left(\frac{1}{2} - C_0\right)^{n-1}}{\left(\frac{1}{2} + C_0\right)^{n}}. \tag{II.148}$$

The remainder of the proof is a complete re-run of the proof of Proposition II.2 and we skip the details. □

Step 3: We estimate the lower bound on $\Sigma(t)$, for $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$.

Proof of Proposition II.4. The proof is similar to those of Proposition II.2 and II.3. We only perform the crucial steps. This time the difference process is given by the difference of $\Sigma(t)$ and $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma_\Delta x^{2i}\varepsilon^2$. Let

$$X_t^{l\Sigma} = \Sigma(t)K - \left(\bar{n}_A - \tfrac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2\right)K, \tag{II.149}$$
$$T_0^{X,l\Sigma} \equiv \inf\{t \geq 0 : X_t^{l\Sigma} = 0\}, \tag{II.150}$$
$$T_{\alpha,M_\Sigma}^{X,l\Sigma} \equiv \inf\{t \geq 0 : X_t^{l\Sigma} \leq -M_\Sigma K^{1/2+\alpha/2}\}. \tag{II.151}$$

As described in Step 1 we construct the discrete process $Y_n^{l\Sigma}$ associated to $X_t^{l\Sigma}$. We show that $Y_n^{l\Sigma}$ jumps down with a probability less than $\frac{1}{2}$:

Lemma II.7. *For $n \in \mathbb{N}$ such that $\vartheta_n \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$ and $1 \leq k \leq M_\Sigma K^{1/2+\alpha/2}$ there exists a constant $C_0 > 0$ such that*

$$\mathbb{P}[Y_{n+1} = -k-1 | Y_n = -k] \leq \frac{1}{2} - C_0 x^{2i}\varepsilon^2 - C_1 k K^{-1} \equiv p_\Sigma. \tag{II.152}$$

Proof. Using the rates of the sum process (II.104) and the upper bound on n_{aa} (Proposition II.3), this is a simple computation and we skip the details. □

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:

$$(1)\quad Z_0 = Y_0 \vee 0, \tag{II.153}$$

$$(2)\quad \mathbb{P}[Z_{n+1}^{l\Sigma} = -k+1 | Y_n^{l\Sigma} > Z_n^{l\Sigma} = -k] = 1 - p_\Sigma, \tag{II.154}$$

$$(3)\quad \mathbb{P}[Z_{n+1}^{l\Sigma} = -k-1 | Y_n^{l\Sigma} > Z_n^{l\Sigma} = -k] = p_\Sigma, \tag{II.155}$$

$$(4)\quad \mathbb{P}[Z_{n+1}^{l\Sigma} = -k-1 | Y_{n+1}^{l\Sigma} = -k-1, Y_n^{l\Sigma} = Z_n^{l\Sigma} = -k] = 1, \tag{II.156}$$

$$(5)\quad \mathbb{P}[Z_{n+1}^{l\Sigma} = -k-1 | Y_{n+1}^{l\Sigma} = -k+1, Y_n^{l\Sigma} = Z_n^{l\Sigma} = -k] = \tfrac{p_\Sigma - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k-1 | Y_n^{l\Sigma} = -k]}{1 - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k-1 | Y_n^{l\Sigma} = -k]}, \tag{II.157}$$

$$(6)\quad \mathbb{P}[Z_{n+1}^{l\Sigma} = -k+1 | Y_{n+1}^{l\Sigma} = -k+1, Y_n^{l\Sigma} = Z_n^{l\Sigma} = -k] = 1 - \tfrac{p_\Sigma - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k-1 | Y_n^{l\Sigma} = -k]}{1 - \mathbb{P}[Y_{n+1}^{l\Sigma} = -k-1 | Y_n^{l\Sigma} = -k]}. \tag{II.158}$$

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

$$\mathbb{P}\left[Z_{n+1}^{l\Sigma} = k+1 | Z_n^{l\Sigma} = k\right] = p_\Sigma, \tag{II.159}$$

$$\mathbb{P}\left[Z_{n+1}^{l\Sigma} = k-1 | Z_n^{l\Sigma} = k\right] = 1 - p_\Sigma, \tag{II.160}$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1}\left(\frac{1}{2} - C_0 x^{2i}\varepsilon^2 - C_1 k K^{-1}\right)}{\prod_{k=1}^{n}\left(\frac{1}{2} + C_0 x^{2i}\varepsilon^2 + C_1 k K^{-1}\right)}. \tag{II.161}$$

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 $M_\Sigma K^{1/2+\alpha/2}$ and calculate the duration of one zero-return to get the desired result. □

Step 4: With Propositions II.2-II.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 $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$.

Step 4.1: We now prove Proposition II.5, the lower bound on n_{AA}.

Proof of Proposition II.5. From Proposition II.4 we know that $\Sigma(t) \geq \bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{1/2+\alpha/2}$. With the upper bound in Proposition II.3 for n_{aa} and the settings for the steps used for the aA population, we get

$$\begin{aligned} n_{AA}(t) &= \Sigma(t) - n_{aA}(t) - n_{aa}(t) \\ &\geq \bar{n}_A - x^i\varepsilon - x^{2i}\varepsilon^2\left(\tfrac{\Delta+\vartheta}{c\bar{n}_A}\gamma + \gamma_\Delta\right) - M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha} - \mathcal{O}\left(K^{-1/2+3/4\alpha}\right) \\ &\geq \bar{n}_A - x^i\varepsilon - \mathcal{O}\left(x^{2i}\varepsilon^2 + K^{-1/2+3/4\alpha}\right). \end{aligned} \tag{II.162}$$

□

Step 4.2: We prove Proposition II.6, the upper bound on n_{AA}.

Proof of Proposition II.6. From Proposition II.2 we know that $\Sigma(t) \leq \bar{n}_A + M_\Sigma K^{-1/2+\alpha/2}$. With the lower bound on $n_{aA}(t)$ defined in the settings we get

$$\begin{aligned} n_{AA}(t) &= \Sigma(t) - n_{aA}(t) - n_{aa}(t) \\ &\leq \bar{n}_A + M_\Sigma K^{-1/2+\alpha/2} - x^{i+1}\varepsilon \\ &\leq \bar{n}_A - x^{i+1}\varepsilon + \mathcal{O}(K^{-1/2+\alpha/2}). \end{aligned} \tag{II.163}$$

□

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 $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$. 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 $x^i\varepsilon$ and show that with high probability the aA- population decreases to $x^{i+1}\varepsilon$ before it exceeds again the $x^i\varepsilon + K^{-1/2+3/4\alpha}$-value (Proposition II.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 $x^{i+1}\varepsilon$ before going back to $x^i\varepsilon + K^{-1/2+3/4\alpha}$.

Proof of Proposition II.7. As before let Y_n^{aA} (cf. Step 1) be the associated discrete process to $N_{aA}(t)$. We start by coupling Y_n^{aA} with a Markov chain Z_n such that $Z_n^u \succcurlyeq Y_n^{aA}$, a.s..

Lemma II.8. *For $n \in \mathbb{N}$ such that $\vartheta_n \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$ there exists a constant $C_0 > 0$ such that*

$$\mathbb{P}\left[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k\right] \leq \frac{1}{2} - C_0 x^{i+1}\varepsilon \equiv p_{aA}^u. \tag{II.164}$$

Proof. For $t < \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}$, we have

$$\begin{aligned} \mathbb{P}\left[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k\right] &= \frac{fk\left(1 - \frac{kK^{-1}}{2\Sigma(t)}\right) + 2fN_{aa}(t)\frac{n_{AA}(t)}{\Sigma(t)}}{fk\left(1 - \frac{kK^{-1}}{2\Sigma(t)}\right) + 2fN_{aa}(t)\frac{n_{AA}(t)}{\Sigma(t)} + k[D + c\Sigma(t)]} \\ &\leq \frac{1}{2} + \frac{\frac{1}{2}f - \frac{f}{4\Sigma(t)}kK^{-1} + \frac{fn_{AA}(t)}{k\Sigma(t)}N_{aa}(t) - \frac{1}{2}D - \frac{1}{2}c\Sigma(t)}{f\left(1 - \frac{kK^{-1}}{2\Sigma(t)}\right) + 2\frac{fN_{aa}(t)}{k}\frac{n_{AA}(t)}{\Sigma(t)} + [D + c\Sigma(t)]}. \end{aligned} \tag{II.165}$$

As in the previous steps, we bound the nominator of the second summand in (II.165) using Propositions II.2, II.4, and II.6, from above by

$$\begin{aligned} &\tfrac{1}{2}f - \tfrac{f}{4\bar{n}_A}kK^{-1} + \frac{f\left(\bar{n}_A - x^{i+1}\varepsilon + \mathcal{O}\left(K^{-1/2+\alpha/2}\right)\right)}{kK^{-1}\left(\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}\right)}\gamma_\Delta x^{2i}\varepsilon^2 \\ &\qquad - \tfrac{1}{2}D - \tfrac{1}{2}c\left(\bar{n}_A - \tfrac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}\right) + \mathcal{O}\left(\left(x^i\varepsilon\right)^{1+2\alpha}\right) \\ &\leq -\frac{f\left(\vartheta - \frac{\Delta}{2}\right)}{4\bar{n}_A(f+\vartheta)}x^{i+1}\varepsilon + \mathcal{O}\left(\left(x^i\varepsilon\right)^{1+2\alpha} + K^{-1/2+\alpha/2}\right). \end{aligned} \tag{II.166}$$

This term is negative since $\frac{\Delta}{2} < \vartheta$. Hence, we get

$$\mathbb{P}\left[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k\right] \leq \frac{1}{2} - C_0 x^{i+1}\varepsilon. \tag{II.167}$$

□

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

$$(1)\quad Z_0^u = Y_0^{aA}, \tag{II.168}$$
$$(2)\quad \mathbb{P}[Z_{n+1}^u = k+1 | Y_n^{aA} < Z_n^u = k] = p_{aA}^u, \tag{II.169}$$
$$(3)\quad \mathbb{P}[Z_{n+1}^u = k-1 | Y_n^{aA} < Z_n^u = k] = 1 - p_{aA}^u, \tag{II.170}$$
$$(4)\quad \mathbb{P}[Z_{n+1}^u = k+1 | Y_{n+1}^{aA} = k+1, Y_n^{aA} = Z_n^u = k] = 1, \tag{II.171}$$
$$(5)\quad \mathbb{P}[Z_{n+1}^u = k+1 | Y_{n+1}^{aA} = k-1, Y_n^{aA} = Z_n^u = k] = \frac{p_{aA}^u - \mathbb{P}[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k]}{1 - \mathbb{P}[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k]}, \tag{II.172}$$
$$(6)\quad \mathbb{P}[Z_{n+1}^u = k-1 | Y_{n+1}^{aA} = k-1, Y_n^{aA} = Z_n^u = k] = 1 - \frac{p_{aA}^u - \mathbb{P}[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k]}{1 - \mathbb{P}[Y_{n+1}^{aA} = k+1 | Y_n^{aA} = k]}. \tag{II.173}$$

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

$$\mathbb{P}[Z_{n+1}^u = k+1 | Z_n^u = k] = p_{aA}^u, \tag{II.174}$$
$$\mathbb{P}[Z_{n+1}^u = k-1 | Z_n^u = k] = 1 - p_{aA}^u, \tag{II.175}$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1}\left(\frac{1}{2} - C_0 x^{i+1}\varepsilon\right)}{\prod_{k=1}^{n}\left(\frac{1}{2} + C_0 x^{i+1}\varepsilon\right)} = \frac{\left(\frac{1}{2} - C_0 x^{i+1}\varepsilon\right)^{n-1}}{\left(\frac{1}{2} + C_0 x^{i+1}\varepsilon\right)^{n}}. \tag{II.176}$$

We define the stopping times

$$T_{i+}^Z \equiv \inf\left\{\vartheta_n \geq 0 : Z_n \geq x^i \varepsilon K + K^{1/2+3/4\alpha}\right\}, \tag{II.177}$$
$$T_{(i+1)-}^Z \equiv \inf\left\{\vartheta_n \geq 0 : Z_n \leq x^{i+1}\varepsilon K\right\}. \tag{II.178}$$

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

$$\begin{aligned}\mathbb{P}_{zK}\left[T_{i+}^{Z^u} < T_{(i+1)-}^{Z^u}\right] &= \frac{\sum_{n=x^{i+1}\varepsilon K+1}^{zK}\left(\frac{1+2C_0x^{i+1}\varepsilon}{1-2C_0x^{i+1}\varepsilon}\right)^{n-1}}{\sum_{n=x^{i+1}\varepsilon K+1}^{x^i\varepsilon K + K\max\{x^{2i}\varepsilon^2, K^{-1/2}\}}\left(\frac{1+2C_0x^{i+1}\varepsilon}{1-2C_0x^{i+1}\varepsilon}\right)^{n-1}} \\ &\leq K^{1/2-3/4\alpha}\exp\left(-\tilde{C}K^{7/4\alpha}\right).\end{aligned} \tag{II.179}$$

Now we couple Y_n^{aA} with a Markov chain Z_n^l such that $Z_n^l \preccurlyeq Y_n^{aA}$, a.s..

Lemma II.9. *For $n \in \mathbb{N}$ such that $\vartheta_n \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$ there exists a constant $C_1 > 0$ such that*

$$\mathbb{P}\left[Y_{n+1} = k+1 | Y_n^{aA} = k\right] \geq \frac{1}{2} - C_1 x^i \varepsilon \equiv p_{aA}^l. \tag{II.180}$$

Proof. The proof is completely analogous to the proof of Lemma II.8 and we skip the details. □

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

$$(1)\quad Z_0^l = Y_0, \tag{II.181}$$

$$(2)\quad \mathbb{P}[Z_{n+1}^l = k+1 | Y_n^{aA} > Z_n^l = k] = p_{aA}^l, \tag{II.182}$$

$$(3)\quad \mathbb{P}[Z_{n+1}^l = k-1 | Y_n^{aA} > Z_n^l = k] = 1 - p_{aA}^l, \tag{II.183}$$

$$(4)\quad \mathbb{P}[Z_{n+1}^l = k-1 | Y_{n+1} = k-1, Y_n^{aA} = Z_n^l = k] = 1, \tag{II.184}$$

$$(5)\quad \mathbb{P}[Z_{n+1}^l = k-1 | Y_{n+1} = k+1, Y_n^{aA} = Z_n^l = k] = \frac{\mathbb{P}[Y_{n+1}=k+1|Y_n^{aA}=k]-p_{aA}^l}{\mathbb{P}[Y_{n+1}=k+1|Y_n^{aA}=k]}, \tag{II.185}$$

$$(6)\quad \mathbb{P}[Z_{n+1}^l = k+1 | Y_{n+1} = k+1, Y_n^{aA} = Z_n^l = k] = 1 - \frac{\mathbb{P}[Y_{n+1}=k+1|Y_n^{aA}=k]-p_{aA}^l}{\mathbb{P}[Y_{n+1}=k+1|Y_n^{aA}=k]}. \tag{II.186}$$

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

$$\mathbb{P}[Z_{n+1}^l = k+1 | Z_n^l = k] = p_{aA}^l, \tag{II.187}$$

$$\mathbb{P}[Z_{n+1}^l = k-1 | Z_n^l = k] = 1 - p_{aA}^l. \tag{II.188}$$

Similar to the upper process, we can show that the lower process reaches $x^{i+1}\varepsilon K$ before returning to $x^i\varepsilon K + K^{1/2+3/4\alpha}$, with high probability. This concludes the proof of the proposition. □

Step 6: In this step we calculate the time which $n_{aA}(t)$ needs to decrease from $x^i\varepsilon$ to $x^{i+1}\varepsilon$ (Proposition II.8).

Proof of Proposition II.8. Let $Z_n^l \preccurlyeq Y_n^{aA} \preccurlyeq Z_n^u$ be defined as in the step before and $Y_0 = Z_0^l = Z_0^u = x^i\varepsilon K$.
Recalling (II.106), we get

$$\begin{aligned}\lambda_{aA}(t) &= fN_{aA}(t)\left(1 - \frac{n_{aA}(t)}{2\Sigma(t)}\right) + 2fN_{aa}(t)\frac{n_{AA}(t)}{\Sigma(t)} + N_{aA}(t)[D + c\Sigma(t)] \\ &\geq 2fx^{i+1}\varepsilon K + \mathcal{O}(x^{2i}\varepsilon^2 K) \equiv C_\lambda x^{i+1}\varepsilon K \equiv \lambda_{aA}^l,\end{aligned} \tag{II.189}$$

$$\begin{aligned}\lambda_{aA}(t) &= fN_{aA}(t)\left(1 - \frac{n_{aA}(t)}{\Sigma(t)}\right) + 2fN_{aa}(t)\frac{n_{AA}(t)}{\Sigma(t)} + N_{aA}(t)[D + c\Sigma(t)] \\ &\leq 2fx^{i}\varepsilon K + \mathcal{O}(x^{2i}\varepsilon^2 K) \equiv C_\lambda x^{i}\varepsilon K \equiv \lambda_{aA}^u.\end{aligned} \tag{II.190}$$

Let $n_* := \inf\left\{n \geq 0 : Y_n^{aA} - Y_0^{aA} \leq -(1-x)x^i\varepsilon K\right\}$ be the random variable which counts the number of jumps $Y_n^{aA} - Y_0^{aA}$ makes until it is smaller than $-(1-x)x^i\varepsilon K$. The time between two jumps of $n_{aA}(t)$ is given by $\tau_m - \tau_{m-1}$. It holds that $J_m^u \preceq \tau_m - \tau_{m-1} \preceq J_m^l$, where J_m^l (resp. J_m^u) are i.i.d. exponential distributed random variables with parameter λ_{aA}^l (resp. λ_{aA}^u). We want to estimate bounds for the times that the processes Z_n^u, resp.

Z_n^l, need to decrease from $x^i\varepsilon K$ to $x^{i+1}\varepsilon K$. Thus we show, for constants $C_u, C_l > 0$, that

$$(i) \quad \mathbb{P}\left[\sum_{m=1}^{n_*} J_m^u > \frac{2C_u}{C_\lambda x^{i+1}\varepsilon}\right] \leq \exp(-MK^{2\alpha}), \tag{II.191}$$

$$(ii) \quad \mathbb{P}\left[\sum_{m=1}^{n_*} J_m^l < \frac{C_l}{2C_\lambda x^i\varepsilon}\right] \leq \exp(-MK^{2\alpha}). \tag{II.192}$$

We start by showing (II.191). We need to find N such that, with high probability, $n_* \leq N$. To do this, we use the majorising process Z^u. Let W_k^u be i.i.d. random variables with

$$\mathbb{P}[W_k^u = 1] = \frac{1}{2} - C_0 x^{i+1}\varepsilon,\ \mathbb{P}[W_k^u = -1] = \frac{1}{2} + C_0 x^{i+1}\varepsilon \text{ and } \mathbb{E}[W_1^u] = -2C_0 x^{i+1}\varepsilon. \tag{II.193}$$

W_k^u records a birth or a death event of the process Z_n^u. From Lemma II.8 we get

$$\begin{aligned}
\mathbb{P}[n_* \leq N] &\geq \mathbb{P}\left[\exists n \leq N : \sum_{k=1}^{n} W_k \leq -\left\lfloor (1-x)x^i\varepsilon K\right\rfloor\right] \\
&\geq \mathbb{P}\left[\sum_{k=1}^{N} W_k \leq -\left\lfloor (1-x)x^i\varepsilon K\right\rfloor\right] \\
&\geq 1 - \mathbb{P}\left[\sum_{k=1}^{N}(W_k - \mathbb{E}W_k) \geq 2NC_0x^{i+1}\varepsilon - \left\lfloor (1-x)x^i\varepsilon K\right\rfloor\right].
\end{aligned} \tag{II.194}$$

By Hoeffding's inequality and choosing $N = \frac{1-x}{C_0 x}K =: C_u K$, we get

$$\mathbb{P}[n_* \leq C_u K] \geq 1 - \exp\left(-\frac{C_0\left(x^i\varepsilon\right)^2 Kx(1-x)}{2}\right) \geq 1 - \exp\left(-K^{2\alpha}C_0(1-x)x/2\right), \tag{II.195}$$

where we used that $x^i\varepsilon \geq K^{-1/2+\alpha}$. Thus

$$\mathbb{P}\left[\sum_{m=1}^{n_*} J_m^u > \frac{2C_u}{C_\lambda x^{i+1}\varepsilon}\right] \leq \mathbb{P}\left[\sum_{m=1}^{C_uK} J_m^u > \frac{2C_u}{C_\lambda x^{i+1}\varepsilon}\right] + \exp(-K^{2\alpha}C_0x(1-x)/2). \tag{II.196}$$

By applying the exponential Chebyshev inequality we get

$$\mathbb{P}\left[\sum_{m=1}^{C_uK} J_m^u > \frac{2C_u}{C_\lambda x^{i+1}\varepsilon}\right] \leq \exp\left(-C_uK/2\right). \tag{II.197}$$

Next we show (II.192). 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_k^l be i.i.d. random variables which record a birth or a death event of the process Z_n^l. They satisfy

$$\mathbb{P}[W_k^l = 1] = \frac{1}{2} - C_1x^i\varepsilon, \quad \mathbb{P}[W_k^l = -1] = \frac{1}{2} + C_1x^i\varepsilon \quad \text{and} \quad \mathbb{E}[W_1^l] = -2C_1x^i\varepsilon. \tag{II.198}$$

Note that from Lemma II.9

$$\mathbb{P}[n_* \le N] \le \mathbb{P}\left[\inf\left\{n \ge 0 : Z_n^l - Z_0^l \le -\left\lfloor (1-x)x^i\varepsilon K\right\rfloor\right\} \le N\right] \tag{II.199}$$

$$= \mathbb{P}\left[\exists n \le N : \sum_{k=1}^{n}(W_k - \mathbb{E}W_k) \le 2nC_1x^i\varepsilon - \left\lfloor (1-x)x^i\varepsilon K\right\rfloor\right]$$

$$\le \sum_{n=0}^{N}\mathbb{P}\left[\sum_{k=1}^{n}(W_k - \mathbb{E}W_k) \le 2nC_1x^i\varepsilon - \left\lfloor (1-x)x^i\varepsilon K\right\rfloor\right]. \tag{II.200}$$

If we choose $N = \frac{1-x}{4C_1}K =: C_lK$, using Hoeffding's inequality, we get, for all $n \le N$,

$$\mathbb{P}\left[\sum_{k=1}^{n}(W_k - \mathbb{E}W_k) \ge \left\lceil (1-x)x^i\varepsilon K\right\rceil - \frac{1}{2}(1-x)x^i\varepsilon K\right] \le \exp(-K^{2\alpha}C_1(1-x)/2). \tag{II.201}$$

Thus

$$\mathbb{P}\left[\sum_{m=1}^{n_*} J_m^l < \frac{C_l}{2C_\lambda x^i\varepsilon}\right] \le \mathbb{P}\left[\sum_{m=1}^{C_lK} J_m^l < \frac{C_l}{2C_\lambda x^i\varepsilon}\right] + C_lK\exp\left(-K^{2\alpha}C_1(1-x)/2\right). \tag{II.202}$$

It holds that

$$\mathbb{P}\left[\sum_{m=1}^{C_lK} J_m^l > \frac{C_l}{2C_\lambda x^i\varepsilon}\right] = 1 - \mathbb{P}\left[\sum_{m=1}^{C_lK} J_m^l < \frac{C_l}{2C_\lambda x^i\varepsilon}\right] = 1 - \mathbb{P}\left[-\sum_{m=1}^{C_lK} J_m^l > -\frac{C_l}{2C_\lambda x^i\varepsilon}\right]. \tag{II.203}$$

A simple use of the exponential Chebyshev inequality shows that

$$\mathbb{P}\left[\sum_{m=1}^{C_lK} J_m^l < \frac{C_l}{2C_\lambda x^i\varepsilon}\right] \le \exp\left(-C_lK/2\right). \tag{II.204}$$

Thus we have that $\mathbb{P}\left[\sum_{m=1}^{n_*} J_m^l < \frac{C_l}{2C_\lambda x^i\varepsilon}\right] \le \exp\left(-MK^{2\alpha}\right)$, for some constant $M > 0$. □

Step 7: In this step it is shown that $n_{aa}(t)$ decreases under the upper bound $\gamma_\Delta x^{2i+2}\varepsilon^2$, which we need to proceed the next iteration step, in at least the time $n_{aA}(t)$ needs to decrease from $x^i\varepsilon$ to $x^{i+1}\varepsilon$ (Proposition II.9). We set the time to zero when $n_{aA}(t)$ hits $x^i\varepsilon$. Hence $n_{aa}(0) \le \gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}$. Remember, $\gamma_{\Delta/2} = \frac{f+\frac{\Delta}{2}(1-\vartheta)}{4\bar{n}_A(f+\Delta)}$ and let

$$\theta_i^+(aa) \equiv \inf\left\{t \ge 0 : n_{aa}(t) \ge \gamma_\Delta x^{2i}\varepsilon^2 + 3M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}\right\}, \tag{II.205}$$

$$\theta_i^-(aa) \equiv \inf\left\{t \ge 0 : n_{aa}(t) \le \gamma_{\Delta/2}x^{2i+2}\varepsilon^2\right\}. \tag{II.206}$$

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 $\gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}$ to $\gamma_{\Delta/2}x^{2i+2}\varepsilon^2$. As result we will get that $n_{aa}(t)$ reaches the next upper bound before the aA population decreases to $x^{i+1}\varepsilon$. Thus in the third part we ensure that $n_{aa}(t)$ stays below the bound $\gamma_\Delta x^{2i+2}\varepsilon^2$ until $n_{aA}(t)$ reaches $x^{i+1}\varepsilon$.

Part 1: We show

Proposition II.13. *For* $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^\alpha/2}\right]$ *there are constants* $\bar{C}, \tilde{C} > 0$ *such that*

$$\mathbb{P}\left[\theta_i^+(aa) < \theta_i^-(aa) | n_{aa}(0) \leq \gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}\right] \leq \exp(-\tilde{C}K^\alpha). \tag{II.207}$$

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 $\gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}$ to $\gamma_{\Delta/2}x^{2i+2}\varepsilon^2$ with high probability.

Proof. As before, let Y_n^{aa} be the discretisation of $N_{aa}(t)$. We start with the construction of the upper process.

Lemma II.10. *For* $n \in \mathbb{N}$ *such that* $\vartheta_n \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^\alpha/2}\right]$ *there exists a constant* $C_u > 0$ *such that*

$$\mathbb{P}\left[Y_{n+1}^{aa} = k+1 | Y_n^{aa} = k\right] \leq \frac{1}{2} - C_u \equiv p_u. \tag{II.208}$$

Proof. We know that, for $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^\alpha/2}\right]$, $n_{aA}(t) \in \left[x^{i+1}\varepsilon, x^i\varepsilon + K^{-1/2+3/4\alpha}\right]$. Again with (II.105) we have

$$\begin{aligned}\mathbb{P}\left[Y_{n+1}^{aa} = k+1 | Y_n = k\right] &= \frac{fk\left(1 - \frac{n_{AA}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}n_{aA}^2(t)}{fk\left(1 - \frac{n_{AA}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}n_{aA}^2(t) + k[D + \Delta + c\Sigma(t)]} \\ &\leq \frac{1}{2} + \frac{\frac{1}{2}f - \frac{f}{2\Sigma(t)}n_{AA}(t) + \frac{fK}{8\Sigma(t)}\frac{n_{aA}^2(t)}{k} - \frac{1}{2}D - \frac{1}{2}c\Sigma(t) - \frac{1}{2}\Delta}{f\left(1 - \frac{n_{AA}(t)}{\Sigma(t)}\right) + \frac{fK}{4\Sigma(t)}\frac{n_{aA}^2(t)}{k} + [D + \Delta + c\Sigma(t)]}.\end{aligned} \tag{II.209}$$

Using Proposition II.2, II.4 and II.6, we bound the numerator in the second summand from above by

$$\begin{aligned}&-\frac{f+\Delta}{2} + \frac{f}{8\bar{n}_A\gamma_{\Delta/2}x^2} + \mathcal{O}\left(x^i\varepsilon^1 + K^{-1/4\alpha}\right) \\ &\qquad \leq -\frac{(f+\Delta)\left(f\left(\vartheta - \frac{\Delta}{2} - \frac{\Delta\vartheta}{2}\right) + \frac{\Delta\vartheta}{2}(1-\vartheta)\right)}{2(f + \frac{\Delta}{2}(1-\vartheta))(f+\vartheta)} + \mathcal{O}\left(x^i\varepsilon + K^{-1/4\alpha}\right).\end{aligned} \tag{II.210}$$

Since $\frac{\Delta}{2} < \vartheta$ there exists a constant $C_u > 0$ such that

$$\mathbb{P}[Y_{n+1} = k+1 | Y_n = k] \leq \frac{1}{2} - C_u \equiv p_u. \tag{II.211}$$

□

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 $Z_n^u \succcurlyeq Y_n^{aa}$, a.s.. The marginal distribution of Z_n^u is the desired Markov chain with transition probabilities

$$\mathbb{P}[Z_{n+1}^u = k+1 | Z_n^u = k] = p_u, \tag{II.212}$$
$$\mathbb{P}[Z_{n+1}^u = k-1 | Z_n^u = k] = 1-p_u, \tag{II.213}$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1}\left(\frac{1}{2}-C_u\right)}{\prod_{k=1}^{n}\left(\frac{1}{2}+C_u\right)} = \frac{\left(\frac{1}{2}-C_u\right)^{n-1}}{\left(\frac{1}{2}+C_u\right)^{n}}. \tag{II.214}$$

We define the stopping times

$$T_i^+(aa) \equiv \inf\left\{t \geq 0 : Z_n^u \geq \gamma_\Delta x^{2i}\varepsilon^2 K + 3M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}K\right\}, \tag{II.215}$$
$$T_i^-(aa) \equiv \inf\left\{t \geq 0 : Z_n^u \leq \gamma_{\Delta/2} x^{2i+2}\varepsilon^2 K\right\}. \tag{II.216}$$

Again with the formula of the equilibrium potential we estimate for $\gamma_{\Delta/2}x^{2i+2}\varepsilon^2 K \leq zK < \gamma_\Delta x^{2i}\varepsilon^2 K + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}K$

$$\mathbb{P}_{zK}\left[T_i^+(aa) < T_i^-(aa)\right] = \frac{\sum_{n=\gamma_{\Delta/2}x^{2i+2}\varepsilon^2 K+1}^{zK}\left(\frac{1+2C_u}{1-2C_u}\right)^{n-1}}{\sum_{n=\gamma_{\Delta/2}x^{2i+2}\varepsilon^2 K+1}^{\gamma_\Delta x^{2i}\varepsilon^2 K+3M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}K}\left(\frac{1+2C_u}{1-2C_u}\right)^{n-1}} \leq \exp(-\tilde{C}K^\alpha). \tag{II.217}$$

□

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

Proposition II.14. *Let*

$$\theta_i(aa) \equiv \inf\left\{t \geq 0 : n_{aa}(t) \leq \gamma_{\Delta/2}x^{2i+2}\varepsilon^2 | n_{aa}(0) = \gamma_\Delta x^{2i}\varepsilon^2 + M_{aa}(x^{2i}\varepsilon^2)^{1+\alpha}\right\}, \tag{II.218}$$

be the decay time of $n_{aa}(t)$, for $t \in \left[\tau_{aA}^{i-}, \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$. Then there exist finite, positive constants C_u^{aa} and M, such that for all $0 \leq i \leq \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)} - 1$

$$\mathbb{P}\left[\theta_i(aa) > C_u^{aa}\right] \leq \exp\left(-MK^{2\alpha}\right). \tag{II.219}$$

Proof. The proof works like the one of Proposition II.8. We calculate an upper bound on the decay time of the majorising process Z_n^u. Let Y_n^{aa} and Z_n^u be defined as in the step before and let W^u be i.i.d. random variables with

$$\mathbb{P}[W^u = 1] = \frac{1}{2} - C_u, \quad \mathbb{P}[W^u = -1] = \frac{1}{2} + C_u \quad \text{and} \quad \mathbb{E}[W_1^u] = -2C_u. \tag{II.220}$$

The W^u's record a birth or a death event of Z_n^u. Similar as in Step 6, we choose $N = \frac{\gamma_\Delta - x^2\gamma_{\Delta/2} + M_{aa}(x^i\varepsilon)^{2\alpha}}{C_u}x^{2i}\varepsilon^2 K =: \frac{\bar{C}}{C_u}x^{2i}\varepsilon^2 K$ and show that $\mathbb{P}[n_* \leq \frac{\bar{C}}{C_u}x^{2i}\varepsilon^2 K] \geq 1 - \exp(-K^{2\alpha}\bar{C}C_u/2)$. It holds

$$\begin{aligned}\lambda_{aa}(t) &= fN_{aa}(t)\left(1 - \tfrac{n_{AA}(t)}{\Sigma(t)}\right) + \tfrac{fK}{4\Sigma(t)}n_{aA}^2(t) + N_{aa}(t)[D + \Delta + c\Sigma(t)] \\ &\leq C_\lambda x^{2i+2}\varepsilon^2 K \equiv \lambda_{aa}^l.\end{aligned} \tag{II.221}$$

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 λ_{aa}^l. As in Step 6, using the exponential Chebyshev inequality, we get that

$$\mathbb{P}\left[\sum_{m=1}^{\bar{C}/C_u x^{2i}\varepsilon^2 K} J_m^{aa} > \frac{2\bar{C}}{C_u C_\lambda x^2}\right] \leq \exp\left(-\tfrac{\bar{C}}{2C_u}K^{2\alpha}\right), \tag{II.222}$$

and hence

$$\mathbb{P}\left[\sum_{m=1}^{n_*} J_m^{aa} > \frac{2\bar{C}}{C_u C_\lambda x^2}\right] \leq \exp\left(-MK^{2a}\right). \tag{II.223}$$

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

Step 8: For the iteration we have to ensure that the sum process increases on the given time interval from the value $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}$ to $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ and stays there until the next iteration-step (Proposition II.10). We set the time to zero when the aA population hits $x^i\varepsilon$. Hence $\Sigma(0) \geq \bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}$. Observe that we have to carry out this step only until $i \leq \frac{-\ln\left(\varepsilon K^{1/4-\alpha/4}\right)}{\ln(x)}$, since afterwards $x^{2i}\varepsilon^2 \leq K^{-1/2+\alpha/2}$ and thus $\Sigma(t) \geq \bar{n}_A - M_\Sigma K^{-1/2+\alpha/2}$ which is enough for the further iteration. As in the proof of Proposition II.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 $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ before going back to $\bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - 3M_\Sigma K^{-1/2+\alpha/2}$. We define stopping times on $\Sigma(t)$:

$$\tau_\Sigma^{i-} \equiv \inf\left\{t \geq 0 : \Sigma(t) \leq \bar{n}_A - \tfrac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - 3M_\Sigma K^{-1/2+\alpha/2} \equiv v(\Sigma_-) - 3M_\Sigma K^{-1/2+\alpha/2}\right\}, \tag{II.224}$$

$$\tau_\Sigma^{i+} \equiv \inf\left\{t \geq 0 : \Sigma(t) \geq \bar{n}_A - \tfrac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2 \equiv v(\Sigma_+)\right\}, \tag{II.225}$$

where $M_\Sigma > 0$.

Proposition II.15. *There are constants $\bar{C}, \tilde{C} > 0$ such that for all $0 \leq i \leq \frac{-\ln\left(\varepsilon K^{1/4-\alpha/4}\right)}{\ln(x)}$*

$$\mathbb{P}[\tau_\Sigma^{i-} < \tau_\Sigma^{i+} | \Sigma(0) \geq \bar{n}_A - \tfrac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 + M_\Sigma(x^{2i}\varepsilon^2)^{1+\alpha}] \leq \exp\left(-\tilde{C}K^\alpha\right). \tag{II.226}$$

Proof. From the step before we know that $n_{aa}(t)$ decreases under the value $\gamma_\Delta x^{2i+2}\varepsilon^2$ 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 $t \in \left[\tau_{aA}^{i-} + \theta_i(aa), \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-}\right]$, the sum process has the tendency to increase and exceed the lower bound $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ before $n_{aA}(t)$ hits $x^{i+1}\varepsilon$. As before, let Y_n^Σ be the associated discrete process to $\Sigma(t)$.

Lemma II.11. *For $n \in \mathbb{N}$ such that $\vartheta_n \in \left[\tau_{aA}^{i-} + \theta_i(aa), \tau_{aA}^{i+} \wedge \tau_{aA}^{(i+1)-} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$ there exists a constant $C_0 > 0$ such that for all $0 \le i \le \frac{-\ln\left(\varepsilon K^{1/4-\alpha/4}\right)}{\ln(x)}$*

$$\mathbb{P}\left[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k\right] \le \frac{1}{2} - C_0 x^{2i+2}\varepsilon^2 \equiv p_0(\Sigma). \tag{II.227}$$

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

$$\mathbb{P}\left[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k\right] \le \frac{1}{2} + \frac{-\frac{1}{2}(f-D) + \frac{1}{2}ckK^{-1} + \frac{\Delta}{2k}N_{aa}(t)}{f + D + ckK^{-1} + \frac{\Delta}{k}N_{aa}(t)}. \tag{II.228}$$

We estimate the nominator

$$\le -\frac{f-D}{2} + \frac{1}{2}ckK^{-1} + \frac{\Delta}{2}\gamma_\Delta x^{2i+2}\varepsilon^2 k^{-1}K. \tag{II.229}$$

This term assumes its maximum at $k = v(\Sigma_+)$. If we insert this bound, $\bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$, we can estimate

$$\begin{aligned} &\le -\frac{f-D}{2} + \frac{c}{2}\bar{n}_A - \frac{\Delta+\vartheta/2}{2\bar{n}_A}\gamma x^{2i+2}\varepsilon^2 + \frac{\Delta}{2\bar{n}_A}\gamma_\Delta x^{2i+2}\varepsilon^2 + \mathcal{O}(x^{4i+4}\varepsilon^4) \\ &\le -\frac{\vartheta f - \Delta^2}{16\bar{n}_A^2(f+\Delta)}x^{2i+2}\varepsilon^2 + \mathcal{O}(x^{4i+4}\varepsilon^4). \end{aligned} \tag{II.230}$$

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

$$\mathbb{P}\left[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k\right] \le \frac{1}{2} - C_0 x^{2i+2}\varepsilon^2. \tag{II.231}$$

□

Again we couple Y_n^{Σ} with Z_n^l via:

(1) $Z_0^l = Y_0^{\Sigma}$, (II.232)

(2) $\mathbb{P}[Z_{n+1}^l = k+1 | Y_n^{\Sigma} > Z_n^l = k] = 1 - p_0(\Sigma)$, (II.233)

(3) $\mathbb{P}[Z_{n+1}^l = k-1 | Y_n^{\Sigma} > Z_n^l = k] = p_0(\Sigma)$, (II.234)

(4) $\mathbb{P}[Z_{n+1}^l = k-1 | Y_{n+1}^{\Sigma} = k-1, Y_n^{\Sigma} = Z_n^l = k] = 1$, (II.235)

(5) $\mathbb{P}[Z_{n+1}^l = k-1 | Y_{n+1}^{\Sigma} = k+1, Y_n^{\Sigma} = Z_n^l = k] = \frac{p_0(\Sigma) - \mathbb{P}[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k]}{1 - \mathbb{P}[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k]}$, (II.236)

(6) $\mathbb{P}[Z_{n+1}^l = k+1 | Y_{n+1}^{\Sigma} = k+1, Y_n^{\Sigma} = Z_n^l = k] = 1 - \frac{p_0(\Sigma) - \mathbb{P}[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k]}{1 - \mathbb{P}[Y_{n+1}^{\Sigma} = k-1 | Y_n^{\Sigma} = k]}$. (II.237)

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

$$\mathbb{P}\left[Z_{n+1}^l = k+1 | Z_n^l = k\right] = 1 - p_0(\Sigma), \tag{II.238}$$

$$\mathbb{P}\left[Z_{n+1}^l = k-1 | Z_n^l = k\right] = p_0(\Sigma), \tag{II.239}$$

and invariant measure

$$\mu(n) = \frac{\prod_{k=1}^{n-1}\left(\frac{1}{2} + C_0 x^{2i+2}\varepsilon^2\right)}{\prod_{k=1}^{n}\left(\frac{1}{2} - C_0 x^{2i+2}\varepsilon^2\right)} = \frac{\left(\frac{1}{2} + C_0 x^{2i+2}\varepsilon^2\right)^{n-1}}{\left(\frac{1}{2} - C_0 x^{2i+2}\varepsilon^2\right)^{n}}. \tag{II.240}$$

Again we get a bound on the harmonic function, for $\bar{n}_A - \frac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2} \leq z < \bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma_x^{2i+2}\varepsilon^2$,

$$\mathbb{P}_{zK}\left[T_\Sigma^{i-} < T_\Sigma^{i+}\right] = \frac{\sum_{n=zK+1}^{Kv(\Sigma_+)}\left(\frac{1-2C_0x^{2i+2}\varepsilon^2}{1+2C_0x^{2i+2}\varepsilon^2}\right)^{n-1}}{\sum_{n=Kv(\Sigma_-)-3M_\Sigma K^{1/2+\alpha/2}+1}^{Kv(\Sigma_+)}\left(\frac{1-2C_0x^{2i+2}\varepsilon^2}{1+2C_0x^{2i+2}\varepsilon^2}\right)^{n-1}} \leq \exp\left(-\tilde{C}K^\alpha\right). \tag{II.241}$$

□

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

Proposition II.16. *Let*

$$\theta_i(\Sigma) \equiv \inf\Big\{t \geq 0 : \Sigma(t) \geq \bar{n}_A - \tfrac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2\Big| \Sigma(0) = \bar{n}_A - \tfrac{\Delta+\vartheta}{c\bar{n}_A}\gamma x^{2i}\varepsilon^2 - M_\Sigma K^{-1/2+\alpha/2}\Big\}, \tag{II.242}$$

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

$$\mathbb{P}\left[\theta_i(\Sigma) > C_l^\Sigma\right] \leq \exp\left(-MK^{4\alpha}\right). \tag{II.243}$$

Proof. Let Y_n^Σ be defined as in the step before and let W_k be i.i.d. random variables with

$$\mathbb{P}[W_k = 1] = \frac{1}{2} + C_0 x^{2i+2}\varepsilon^2,\ \mathbb{P}[W_k = -1] = \frac{1}{2} - C_0 x^{2i+2}\varepsilon^2 \text{ and } \mathbb{E}[W_1] = 2C_0 x^{2i+2}\varepsilon^2, \tag{II.244}$$

which record a birth or a death event of the lower process Z_n^l. We choose $N = \frac{\tilde{C}K}{4C_0x^2}$, with $\tilde{C} = \frac{\Delta+\vartheta-x^2(\Delta+\vartheta/2)}{c\bar{n}_A}$, and show that $\mathbb{P}\left[n_* \leq N\right] \geq 1 - \exp\left(-K^\alpha C_0\tilde{C}x^2/2\right)$. We estimate from above the time the process $\Sigma(t)$ needs to make one jump:

$$\lambda_\Sigma(t) = f\Sigma(t)K + \Sigma(t)K[D + \Delta + c\Sigma(t)] + \Delta N_{aa}(t) \geq C_\lambda K \equiv \lambda_\Sigma^l. \tag{II.245}$$

As before let $\tau_m - \tau_{m-1}$ the time between two jumps of $\Sigma(t)$ and let J_m^Σ are i.i.d. exponential random variables with parameter λ_Σ^l. As in Step 6, by applying the exponential Chebyshev inequality, we get that $\mathbb{P}\left[\sum_{m=1}^{\tilde{C}K} J_m^\Sigma > \frac{\tilde{C}}{2C_0x^2C_\lambda}\right] \leq \exp\left(-\frac{\tilde{C}}{2C_0x^2}K\right)$ and hence

$$\mathbb{P}\left[\sum_{m=1}^{n_*} J_m^\Sigma > \frac{\tilde{C}}{2C_0x^2C_\lambda}\right] \leq \exp\left(-MK^\alpha\right). \tag{II.246}$$

□

Part 3: For the iteration we have to ensure that once $\Sigma(t)$ hits the upper bound $\bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ it will stay close to it for an exponential time on the given time interval. This ensures that the sum process is larger than $\bar{n}_A - \frac{\vartheta+\Delta}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$ when $n_{aA}(t)$ hits $x^{i+1}\varepsilon$ and the next iteration step can start.

Proposition II.17. *Assume that* $\Sigma(0) = \bar{n}_A - \frac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2$. *Let*

$$\hat{\tau}^{\alpha}_{\Sigma} \equiv \inf\left\{t > 0 : \Sigma(t) - \bar{n}_A + \tfrac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2 \leq -\tfrac{\vartheta}{2c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2\right\}. \tag{II.247}$$

Then, for all $0 \leq i \leq \frac{-\ln\left(\varepsilon K^{1/4-\alpha/4}\right)}{\ln(x)}$

$$\mathbb{P}[\hat{\tau}^{\alpha}_{\Sigma} < \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}}] = o(K^{-\alpha}). \tag{II.248}$$

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

$$\begin{aligned}
\hat{X}_t &= \Sigma(t)K - \bar{n}_A K + \tfrac{\Delta+\vartheta/2}{c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2 K, && \text{(II.249)}\\
\hat{T}^{X}_{0} &\equiv \inf\{t \geq 0 : \hat{X}_t = 0\}, && \text{(II.250)}\\
\hat{T}^{X}_{\alpha,M_\Sigma} &\equiv \{t \geq 0 : \hat{X}_t \leq -\tfrac{\vartheta}{2c\bar{n}_A}\gamma x^{2i+2}\varepsilon^2 K\}. && \text{(II.251)}
\end{aligned}$$

Let $\hat{Y}_n$ be the discrete process associated to $\hat{X}_t$, obtained as described in Step 1.

Lemma II.12. *For* $n \in \mathbb{N}$ *such that* $\vartheta_n \in \left[\tau^{i-}_{aA} + \theta_i(aa), \tau^{i+}_{aA} \wedge \tau^{(i+1)-}_{aA} \wedge \mathrm{e}^{VK^{\alpha/2}}\right]$, *there exists a constant* $\hat{C}_0 > 0$ *such that for all* $0 \leq i \leq \frac{-\ln\left(\varepsilon K^{1/4-\alpha/4}\right)}{\ln(x)}$

$$\mathbb{P}[\hat{Y}_{n+1} = -k-1 | \hat{Y}_n = -k] \leq \frac{1}{2} - \hat{C}_0 x^{2i+2}\varepsilon^2 \equiv \hat{p}_\Sigma. \tag{II.252}$$

The proof is a re-run of Step 3 by using the rates (II.104). The rest of the proof of Proposition II.17 is similar to Step 3. □

Final Step:

Calculation of the Decay Time of n_{aA}: The following proves Theorem II.2 (ii). Set

$$\sigma \equiv \frac{-\ln(\varepsilon K^{1/2-\alpha})}{\ln(x)}. \tag{II.253}$$

Observe that $x^\sigma\varepsilon \geq K^{-1/2+\alpha}$ 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 $i = \sigma - 1$. Observe that $\sum_{j=0}^{\sigma-1}\frac{C_u}{x^j\varepsilon} = \frac{C_u x}{1-x}(K^{1/2-\alpha} - \varepsilon^{-1}) \geq \frac{C_l x}{1-x}(K^{1/2-\alpha} -$

$\varepsilon^{-1}) = \sum_{j=0}^{\sigma-1} \frac{C_l}{x^j \varepsilon}$ 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

$$\mathbb{P}\left[\frac{C_u x}{1-x}\left(K^{1/2-\alpha} - \varepsilon^{-1}\right) \geq \tau_{aA}^{\sigma-} - \tau_{aA}^{0+} \geq \frac{C_l x}{1-x}\left(K^{1/2-\alpha} - \varepsilon^{-1}\right)\right] \tag{II.254}$$

$$= \mathbb{P}\left[\frac{C_u x}{1-x}\left(K^{1/2-\alpha} - \varepsilon^{-1}\right) \geq \sum_{j=0}^{\sigma-1}\left(\tau_{aA}^{(j+1)-} - \tau_{aA}^{j+}\right) \geq \frac{C_l x}{1-x}\left(K^{1/2-\alpha} - \varepsilon^{-1}\right)\right] \tag{II.255}$$

$$\geq \mathbb{P}\left[\frac{C_u}{x\varepsilon} \geq \theta_1(aA) \geq \frac{C_l}{x\varepsilon}, ..., \frac{C_u}{x^\sigma \varepsilon} \geq \theta_\sigma(aA) \geq \frac{C_l}{x^\sigma \varepsilon}\right] \tag{II.256}$$

$$\geq 1 - \sigma \exp(-MK^{2\alpha}) \geq 1 - \exp(-\tilde{M}K^{2\alpha}). \tag{II.257}$$

Remark II.5. *A function $f(t) = \frac{1}{t}$ needs a time $t_1 - t_0 = K^{1/2-\alpha} - \frac{1}{\varepsilon}$ to decrease from $f(t_0) = \varepsilon$ to $f(t_1) = K^{-1/2+\alpha}$. 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 II.3. We already know that there is no mutation before a time of order $\frac{1}{K\mu_K}$ (cf. Lemma II.1 and (II.23)). Since we have seen that the duration of the first step is $\mathcal{O}(\ln K)$ and the time needed for the second step is bounded, the left hand side of (II.15) ensures that the first two phases are ended before the occurrence of a new mutation. Thus we get the first statement of (II.21).
To justify the second statement of (II.21), 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 $K^{-1/2+\alpha}$. Thus we have to verify that the mutation time is smaller than $\mathcal{O}\left(K^{1/2-\alpha}\right)$, the time the process n_{aA} needs to decrease to $K^{-1/2+\alpha}$.
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

$$p_A(t) = \frac{n_{AA}(t) + \frac{1}{2}n_{aA}(t)}{n_{aa}(t) + n_{aA}(t) + n_{AA}(t)} \tag{II.258}$$

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

$$\mu_K^{AA} = \mu_K f p_A(t) n_{AA}(t) K. \tag{II.259}$$

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

$$\mu_K^{AA} \geq \mu_K f(\bar{n}_A - \mathcal{O}(\varepsilon))K. \tag{II.260}$$

From (II.259) we get that the time of a new mutation is smaller or equal to $\frac{1}{f(n_A - \mathcal{O}(\varepsilon)K\mu_K}$. Thus to ensure that we still have a alleles in the population, we have to choose μ_K in such a way that

$$\frac{1}{K\mu_K} \ll K^{1/2-\alpha}, \tag{II.261}$$

since we can ensure the survival of $n_{aA}(t)$ until the time $\frac{C_l x}{1-x}\left(K^{1/2-\alpha}-\varepsilon^{-1}\right)$ (cf. (II.254)). Hence, the right hand side of (II.15) gives us that a new mutation occurs before the aA population died out. This finishes the proof of Theorem II.3.

III. Chapter

The recovery of a recessive allele in a Mendelian diploid model

Anton Bovier, Loren Coquille and Rebecca Neukirch

Abstract We study the large population limit of a stochastic individual-based model which describes the time evolution of a diploid hermaphroditic population reproducing according to Mendelian rules. In [84] it is proved that sexual reproduction allows unfit alleles to survive in individuals with mixed genotype much longer than they would in populations reproducing asexually. In the present paper we prove that this indeed opens the possibility that individuals with a pure genotype can reinvade in the population after the appearance of further mutations. We thus expose a rigorous description of a mechanism by which a recessive allele can re-emerge in a population. This can be seen as a statement of genetic robustness exhibited by diploid populations performing sexual reproduction.

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*. L.C. has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) through the Exploratory Project *CanDyPop* and by the Swiss National Science Foundation through the grant No. P300P2_161031.
We would like to thank Pierre Collet and Vincent Beffara for their help on the theory of dynamical systems and fruitful discussions.

1. Introduction

In *population genetics*, the study of Mendelian diploid models of fixed population size began more than a century ago (see e.g. [12, 23, 32, 34, 45, 46, 82, 103, 107]), while their counterparts of variable population size models were studied in the context of adaptive dynamics from 1999 onwards [59]. The approach of adaptive dynamics is to introduce competition kernels to regulate the population size instead of maintaining it constant, see [52, 72, 76].

Stochastic individual-based versions of these models appeared in the 1990s, see [14–17, 28, 36]. They assume single events of reproduction, mutation, natural death, and death by competition happen at random times to each individual in the population. An important and interesting feature of these models is that different limiting processes on different time-scales appear as the

carrying capacity tends to infinity while mutation rates and mutation step-size tend to zero (see [2, 14, 17, 28, 75]). One of the major results in this context is the convergence of a properly rescaled process to the so called *Trait Substitution Sequence* (TSS) process, which describes the evolution of a monomorphic population as a jump process between monomorphic equilibria. More generally, Champagnat and Méléard [17] obtained the convergence to a *Polymorphic Evolution Sequence* (PES), where jumps occur between equilibria that may include populations that have multiple co-existing phenotypes. The appearance of co-existing phenotypes is, however, exceptional and happens only at so-called *evolutionary singularities*. From a biological point of view, this is somewhat unsatisfactory, as it apparently fails to explain the biodiversity seen in real biological systems.

Most of the models considered in this context assume haploid populations with a-sexual reproduction. Exceptions are the paper [18] by Collet, Méléard and Metz from 2013 and a series of papers by Coron and co-authors [19–21] following it. In [18], the *Trait Substitution Sequence* is derived in a Mendelian diploid model under the assumption that the fitter mutant allele and the resident allele are co-dominant.

The main reason why both in haploid models and in the model considered in [18] the evolution along monomorphic populations is typical is that the time scales for the fixation of a new trait and the extinction of the resident trait are the same (both of order $\ln K$) (unless some very special fine-tuning of parameters occurs that allows for co-existence). This precludes (at least in the rare mutation scenarios considered) that an initially less fit trait survives long enough until after possibly several new mutations occurred that might create a situation where this trait may become fit again and recover.

In a follow-up paper to [18], two of the present authors [84], it was shown that, if instead one assumes that the resident allele is recessive, the time to extinction of this allele is dramatically increased. This will be discussed in detail in Section 1.2 and paves the way for the appearance of a richer limiting process.

The general framework in [18] and [84] is the following. Each individual is characterised by a reproduction and death rate which depend on a phenotypic trait determined by its genotype, which here is determined by two alleles (e.g. A and a) on one single locus. The evolution of the trait distribution of the three genotypes aa, aA and AA is studied under the action of (1) heredity, which transmits traits to new offsprings according to Mendelian rules, (2) mutation, which produces variations in the trait values in the population onto which selection is acting, and (3) of competition for resources between individuals.

The paper [84] proves that sexual reproduction allows unfit alleles to survive in individuals with mixed genotype much longer than they would in populations reproducing asexually. This opens the possibility that while this allele is still alive in the population, the appearance of new mutants alters the fitness landscape in such a way that is favourable for this allele and allow it to reinvade in the population, leading to a new equilibrium with co-existing phenotypes. The goal of this paper is to rigorously prove that such a scenario indeed occurs under fairly natural assumptions. Recently, Billiard and Smadi [7] considered related questions for haploid individuals (performing clonal reproduction). The authors show that a deleterious allele can reinvade after a new mutation, but the range of parameters allowing this behaviour is though very small.

1.1. The stochastic model

The individual-based microscopic Mendelian diploid model is a non-linear birth-and-death process. We consider a model for a population of a finite number of hermaphroditic individuals which

reproduce sexually. Each individual i is characterised by two alleles, $u_1^i u_2^i$, taken from some allele space $\mathcal{U} \subset \mathbb{R}$. These two alleles define the genotype of the individual i. We suppress parental effects, which means that we identify individuals with genotype u_1u_2 and u_2u_1. Each individual has a Mendelian reproduction rate with possible mutations and a natural death rate. Moreover, there is an additional death rate due to ecological competition with the other individuals in the population. Let

$f_{u_1u_2} \in \mathbb{R}_+$	the per capita birth rate (fertility) of an individual with genotype u_1u_2,
$D_{u_1u_2} \in \mathbb{R}_+$	the per capita natural death rate of an individual with genotype u_1u_2,
$K \in \mathbb{N}$	the carrying capacity, a parameter which scales the population size,
$\frac{c_{u_1u_2,v_1v_2}}{K} \in \mathbb{R}_+$	the competition effect felt by an individual with genotype u_1u_2 from an individual of genotype v_1v_2,
$R_{u_1u_2}(v_1v_2) \in \{0,1\}$	the reproductive compatibility of the genotype v_1v_2 with u_1u_2,
$\mu_K \in \mathbb{R}_+$	the mutation probability per birth event. Here it is independent of the genotype,
$m(u, dh)$	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. We are interested in asymptotic results when K is large. We assume rare mutation, i.e. $\mu_K \ll 1$. If a mutation occurs at a birth event, only one allele changes from u to $u + h$ where h is a random variable with law $m(u, dh)$.

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

$$\nu_t = \frac{1}{K} \sum_{i=1}^{N_t} \delta_{u_1^i(t)u_2^i(t)}. \tag{III.1}$$

Formally, ν_t takes values in the set of re-scaled point measures

$$\mathcal{M}^K = \left\{ \frac{1}{K} \sum_{i=1}^{n} \delta_{u_1^i u_2^i} \;\middle|\; n \geq 0, u_1^1 u_2^1, ..., u_1^n u_2^n \in \mathcal{U}^2 \right\}, \tag{III.2}$$

on $\mathcal{U}^2$, equipped with the vague topology. Define $\langle \nu, g \rangle$ as the integral of the measurable function $g : \mathcal{U}^2 \to \mathbb{R}$ with respect to the measure $\nu \in \mathcal{M}^K$. Then $\langle \nu_t, \mathbb{1} \rangle = \frac{N_t}{K}$ and for any $u_1u_2 \in \mathcal{U}^2$, the positive number $\langle \nu_t, \mathbb{1}_{u_1u_2} \rangle$ is called the *density* at time t of the genotype u_1u_2. The generator of the process is defined as in [18]: first we define, for the genotypes u_1u_2, v_1v_2 and a point measure ν, the Mendelian reproduction operator:

$$\begin{aligned}
&(A_{u_1u_2,v_1v_2}F)(\nu) \\
&\quad = \frac{1}{4}\left[F\left(\nu + \frac{\delta_{u_1v_1}}{K}\right) + F\left(\nu + \frac{\delta_{u_1v_2}}{K}\right) + F\left(\nu + \frac{\delta_{u_2v_1}}{K}\right) + F\left(\nu + \frac{\delta_{u_2v_2}}{K}\right)\right] - F(\nu),
\end{aligned} \tag{III.3}$$

and the Mendelian reproduction-cum-mutation operator:

$$\begin{aligned}(M_{u_1u_2,v_1v_2}F)(\nu) =& \frac{1}{8}\int_{\mathbb{R}}\Bigg[\left(F\left(\nu+\frac{\delta_{u_1+h,v_1}}{K}\right)+F\left(\nu+\frac{\delta_{u_1+h,v_2}}{K}\right)\right)m(u_1,h)\\ &+\left(F\left(\nu+\frac{\delta_{u_2+h,v_1}}{K}\right)+F\left(\nu+\frac{\delta_{u_2+h,v_2}}{K}\right)\right)m(u_2,h)\\ &+\left(F\left(\nu+\frac{\delta_{u_1,v_1+h}}{K}\right)+F\left(\nu+\frac{\delta_{u_2,v_1+h}}{K}\right)\right)m(v_1,h)\\ &+\left(F\left(\nu+\frac{\delta_{u_1,v_2+h}}{K}\right)+F\left(\nu+\frac{\delta_{u_2,v_2+h}}{K}\right)\right)m(v_2,h)\Bigg]dh-F(\nu).\end{aligned} \tag{III.4}$$

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

$$\begin{aligned}(L^KF)(\nu) =& \int_{\mathcal{U}^2}\left(D_{u_1u_2}+\int_{\mathcal{U}^2}c_{u_1u_2,v_1v_2}\nu(d(v_1v_2))\right)\left(F\left(\nu-\frac{\delta_{u_1u_2}}{K}\right)-F(\nu)\right)K\nu(d(u_1u_2))\\ &+\int_{\mathcal{U}^2}(1-\mu_K)f_{u_1u_2}\left(\int_{\mathcal{U}^2}\frac{f_{v_1v_2}R_{u_1u_2}(v_1v_2)}{\langle\nu R_{u_1u_2},f\rangle}(A_{u_1u_2,v_1v_2}F)(\nu)\nu(d(v_1v_2))\right)K\nu(d(u_1u_2))\\ &+\int_{\mathcal{U}^2}\mu_Kf_{u_1u_2}\left(\int_{\mathcal{U}^2}\frac{f_{v_1v_2}R_{u_1u_2}(v_1v_2)}{\langle\nu R_{u_1u_2},f\rangle}(M_{u_1u_2,v_1v_2}F)(\nu)\nu(d(v_1v_2))\right)K\nu(d(u_1u_2)).\end{aligned} \tag{III.5}$$

The first non-linear term describes the competition between individuals. The second and last non-linear terms describe the birth with and without mutation. There, $f_{u_1u_2}\frac{f_{v_1v_2}R_{u_1u_2}(v_1v_2)}{K\langle\nu R_{u_1u_2},f\rangle}$ is the reproduction rate of an individual with genotype u_1u_2 with an individual with genotype v_1v_2. Note that $\nu R_{u_1u_2}$ is the population restricted to the pool of potential partners of an individual of genotype u_1u_2.

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 exists $\bar{f},\bar{D},\bar{c}<\infty$ such that

$$0\leq f_{u_1u_2}\leq\bar{f},\quad 0\leq D_{u_1u_2}\leq\bar{D}\quad\text{and}\quad 0\leq c_{u_1u_2,v_1v_2}\leq\bar{c}. \tag{III.6}$$

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

(A3) There exists a function, $\bar{m}:\mathbb{R}\to\mathbb{R}_+$, such that $\int\bar{m}(h)dh<\infty$ and $m(u,h)\leq\bar{m}(h)$ for any $u\in\mathcal{U}$ and $h\in\mathbb{R}$.

For fixed K, under the Assumptions **(A1)+(A3)** and assuming that $\mathbb{E}(\langle\nu_0,\mathbb{1}\rangle)<\infty$, Fournier and Méléard [36] have shown existence and uniqueness in law of a process with infinitesimal generator L^K. For $K\to\infty$, under mild restrictive assumptions, they prove the convergence of the process ν^K in the space $\mathbb{D}(\mathbb{R}_+,\mathcal{M}^K)$ of càdlàg functions from $\mathbb{R}^+$ to $\mathcal{M}_K$, to a deterministic process, which is the solution to a non-linear integro-differential equation. Assumption **(A2)** ensures that the population does not tend to infinity in finite time or becomes extinct too fast.

1.2. Previous works

Consider the process starting with a monomorphic aa population, with one additional mutant individual of genotype aA. Assume that the phenotype difference between the mutant and the resident population is small. The phenotype difference is assumed to be a slightly smaller death rate compared to the resident population, namely:

$$D_{aa} = D, \quad D_{aA} = D - \Delta. \tag{III.7}$$

for some small enough $\Delta > 0$. The mutation probability for an individual with genotype u_1u_2 is given by μ_K. Hence, the time until the next mutation in the whole population is of order $\frac{1}{K\mu_K}$. Now assume that the demographic parameters introduced in Section 1.1 depend continuously on the phenotype. In particular, they are the same for individuals bearing the same phenotype.

In [18] it is proved that if the two alleles a and A are co-dominant and if the allele A is slightly fitter than the allele a, namely

$$D_{aa} = D, \quad D_{aA} = D - \Delta, \quad D_{AA} = D - 2\Delta, \tag{III.8}$$

then in the limit of large population and rare mutations ($\ln K \ll \frac{1}{\mu_K K} \ll e^{VK}$ for some $V > 0$), the suitably time-rescaled process converges to the TSS model of adaptive dynamics, essentially as shown in [14] in the haploid case. In particular, the genotypes containing the unfit allele a decay exponentially fast after the invasion of AA (see Figure III.1).

If in place of co-dominance we assume, as in [84], that the fittest phenotype A is dominant, namely

$$D_{aa} = D, \quad D_{aA} = D - \Delta, \quad D_{AA} = D - \Delta, \tag{III.9}$$

then 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. Indeed, it was know for some time (see e.g. [82]) that in this case the unique stable fixpoint $(0, 0, \bar{n}_{AA})$ corresponding to a monomorphic AA population is degenerate, i.e. its Jacobian matrix has zero-eigenvalue. This implies that in the deterministic system, the aa and aA populations decay in time only polynomially fast to zero, namely like $1/t^2$ and $1/t$, respectively. This is in contrast to the exponential decay in the co-dominant scenario (see Figure III.1). In [84] it was shown that the deterministic system remains a good approximation of the stochastic system as long as the size of the aA population remains much larger than $K^{1/2}$ and therefore that the a allele survives for a time of order at least $K^{1/2-\alpha}$, for any $\alpha > 0$[1]. Note that this statement is a non trivial fact, since it is not a consequence of the law of large numbers, because the time window diverges as K grows. In summary, the unfit recessive a allele survives in the population much longer due to the slow decay of the aA population.

It is argued in [84] that 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, i.e.

$$\ln K \ll \frac{1}{\mu_K K} \ll K^{1/2-\alpha} \quad \text{as } K \to \infty, \text{for some } \alpha > 0, \tag{III.10}$$

and if the new mutant can coexist with the unfit aa individuals, then the aa population can potentially recover. This is the starting point of the present paper.

[1] In [84] only state that survival occurs up to time $K^{1/4-\alpha}$. However, taking into account that it is really only the survival of the aA population that needs to be ensured, one can easily improve this to $K^{1/2-\alpha}$

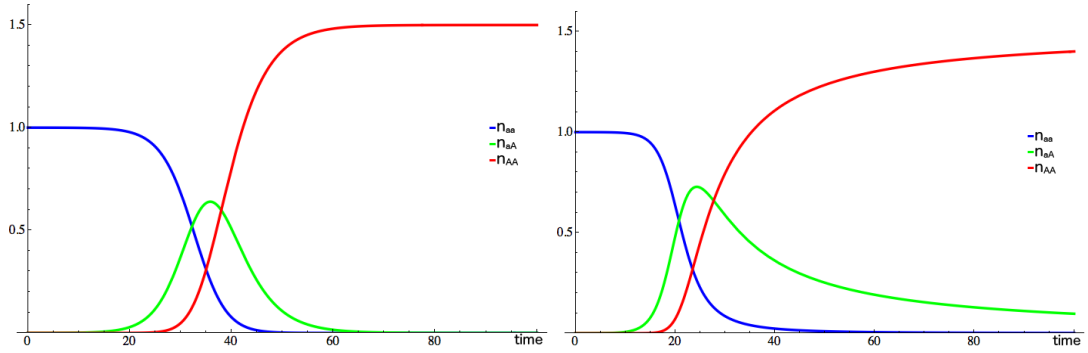

Figure III.1.: Evolution of the model from a resident aa population at equilibrium with a small amount of mutant aA, and when the alleles a and A are co-dominant (left) or when the mutant phenotype A is dominant (right).

1.3. Goal of the paper

The goal of this paper is to show that under reasonable hypothesis, the prolonged survival of the a allele after the invasion of the A allele can indeed lead to a recovery of the aa-type. To do this, we assume that there will occur a new mutant allele, B, that on the one hand has a higher fitness than the AA-phenotype but that (for simplicity) has no competition with the aa-type. The possible genotypes after this mutation are aa, aA, AA, aB, AB, and BB, so that even for the deterministic system we have now to deal with a 6-dimensional dynamical system whose analysis if far from simple.

Under the assumption of dominance of the fittest phenotype, and mutation rate satisfying (III.10), we consider the model described in Section 1.1 starting at the time of the second mutation, that is (with probability converging to 1 as $K \to \infty$) the AA population being close to its equilibrium and the aA population having decreased to a size of order $K\mu_K$, while the aa population is of the order of the square of the aA population. We assume that there just occurred a mutation to a fitter (and most dominant) allele B: we thus start with a quantity $\frac{1}{K}$ of genotype AB. We will start with a population where AA is close to its equilibrium, the populations of aa and aA are already small (of order ε^2 and ε), and by mutation a single individual of genotype AB appears.

By using well known techniques [14, 17, 18], we know that the AB population behaves as a super-critical branching process and reaches the level ε with positive probability in a time of order $\ln K$, without perturbing the 3-system (aa, aA, AA).

We see in numerical solutions to the deterministic system that a reduced fertility together with a reduced competition between a and B phenotypes constitutes a sufficient condition for the recovery of the aa population. For simplicity and in order to prove rigorous results, we suppose that there can be no reproduction between individuals of phenotypes a and B, nor competition between them, and we reduce the number of remaining parameters as much as possible (see Section 2). We study the deterministic system which corresponds to the large population limit of the stochastic counterpart, and we show that (for an initial quantity ε of aA, ε^2 of aa and ε^3 of AB) the system converges to a fixed point denoted by p_{aB} consisting of the two coexisting populations aa and BB. If no further assumptions are made, we will show that the number of individuals bearing an a allele decreases to level $\varepsilon^{1+\Delta/(1-\Delta)}$ (where Δ is defined in (III.7)) before aa grows and stabilises at order 1.

If $\Delta < \frac{\alpha}{1-2\alpha}$, this control on the a allele is in principle sufficient in order for the *stochastic* system to exhibit the recovery of aa with positive probability in the large population limit. Indeed, if the mutation time is of order $K^{\frac{1}{2}-\alpha}$, then the initial amount of aa and aA genotypes is close to the typical fluctuations of those populations. Following the heuristics of [84] (although the six-dimensional stochastic process is surely much more tedious to study), the deterministic system should constitute a good approximation of the process if the typical fluctuations of populations containing an a allele do not bring them to extinction. If $\Delta < \frac{\alpha}{1-2\alpha}$ this ensures that the population containing an a allele is not falling below order $K^{-1/2}$ at any time.

In order to go deeper and control the speed of recovery of the aa population, we look for a parameter regime which ensures that the aa population always grows after the invasion of B. Ensuring this lower bound on aa is not trivial at all, and the solution we found is to introduce an additional parameter η, which lowers the competition between the aA and BB populations, compared to the one between AA and BB. Note that the competition does not depend only on the phenotype, and can be interpreted as a refinement of a phenotypic competition for resources: the strength (or ability to get resources) of an individual not only depends on its phenotype but also on the dominance of its genotype. We show that for η larger than some positive value (of order Δ), the aa population always grows after the invasion of B. The time of convergence to the coexistence fixed point is thus lowered, see Figure III.5. Moreover, we point out the existence of a bifurcation: for η larger than some threshold, the co-existence fixed point p_{aB} becomes unstable and the system converges to another fixed point where all populations coexist.

Our contribution is a rigorous description of a mechanism by which a recessive allele can re-emerge in a population. This can be seen as a statement of genetic robustness exhibited by diploid populations performing sexual reproduction.

The structure of the paper is the following. In Section 2 we describe our assumptions on the parameters of the model, and compute the large population limit; in Section 3 we present our results on the evolution of the deterministic system towards the co-existence fixed point p_{aB}, and we give a heuristic of the proof. Section 4 is dedicated to the proof of these results. The closing Section 5 contain a heuristic considerations and numerical simulations of the model with relaxed assumption on the parameters.

Notation. We write $x = \Theta(y)$ whenever $x = O(y)$ and $y = O(x)$ as $\varepsilon \to 0$.

2. Model setup

Let $\mathcal{G} = \{aa, aA, AA, aB, AB, BB\}$ be the genotype space. Let $n_i(t)$ be the number of individuals with genotype $i \in \mathcal{G}$ in the population at time t and set $n_i^K(t) \equiv \frac{1}{K} n_i(t)$.

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

$$\bar{n}_u = \frac{f_{uu} - D_{uu}}{c_{uu,uu}}. \tag{III.11}$$

Definition III.2. *For $u, v \in \{a, A, B\}$, we call*

$$S_{uv,uu} = f_{uv} - D_{uv} - c_{uv,uu}\bar{n}_u, \tag{III.12}$$

the invasion fitness *of a mutant uv in a resident uu population.*

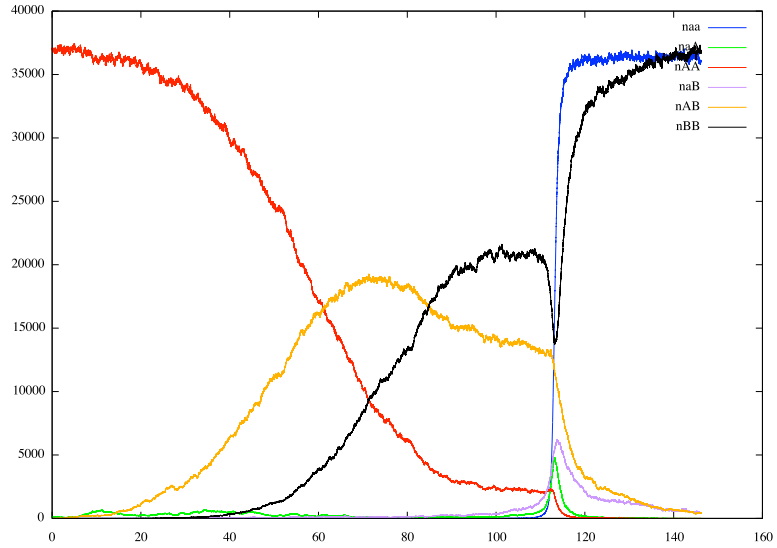

Figure III.2.: Simulation of the stochastic system for $f = 6, D = 0.7, \Delta = 0.1, c = 1, \eta = 0.02, \varepsilon = 0.014$ and $K = 7000$.

We take the phenotypic viewpoint and assume that the B allele is the most dominant one. That means the ascending order of dominance (in the Mendelian sense) is given by $a < A < B$, i.e.

1. phenotype a consists of the genotype aa,
2. phenotype A consists of the genotypes aA, AA,
3. phenotype B consists of the genotypes aB, AB, BB.

For simplicity, we assume that the fertilities are the same for all genotypes, and that natural death rates are the same within the three different phenotypes. Moreover, we assume that there can be no reproduction between a and B phenotypes.
To sumarize, we make the following Assumptions **(B)** on the rates:

(B1) *Fertilities.* For all $i \in \mathcal{G}$, and some $f > 0$

$$f_i \equiv f, \tag{III.13}$$

(B2) *Natural death rates.* The difference in fitness of the three phenotypes is realised by choosing a slightly higher natural death-rate of the a-phenotype and a slightly lower death-rate for the B-phenotype. For some $0 < \Delta < D$,

$$D_{aa} = D + \Delta, \tag{III.14}$$
$$D_{AA} \equiv D_{aA} = D, \tag{III.15}$$
$$D_{aB} \equiv D_{AB} \equiv D_{BB} = D - \Delta. \tag{III.16}$$

(B3) *Competition rates.* We require that phenotypes a and B do not compete with each other. Moreover, we introduce a parameter $\eta \geq 0$ which lowers the competition between BB and aA. For some $0 \leq \eta < c$,

$$(c_{i,j})_{\{i,j\}\in\mathcal{G}\times\mathcal{G}} =: \begin{array}{c|c|cc|ccc} & aa & aA & AA & aB & AB & BB \\ \hline aa & c & c & c & 0 & 0 & 0 \\ \hline aA & c & c & c & c & c & c-\eta \\ AA & c & c & c & c & c & c \\ \hline aB & 0 & c & c & c & c & c \\ AB & 0 & c & c & c & c & c \\ BB & 0 & c-\eta & c & c & c & c \end{array}$$

A biological interpretation for this kind of competition could be that it is coded in the alleles which food an individual with a given genotype prefers. Since an AB individual shares one B allele with a BB individual, they compete stronger for the same food than AA with BB since those have completely different alleles.

(B4) *Reproductive compatibility.* We require that phenotypes a and B do not reproduce with each other,

$$(R_i(j))_{\{i,j\}\in\mathcal{G}\times\mathcal{G}} \equiv \begin{array}{c|c|cc|ccc} & aa & aA & AA & aB & AB & BB \\ \hline aa & 1 & 1 & 1 & 0 & 0 & 0 \\ \hline aA & 1 & 1 & 1 & 1 & 1 & 1 \\ AA & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline aB & 0 & 1 & 1 & 1 & 1 & 1 \\ AB & 0 & 1 & 1 & 1 & 1 & 1 \\ BB & 0 & 1 & 1 & 1 & 1 & 1 \end{array}$$

Observe that, under Assumptions **(B)**,

$$S_{AB,AA} = f - (D - \Delta) - c\bar{n}_{AA} = f - D + \Delta - c\frac{f-D}{c} = \Delta, \tag{III.17}$$

$$S_{aa,BB} = f - D - \Delta. \tag{III.18}$$

Therefore, the mutant AB has a positive invasion fitness in the population AA, as well as aa in the BB population (due to the absence of competition between them).

2.1. Birth rates

We assume that there is no recombination between phenotypes a and B. Thus,

1. the pool of possible partners for the phenotype a consists of phenotypes a and A; the total population of this pool is denoted by

$$\Sigma_3 := n_{aa} + n_{aA} + n_{AA}, \tag{III.19}$$

2. the pool of possible partners for the phenotype A consists of the three phenotypes a, A, and B; the total population of this pool is denoted by

$$\Sigma_6 := n_{aa} + n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}, \tag{III.20}$$

3. the pool of possible partners for the phenotype B consists of phenotypes A and B; the total population of this pool is denoted by

$$\Sigma_5 := n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}. \tag{III.21}$$

Computing the reproduction rates with the Mendelian rules as described in (III.5) leads to the following (time-dependant) birth-rates $b_i = b_i(n(t))$:

$$\begin{aligned} b_{aa} =& f\frac{n_{aa}\left(n_{aa}+\frac{1}{2}n_{aA}\right)}{n_{aa}+n_{aA}+n_{AA}} + f\frac{\frac{1}{2}n_{aB}\left(\frac{1}{2}n_{aA}+\frac{1}{2}n_{aB}\right)}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}} \\ &+ f\frac{\frac{1}{2}n_{aA}\left(n_{aa}+\frac{1}{2}n_{aA}+\frac{1}{2}n_{aB}\right)}{n_{aa}+n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}, \end{aligned} \tag{III.22}$$

$$\begin{aligned} b_{aA} =& f\frac{n_{aa}\left(\frac{1}{2}n_{aA}+n_{AA}\right)}{n_{aa}+n_{aA}+n_{AA}} + f\frac{\frac{1}{2}n_{aA}\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}\right)+\frac{1}{2}n_{aB}\left(n_{AA}+n_{AB}\right)}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}} \\ &+ f\frac{\left(\frac{1}{2}n_{aA}+n_{AA}\right)\left(n_{aa}+n_{aA}+\frac{1}{2}n_{aB}\right)+\frac{1}{4}n_{aA}n_{AB}}{n_{aa}+n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}, \end{aligned} \tag{III.23}$$

$$b_{AA} = f\frac{\frac{1}{2}n_{AB}\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}} + f\frac{\left(\frac{1}{2}n_{aA}+n_{AA}\right)\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{n_{aa}+n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}, \tag{III.24}$$

$$b_{aB} = f\frac{\left(\frac{1}{2}n_{aA}+n_{aB}\right)\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}} + f\frac{\frac{1}{2}n_{aA}\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{n_{aa}+n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}, \tag{III.25}$$

$$\begin{aligned} b_{AB} =& f\frac{\left(\frac{1}{2}n_{aA}+n_{AA}+n_{AB}\right)\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}} \\ &+ f\frac{\left(\frac{1}{2}n_{aA}+n_{AA}\right)\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{n_{aa}+n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}, \end{aligned} \tag{III.26}$$

$$b_{BB} = f\frac{\frac{1}{4}\left(n_{aB}+n_{AB}+2n_{BB}\right)^2}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}. \tag{III.27}$$

2.2. Death rates

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

$$d_{aa} = n_{aa}(D + \Delta + c(n_{aa} + n_{aA} + n_{AA})), \tag{III.28}$$
$$d_{aA} = n_{aA}(D + c(n_{aa} + n_{aA} + n_{AA} + n_{aB} + n_{AB}) + (c - \eta)n_{BB}), \tag{III.29}$$
$$d_{AA} = n_{AA}(D + c(n_{aa} + n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB})), \tag{III.30}$$
$$d_{aB} = n_{aB}(D - \Delta + c(n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB})), \tag{III.31}$$
$$d_{AB} = n_{AB}(D - \Delta + c(n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB})), \tag{III.32}$$

$$d_{BB} = n_{BB}(D - \Delta + (c - \eta)n_{aA} + c(n_{AA} + n_{aB} + n_{AB} + n_{BB})). \tag{III.33}$$

2.3. Large population limit

By [36] or [18], for large populations, the behaviour of the stochastic process is close to the solution of a deterministic equation.

Proposition III.1 (Generalisation of Proposition 3.2 in [18])**.**
Let $T > 0$ and $C \subset \mathbb{R}^6_+$ be a compact set. Assume that the initial condition $n^K(0) = \frac{1}{K}(n_{aa}(0), n_{aA}(0), n_{AA}(0), n_{aB}(0), n_{AB}(0), n_{BB}(0))$ converges almost surely to a deterministic vector $x^0 = (x^0_1, x^0_2, x^0_3, x^0_4, x^0_5, x^0_6) \in C$, as $K \to \infty$.
Let $\tilde{n}(t, x^0)$ denote the solution to

$$\dot{n}(t) = b(n(t)) - d(n(t)) \equiv F(n(t)), \tag{III.34}$$

$$rm\ i.e. \quad \dot{n}_i(t) = b_i(n(t)) - \left(D_i + \sum_{j \in \mathcal{G}} c_{i,j} n_j(t) \right) n_i(t), \quad \text{for all } i \in \mathcal{G}, \tag{III.35}$$

with initial condition x^0, where $(b_i)_{i\in\mathcal{G}}$ and $(d_i)_{i\in\mathcal{G}}$ are given in (III.22)-(III.27) *and* (III.28)-(III.33)*. Then, for all $T > 0$,*

$$\lim_{K\to\infty} \sup_{t\in[0,T]} |n^K_i(t) - \tilde{n}_i(t, x^0))| = 0, \quad a.s., \tag{III.36}$$

for all $i \in \mathcal{G}$.

2.4. Initial condition

Fix $\varepsilon > 0$ sufficiently small. For the results below, we will consider the dynamical system (III.34) starting with the initial condition:

$$\bar{n}_A \geq n_{AA}(0) \geq \bar{n}_A - \Theta(\varepsilon), \tag{III.37}$$

$$n_{aA}(0) = \varepsilon, \tag{III.38}$$

$$n_{aa}(0) = \Theta(\varepsilon^2), \tag{III.39}$$

$$n_{AB}(0) = \varepsilon^3, \tag{III.40}$$

$$n_{BB}(0) = 0, \tag{III.41}$$

$$n_{aB}(0) = 0. \tag{III.42}$$

Remark III.1. *In all the figures below, the choice of parameters is the following:*

$$f = 6, \quad D = 0.7, \quad \Delta = 0.1, \quad c = 1, \quad \varepsilon = 0.01,$$

and the parameter η is specified on each picture.

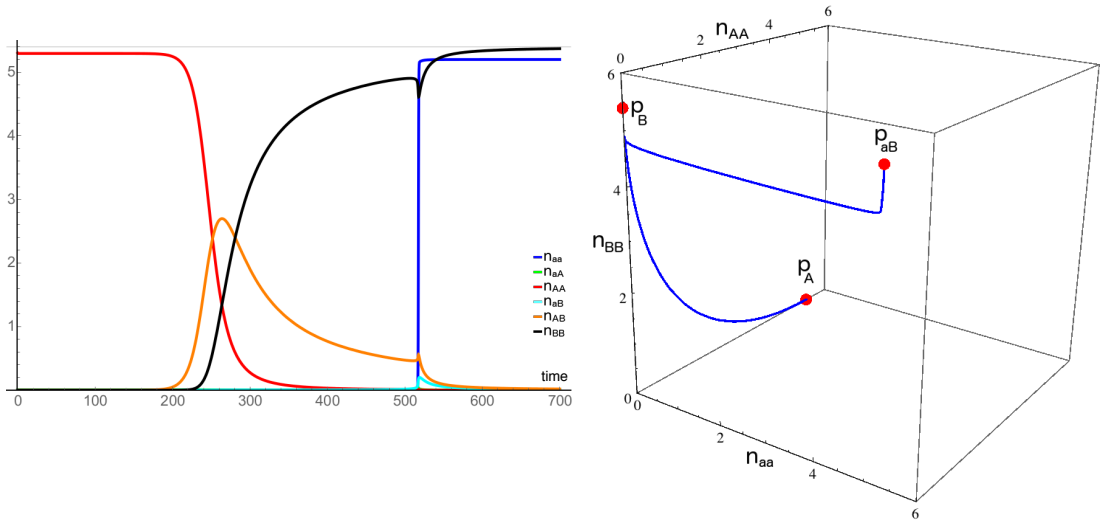

Figure III.3.: General qualitative behaviour of $\{n_i(t), i \in \mathcal{G}\}$ and projection of the dynamical system on the coordinates aa, AA and BB. The re-invasion of the aa population happens sooner and sooner as η grows ($\eta = 0.02$ for both pictures).

3. Results

We are working with a 6-dimensional dynamical system, and computing all the fixed points analytically is impossible for a general choice of the parameters. We can, however, compute those which are relevant for our study. We will call p_A (resp. p_B) the fixed points corresponding to the monomorphic AA (resp. BB) population at equilibrium, and p_{aB} the fixed point corresponding to the coexisting aa and BB populations. Setting the relevant populations to 0 and solving $\dot{n}(t) = 0$, we get:

$$p_A = (0, 0, \bar{n}_A, 0, 0, 0), \tag{III.43}$$
$$p_B = (0, 0, 0, 0, 0, \bar{n}_B), \tag{III.44}$$
$$p_{aB} = (\bar{n}_a, 0, 0, 0, 0, \bar{n}_B), \tag{III.45}$$

where $\bar{n}_a = \frac{f-D-\Delta}{c}, \bar{n}_A = \frac{f-D}{c}$, and $\bar{n}_B = \frac{f-D+\Delta}{c}$. Note that the BB equilibrium population is the same in p_B and p_{aB}. This is due to the non-interaction between phenotypes a and B.

Our general result is that starting with initial conditions (III.37)-(III.42), that is close to p_A (with small coordinates in directions aa, aA and AB), and under minimal assumptions on the parameters, the system gets very close to p_B before finally converging to p_{aB}, see Figure III.3.

Theorem III.1. *Consider the dynamical system* (III.34) *started with initial conditions* (III.37)-(III.42). *Suppose the following Assumptions* ***(C)*** *on the parameters hold:*

(C1) Δ *sufficiently small,*

(C2) f *sufficiently large,*

(C3) $0 \leq \eta < c/2$.

Then the system converges to the fixed point p_{aB}. More precisely, for any fixed $\delta > 0$, as $\varepsilon \to 0$, it reaches a δ-neighbourhood of p_{aB} in a time of order $\Theta(\varepsilon^{-1/(1+\eta\bar{n}_B-\Delta)})$.
Moreover, it holds:

1. *for $\eta = 0$, the amount of allele a in the population decays to $\Theta(\varepsilon^{1+\Delta/(1+\Delta)})$ before reaching $\Theta(1)$,*
2. *for $\eta > \frac{4\Delta}{\bar{n}_B}$, the amount of a allele in the population is bounded below by $\Theta(\varepsilon)$ for all $t > 0$.*

Remark III.2. *For η large, we prove that the fixed point p_{aB} is unstable. We observe numerically that the system is attracted to a fixed point where all the 6 populations coexist, but we do not prove this.*

Let us now briefly discuss the linear stability of the relevant fixed points and give an heuristics of the proof of Theorem III.1.

3.1. Linear stability analysis

The Jacobian matrix $J_F := (\partial F_i/\partial n_j)_{ij}$ of the map F defined in (III.34) can be explicitly computed at p_A and p_{aB} and the situation is as follows:

- The eigenvalues of $J_F(p_A)$ are $0, \Delta > 0$ and $-(f-D), -(f+\Delta), -(f-\Delta)$ (double) which are all strictly negative under Assumptions (C). The fixed point p_A is thus unstable.
- The eigenvalues of $J_F(p_{aB})$ are 0 (double), and $-(2f-D), -(f-D+\Delta), -(f-D-\Delta), -((f-D)(5f-4D)+f\Delta)/(4(f-D)+\eta\bar{n}_B)$ which are strictly negative under Assumptions (C). The linear analysis thus does not imply the stability of p_{aB} but the Phase 4 of the proof does (see Section 4.5) .

It turns out that $J_F(p_B)$ is singular but as the invasion fitness of aa is positive, i.e. $S_{aa,BB} > 0$ (see (III.17)), this implies that a small perturbation in the first coordinate will be amplified, and thus implies the instability of the fixed point p_B.

3.2. Heuristics of the proof

Recall we start the dynamical system (III.34) with initial conditions (III.37)-(III.42). A numerical solution of the system is provided on Figure III.4.

Remark III.3. *Assumption C1 of Theorem III.1 is needed throughout the proof in order to be able to use the results of [84] which rely on the Center Manifold Theorem (a line of fixed points becomes an invariant line under small enough perturbation).*

Phase 1. Time period: until $n_{AB} = \varepsilon_0$.
The mutant population, consisting of all individuals of phenotype B, first grows up to ε_0 exponentially fast with rate Δ without perturbing the behaviour of the 3-system (aa, aA, AA). The rate of growth corresponds to the invasion fitness of AB in the resident population AA, see (III.17). Following [84], AA stays close to $\bar{n}_A$, while aA and aa continue to decay like $1/t$ and $1/t^2$ respectively. The duration T_1 of this phase is such that $\Theta(\varepsilon^3)e^{t\Delta} = \Theta(1) \Leftrightarrow T_1 = \Theta(|\log\varepsilon|)$.

Phase 2. Time period: until $n_{aA} = \Theta(n_{AA})$.

The evolution is a perturbation of an effective 3-system (AA, AB, BB) which behaves exactly the same as in [84], since the parameters satisfy the same hypotheses (slightly lower death rate for phenotype B than for phenotype A, and constant competition parameters). A comparison result (following Theorem III.2 below) shows that this 3-system is almost unperturbed until $n_{aA} = \Theta(n_{AA})$. If that happens in a time T_2 diverging with ε (which we ensure throughout the calculation), we thus know that BB approaches $\bar{n}_B$, while $n_{AB} \propto 1/t$ and $n_{AA} \propto 1/t^2$.

The important fact in this phase is that the amount of allele a in the population decays for η small while it increases for large enough η. Indeed, let us derive some bounds on $\Sigma_{aA,aB} = n_{aA} + n_{aB}$. The population $\Sigma_{aA,aB}$ reproduces by taking the dominant allele in a population of order $\Theta(1)$ and the allele a in itself. Thus its birth rate satisfy $b_{\Sigma_{aA,aB}} \approx f\Sigma_{aA,aB}$. We can compute its death rate exactly and use that $n_{BB} \approx \Sigma_5 \approx \bar{n}_B$:

$$\begin{aligned} d_{\Sigma_{aA,aB}} &= \Sigma_{aA,aB}(D - \Delta + c\Sigma_5) - \eta n_{aA} n_{BB} + \Delta n_{aA} \\ &\approx f\Sigma_{aA,aB} - n_{aA}(\eta\bar{n}_B - \Delta), \qquad \text{(III.46)} \\ \dot{\Sigma}_{aA,aB} &\approx n_{aA}(\eta\bar{n}_B - \Delta) \\ &= \Theta(\Sigma_{aA,aB} \cdot n_{AB})(\eta\bar{n}_B - \Delta). \qquad \text{(III.47)} \end{aligned}$$

The last equality comes from the fact that aA newborns have mainly their a allele coming from $\Sigma_{aA,aB}$ and their A allele coming from AB. Using the $1/t$ decay of AB we get:

$$\dot{\Sigma}_{aA,aB} \approx \frac{\Theta(\Sigma_{aA,aB})}{\Theta(1) + \Theta(1)t}(\eta\bar{n}_B - \Delta). \qquad \text{(III.48)}$$

As $\Sigma_{aA,aB}(T_1) = \Theta(\varepsilon)$ we deduce that $\Sigma_{aA,aB}(t) = \Theta(\varepsilon)(\Theta(1) + \Theta(1)t)^{\Theta(\eta\bar{n}_B - \Delta)}$, and thus $n_{aA} = \Theta(n_{AB} \cdot \Sigma_{aA,aB}) = \Theta(\varepsilon)(\Theta(1) + \Theta(1)t)^{\Theta(\eta\bar{n}_B - \Delta)}/(\Theta(1) + \Theta(1)t)$. By solving $n_{aA} = \Theta(n_{AA}) = \Theta(n_{AB}^2)$ we get the order of magnitude of $T_2 = \Theta\left(\varepsilon^{-1/(1+\eta\bar{n}_B - \Delta)}\right)$. Note that for $\eta = 0$, $\Sigma_{aA,aB}(T_2) = \Theta\left(\varepsilon^{1+\Delta/(1-\Delta)}\right)$. Moreover, (III.47) implies that for $\eta > \Delta/\bar{n}_B$, we have $\dot{\Sigma}_{aA,aB} > 0$, which proves points 1 and 2 of Theorem III.1.

Phase 3. Time period: until aa reaches equilibrium.

The fact that $n_{aA} = \Theta(n_{AA})$ has a crucial effect on the birth rate of aa (see (III.22)) since the term $(n_{aa} + \frac{1}{2}n_{aA})/(n_{aa} + n_{aA} + n_{AA})$ becomes of order $\Theta(1)$. As long as AA stays smaller than $\Theta(\varepsilon)$, we get a lower bound on n_{aa} which grows exponentially fast since f is chosen large enough (Assumption C2):

$$\begin{aligned} b_{aa} &\geq f n_{aa}\Theta(1), \qquad \text{(III.49)} \\ d_{aa} &\leq n_{aa}(D + \Delta + \Theta(\varepsilon)), \qquad \text{(III.50)} \\ \dot{n}_{aa} &\geq n_{aa}(f\Theta(1) - D - \Delta - \Theta(\varepsilon)). \qquad \text{(III.51)} \end{aligned}$$

As aa grows, it makes $\Sigma_{aA,aB}$ grow, and thus AA and AB as well. We have to show that this could not prevent aa from reaching equilibrium. We do not give a detailed argument here, but essentially, the presence of the macroscopic BB population prevents all the non-aa populations to grow too much. Note that if η is too large, then aA could get a positive fitness and grow to a macroscopic level. That is why we have to impose Assumption C3, which

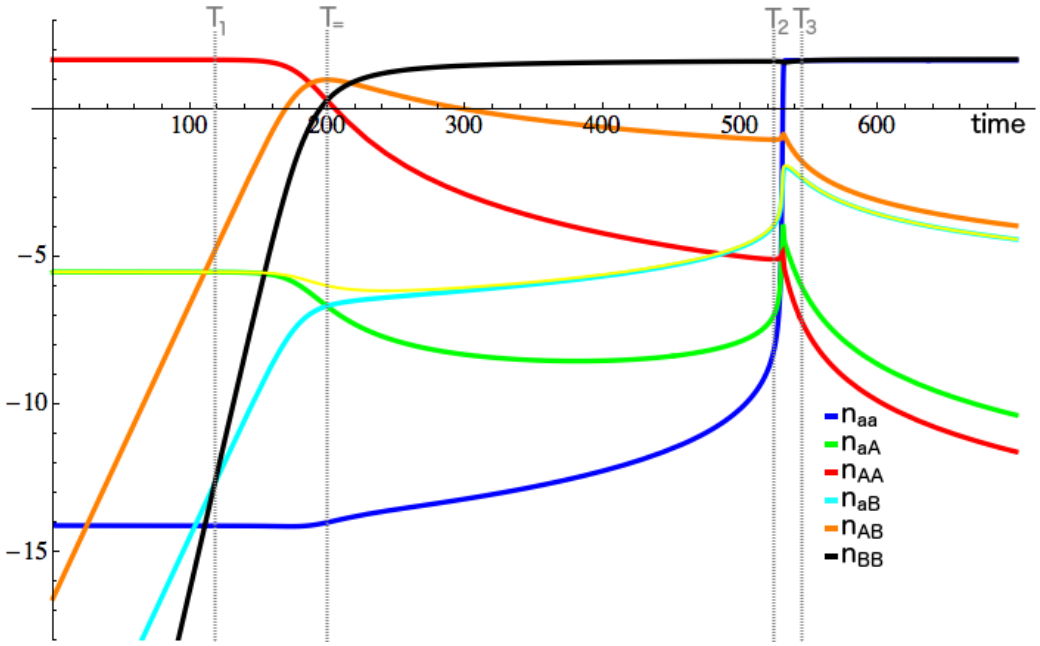

Figure III.4.: Numerical solution of the deterministic system for $\eta = 0.02$, logplot.

will become clearer heuristically in the next phase. We recall that aa does not compete with BB and thus it grows exponentially fast with rate $f - (D + \Delta)$ until an ε_0-neighbourhood of the fixed point where aa and BB coexist. The rate of growth corresponds to the invasion fitness of aa in the resident population BB, see (III.17). Note that, due to Assumption C2, this rate is much larger than the invasion rate of BB into AA. That is why the fourth phase looks very steep on Figure III.4, see the stretched version on Figure III.6. This phase lasts a time $T_3 = \Theta(|\log \varepsilon|)$.

Phase 4. The Jacobian matrix of the field (III.34) at the fixed point p_{aB} has two zero, and 4 negative eigenvalues. p_{aB} is thus a non-hyperbolic equilibrium point of the system and linearisation fails to determine its stability properties. Instead, we use the result of center manifold theory [51, 88] that asserts that the qualitative behaviour of the dynamical system in a neighbourhood of the non-hyperbolic critical point p_{aB} is determined by its behaviour on the center manifold near p_{aB}. Using the Center Manifold Theorem, we show that asymptotically as $f \to \infty$, the field is attractive for $\eta < c \cdot r_{max}$ where $r_{max} \simeq 0.593644$ is the maximum of the rational function (III.322). Thus p_{aB} is a stable fixed point which is approached with speed $\frac{1}{t}$ as long as $\eta < c \cdot r_{max}$. For higher values of η, numerical solutions show that the system converges to a fixed point where the 6 populations co-exist, but we do not prove this.

4. Proof

Definition III.3. *Let $x, y, z \in \{aa, aA, AA, aB, AB, BB\}$ and $h \in \mathbb{R}$. We define*

$$T^{x=y} = \inf\{t > 0 : n_x(t) = n_y(t)\}, \tag{III.52}$$
$$T^{x=\delta y} = \inf\{t > 0 : n_x(t) = \delta n_y(t)\}, \tag{III.53}$$
$$T_h^x = \inf\{t > 0 : n_x(t) > h\}, \tag{III.54}$$
$$T_h^{x+y} = \inf\{t > 0 : n_x(t) + n_y(t) > h\}, \tag{III.55}$$
$$T_h^{x+y+z} = \inf\{t > 0 : n_x(t) + n_y(t) + n_z(t) > h\}. \tag{III.56}$$

Moreover, let

$$\Delta > \varepsilon_0 > \varepsilon > 0. \tag{III.57}$$

The value ε_0 is the small order 1 level in the Phase 1, see the proof heuristics (Section 3.2). We consider Δ fixed and sufficiently small, and will first send $\varepsilon \to 0$ and then $\varepsilon_0 \to 0$.

4.1. Preliminaries

We first prove general facts which will be useful through the proof.

Lemma III.1. *Let $c > 0$ and $n(t)$ be such that*

- *$\dot{n}(t) \leq g(t) - c \cdot n(t)$ for all $t \in \mathcal{T} \subset \mathbb{R}^+$,*
- *$c \cdot n(0) \leq g(0)$,*

if $c \cdot n(t) = g(t) \Rightarrow c \cdot \dot{n}(t) \leq \dot{g}(t)$ for all $t \in \mathcal{T}$ then $c \cdot n(t) \leq g(t)$ for all $t \in \mathcal{T}$.

Proof. This is an easy analysis exercise. □

Proposition III.2. *If $n_{aB}(0) < n_{AB}(0)$ then $n_{aB}(t) \leq n_{AB}(t)$.*

Proof. Intuitively this inequality comes from the fact that phenotype a individuals cannot reproduce with phenotype B. Indeed, if we consider the couples that could give rise to an AB (resp. aB) individual, they are of the form (Ag_1, Bg_2) (resp. (ag_1, Bg_2)), with $g_1, g_2 \in \{a, A, B\}$ and the combination (AA, Bg_2) is possible whereas (aa, Bg_2) is impossible. Here is the rigorous derivation of the result: We compare the birth- and the death-rates of n_{AB} and n_{aB}

$$\frac{d_{aB}}{n_{aB}} = D - \Delta + c(n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}) = \frac{d_{AB}}{n_{AB}}, \tag{III.58}$$

$$b_{aB} = f n_{aB} \frac{\frac{1}{2}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} + I_{aB}, \tag{III.59}$$

$$b_{AB} = f n_{AB} \frac{\frac{1}{2}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} + I_{AB}. \tag{III.60}$$

We see that the death-rates of the two populations are the same, whereas the birth-rates differ only in a factor which comes from the reproduction of the other populations. If we

take a closer look to these factors I_{aB}, I_{AB} under the assumption that $n_{aB} = n_{AB}$ we see that

$$\begin{aligned} I_{AB} &= f\left(\tfrac{1}{2}n_{aA} + n_{AA}\right)\left(\frac{\frac{1}{2}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}}{\Sigma_5} + \frac{\frac{1}{2}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}}{\Sigma_6}\right) \\ &= I_{aB} + fn_{AA}\left(\frac{\frac{1}{2}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}}{\Sigma_5} + \frac{\frac{1}{2}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}}{\Sigma_6}\right). \end{aligned} \tag{III.61}$$

Thus $I_{AB} > I_{aB}$. Hence, $\dot{n}_{AB} > \dot{n}_{aB}$ and $n_{AB}(t)$ stays above $n_{aB}(t)$ for all $t > 0$. □

4.2. Phase 1: Perturbation of the 3-system (aa, aA, AA) until AB reaches $\Theta(1)$

We start with initial conditions given by (III.37)-(III.42). We will show that the mutant population, consisting of all individuals of phenotype B, grows up to some $\varepsilon_0 > \varepsilon$ without perturbing the behaviour of the 3-system (aa, aA, AA) in this time. Let

$$T_1 := T_{\varepsilon_0}^{aB+AB+BB}. \tag{III.62}$$

Proposition III.3. *With the initial conditions* (III.37)-(III.42)*, for all* $t \in [0, T_1]$*, it holds,*

1. $n_{AB}(t)$ *grows exponentially with rate* Δ*. It reaches the level* ε_0 *in a time at most of order* $\Theta\left(\log\left((\varepsilon_0/\varepsilon^3)^{\frac{1}{\Delta-\Theta(\varepsilon_0)}}\right)\right)$.
2. $n_{aB}(t) \le \Theta(\varepsilon^{1-\Theta(\varepsilon_0)}\varepsilon_0)$, $n_{aA}(t) \le \Theta(\varepsilon^{1-\Theta(\varepsilon_0)})$, $n_{aa}(t) \le \Theta(\varepsilon^{2-\Theta(\varepsilon_0)})$ *and* $\bar{n}_A - \Theta(\varepsilon_0) \le n_{AA}(t) \le \bar{n}_A + \Theta(\varepsilon_0)$.
3. $n_{BB}(t) = \Theta\left(n_{AB}^2(t)\right)$.

Proof. Until T_1 the perturbation of the dynamics of the 3-system (aa, aA, AA) is at most of order ε_0. Thus we have $\bar{n}_A - \Theta(\varepsilon_0) \le n_{AA}(t) \le \bar{n}_A + \Theta(\varepsilon_0)$, as well as $n_{aa}, n_{aA} \le \Theta(\varepsilon_0)$. With this rough bounds we will find finer bounds.

1. The Δ reduced death rate of the mutant AB gives it a positive fitness, and the growth is exponential until it reaches a macroscopic level. For an upper bound on the time $T_{\varepsilon_0}^{aB+AB+BB}$, we have to construct a minorising process for n_{AB}. Indeed, let us compare the birth and death rates:

$$b_{AB} \ge \tfrac{1}{2}n_{AB}\frac{2fn_{AA}}{n_{AA} + \Theta(\varepsilon_0)} = n_{AB}(f - \Theta(\varepsilon_0)), \tag{III.63}$$

$$d_{AB} \le n_{AB}(D - \Delta + c\bar{n}_A + \Theta(\varepsilon_0)) = n_{AB}(f - \Delta + \Theta(\varepsilon_0)). \tag{III.64}$$

Hence, we get for the minorising process

$$\dot{n}_{AB} \ge n_{AB}(\Delta - \Theta(\varepsilon_0)), \tag{III.65}$$

$$n_{AB}(t) \ge \varepsilon^3 e^{(\Delta-\Theta(\varepsilon_0))t}, \tag{III.66}$$

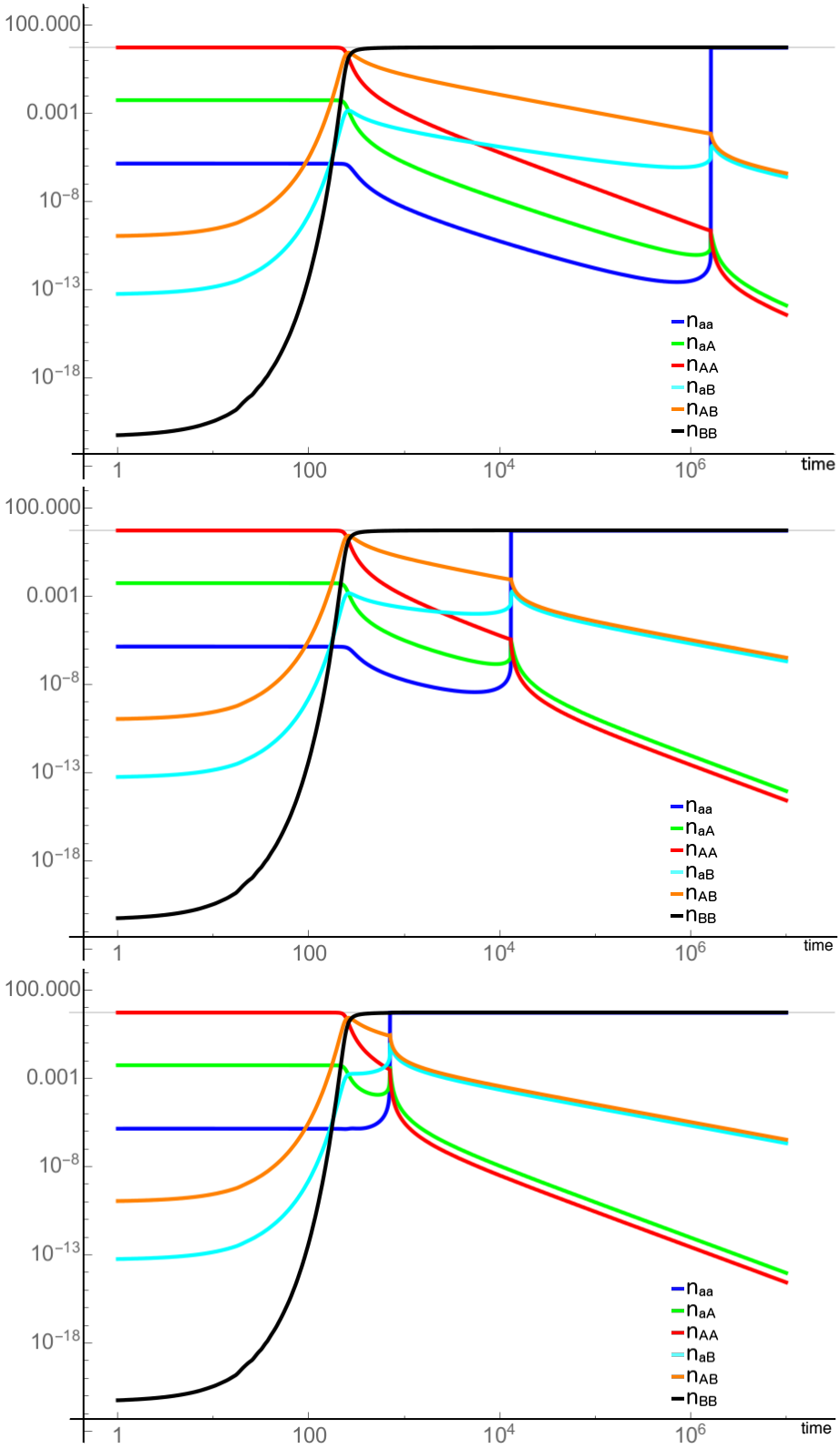

Figure III.5.: Log-plots of $\{n_i(t), i \in \mathcal{G}\}$ for $\eta = 0$ (top), $\eta = 0.003$ (center) and $\eta = 0.014$ (bottom).

and the time T_1 is at most of order $\Theta\left(\log((\varepsilon_0/\varepsilon^3)^{\frac{1}{\Delta-\Theta(\varepsilon_0)}}\right)$. For an lower bound on the time T_1, we have to construct a majorising process for n_{AB}. We compare the birth and death rates:

$$b_{AB} \leq n_{AB}(f + \Theta(\varepsilon_0)), \tag{III.67}$$
$$d_{AB} \geq n_{AB}(D - \Delta + c\bar{n}_A - \Theta(\varepsilon_0)) = n_{AB}(f - \Delta - \Theta(\varepsilon_0)). \tag{III.68}$$

Hence, we get for the majorising process

$$\dot{n}_{AB} \leq n_{AB}(\Delta + \Theta(\varepsilon_0)), \tag{III.69}$$
$$n_{AB}(t) \leq \varepsilon^3 e^{(\Delta+\Theta(\varepsilon_0))t}, \tag{III.70}$$

and the time T_1 is at least of order $\Theta\left(\log((\varepsilon_0/\varepsilon^3)^{\frac{1}{\Delta-\Theta(\varepsilon_0)}}\right)$.

2. Heuristically, the newborns of genotype aA are still in majority produced by recombination of AA and aA, because the mutant population is not large enough to contribute. The newborns of genotype aB are in majority produced by reproduction of the aA population with the B population. Finally, the newborns of genotype aa are in majority produced by recombination of aA and aA, because the only mutant which could perturb it is aB which is of smaller order.

   a) We show that $n_{aa} \leq n_{aA}^2$, or according to Lemma III.1, $\dot{n}_{aa} - 2\dot{n}_{aA}n_{aA} \leq 0$ when $n_{aa} = n_{aA}^2$. Observe that $\dot{n}_{aa} - 2\dot{n}_{aA}n_{aA} = b_{aa} - 2n_{aA}b_{aA} - d_{aa} + 2n_{aA}d_{aA}$. The biggest contributing terms of $b_{aa} - 2n_{aA}b_{aA}$ and $d_{aa} - 2n_{aA}d_{aA}$ at $n_{aa} = n_{aA}^2$ are

   $$b_{aa} - 2n_{aA}b_{aA} = \tfrac{f}{4\Sigma_5}n_{aB}^2 - \tfrac{2f}{\Sigma_6}n_{AA}n_{aA}^2, \tag{III.71}$$
   $$d_{aa} - 2n_{aA}d_{aA} = -n_{aA}^2(f - \Delta + \Theta(\varepsilon_0)). \tag{III.72}$$

   Thus we get, as long as $n_{aB} < n_{aA}$, that

   $$\begin{aligned}\dot{n}_{aa} - 2\dot{n}_{aA}n_{aA} &= b_{aa} - 2n_{aA}b_{aA} - d_{aa} + 2n_{aA}d_{aA} \\ &\leq n_{aA}^2\left(f - \tfrac{2f}{\Sigma_6}n_{AA} - \Delta + \Theta(\varepsilon_0)\right) + \tfrac{f}{4\Sigma_5}n_{aB}^2 < 0.\end{aligned} \tag{III.73}$$

   b) We show that n_{aB} really stays smaller than n_{aA}, precisely we show that $n_{aB} \leq n_{aA}n_{AB}$ or equivalently according to Lemma III.1 $\dot{n}_{aB} - \dot{n}_{aA}n_{AB} - \dot{n}_{AB}n_{aA} \leq 0$ at $n_{aB} = n_{aA}n_{AB}$.
   The biggest contributing terms are

   $$\begin{aligned}b_{aB} - n_{AB}b_{aA} - n_{aA}b_{AB} =& n_{aA}n_{AB}\left(\tfrac{f}{4\Sigma_5} + \tfrac{f}{4\Sigma_6} - \tfrac{f}{\Sigma_6}n_{AA} - \tfrac{f}{2\Sigma_5}n_{AA} - \tfrac{f}{2\Sigma_6}n_{AA}\right) \\ &+ n_{aA}n_{BB}\left(\tfrac{f}{2\Sigma_5} + \tfrac{f}{2\Sigma_6} - \tfrac{f}{\Sigma_5}n_{AA} - \tfrac{f}{\Sigma_6}n_{AA}\right),\end{aligned} \tag{III.74}$$
   $$d_{aB} - n_{AB}d_{aA} - n_{aA}d_{AB} = -n_{aA}n_{AB}(D + c\Sigma_6 - \eta n_{BB}). \tag{III.75}$$

   Thus we get

   $$\begin{aligned}\dot{n}_{aB} - \dot{n}_{aA}n_{AB} - \dot{n}_{AB}n_{aA} \leq& n_{aA}n_{AB}\left(-f + \tfrac{f}{2\Sigma_5} + \Theta(\varepsilon_0)\right) \\ &- n_{aA}n_{BB}\left(2f - \tfrac{f}{\Sigma_5} - \Theta(\varepsilon_0)\right) < 0.\end{aligned} \tag{III.76}$$

c) We show that $n_{aA} \leq \Theta\left(\varepsilon^{1-\varepsilon_0}\right)$.
We construct a majorising process on aA. The biggest contributing terms are

$$b_{aA} \leq \tfrac{f}{\Sigma_6} n_{AA} n_{aA} + n_{aA}\Theta(\varepsilon_0), \tag{III.77}$$

$$d_{aA} \geq n_{aA}(f - \Theta(\varepsilon_0)), \tag{III.78}$$

and we get that

$$\dot{n}_{aA} \leq \Theta(\varepsilon_0) n_{aA}, \tag{III.79}$$

$$n_{aA}(t) \leq \varepsilon e^{\Theta(\varepsilon_0)t}, \tag{III.80}$$

what shows that until time T_1, $n_{aA} \leq \Theta\left(\varepsilon^{1-\varepsilon_0}\right)$.

d) We show $\bar{n}_A - \varepsilon \leq \Sigma_5 \leq \bar{n}_A + 2\Delta\varepsilon_0$.
We construct a minorising and a majorising processes on Σ_5:

$$b_{\Sigma_5} \leq f\Sigma_5 + \Theta(n_{aa}), \tag{III.81}$$

$$b_{\Sigma_5} \geq f\Sigma_5 - \Theta(n_{aA}^2), \tag{III.82}$$

$$d_{\Sigma_5} \geq \Sigma_5(D + c\Sigma_5) - (\Delta + 2\eta n_{aA})(n_{aB} + n_{AB} + n_{BB}), \tag{III.83}$$

$$d_{\Sigma_5} \leq \Sigma_5(D + c\Sigma_5 + cn_{aa}), \tag{III.84}$$

$$\dot{\Sigma}_5 \leq \Sigma_5(f - D - c\Sigma_5) + (\Delta + 2\eta n_{aA})(n_{aB} + n_{AB} + n_{BB}), \tag{III.85}$$

$$\dot{\Sigma}_5 \geq \Sigma_5(f - D - c\Sigma_5 - cn_{aa}). \tag{III.86}$$

At the upper bound we have $\dot{\Sigma}_5 \leq 0$ and at the lower bound $\dot{\Sigma}_5 \geq 0$, which ensure the claimed bounds by Lemma III.1.

3. The newborns of genotype BB are in majority produced by recombination of AB with itself. Indeed, by comparison of the birth- and death-rates,

$$\begin{aligned} b_{BB} &\leq f n_{BB} \frac{n_{aB} + n_{AB} + n_{BB}}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} + \frac{f}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} n_{AB}^2 \\ &\leq f n_{BB}\Theta(\varepsilon_0) + \tfrac{f}{\bar{n}_A} n_{AB}^2 + \Theta(\varepsilon_0^3), \end{aligned} \tag{III.87}$$

$$b_{BB} \geq f n_{BB}\Theta(\varepsilon_0) + \tfrac{f}{4\bar{n}_B} n_{AB}^2, \tag{III.88}$$

$$d_{BB} \geq n_{BB}(D - \Delta + c\bar{n}_A - \Theta(\varepsilon_0)) = n_{BB}(f - \Delta - \Theta(\varepsilon_0)), \tag{III.89}$$

$$d_{BB} \leq n_{BB}(D - \Delta + c\bar{n}_B + \Theta(\varepsilon_0)) = n_{BB}(f + \Theta(\varepsilon_0)). \tag{III.90}$$

We get the upper bound for the process

$$\dot{n}_{BB} \leq -n_{BB}(f(1 - \Theta(\varepsilon_0)) - \Delta - \Theta(\varepsilon_0)) + \tfrac{f}{\bar{n}_A} n_{AB}^2, \tag{III.91}$$

and the lower bound

$$\dot{n}_{BB} \geq -n_{BB}(f + \Theta(\varepsilon_0)) + \tfrac{f}{4\bar{n}_B} n_{AB}^2. \tag{III.92}$$

By applying Lemma III.1 to $n = n_{BB}$ and $g = n_{AB}^2$ (with constants in front), as $n_{BB}(0) = 0 < n_{AB}(0) = \varepsilon^3$ and by Proposition III.3 (2) $\dot{n}_{AB} \geq 0$ for all $t \in [0, T_1]$, we deduce that $n_{BB}(t) \leq \Theta(n_{AB}^2(t))$, for all $t \in [0, T_1]$.

□

Note that Proposition III.3 implies that

$$T_1 = T_{\varepsilon_0}^{aB+AB+BB} = T_{\varepsilon_0}^{AB} \leq \Theta\left(\log\left(\left(\varepsilon_0/\varepsilon^3\right)^{\frac{1}{\Delta - \Theta(\varepsilon_0)}}\right)\right). \tag{III.93}$$

4.3. Phase 2: Perturbation of the 3-system (AA, AB, BB) until $n_{aA} = \Theta(n_{AA})$

Let $\delta > 0$ (to be chosen sufficiently small in the sequel). Let

$$T_2 := T^{aA=\delta AA} \wedge T^{aB=\delta AB} \wedge T^{aa=aA \wedge aB}. \tag{III.94}$$

We will show that for $t \in [T_1, T_2]$ the system behaves as a main 3-system (AA, AB, BB) plus perturbations of order δ. The 3-system (AA, AB, BB) behaves exactly the same as in [84] since the parameters satisfy the same hypotheses (slightly lower death rate for phenotype B than for phenotype A individuals, and constant competition parameters).

Moreover, the crucial role of the parameter η is that the population containing an allele a only continues to grow in this phase when η is large enough. This is due to the smaller competition that aA feels from BB, the aA population is thus higher and induces the growth of aB.

We start by considering how the growth of aA- and aB populations can perturb the 3-system (AA, AB, BB).

Lemma III.2. *Let $n_\cdot^{up}(t)$ be the population of the unperturbed 3- system (AA, AB, BB). The 3-system (AA, AB, BB) satisfies*

$$\dot{n}_{BB} \geq \dot{n}^{up}_{BB} - (n_{aA} + n_{aB}) \left(\frac{f\left(\frac{1}{2}n_{AB}+n_{BB}\right)^2}{(n_{AA}+n_{AB}+n_{BB})^2} + c n_{BB} \right), \tag{III.95}$$

$$\dot{n}_{BB} \leq \dot{n}^{up}_{BB} + (n_{aA} + n_{aB}) \left(\frac{f\left(\frac{1}{4}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}\right)}{\Sigma_5} + c n_{BB} \right), \tag{III.96}$$

$$\dot{n}_{AB} \geq \dot{n}^{up}_{AB} - (n_{aa} + n_{aA} + n_{aB}) \left(\frac{f(n_{AB} + n_{AA})\left(\frac{1}{2}n_{AB} + n_{BB}\right)}{(n_{AA} + n_{AB} + n_{BB})^2} + c n_{AB} \right), \tag{III.97}$$

$$\dot{n}_{AB} \leq \dot{n}^{up}_{AB} + \tfrac{f}{\Sigma_5} n_{aA}\left(\tfrac{1}{2}n_{aB} + \tfrac{1}{2}n_{AB} + n_{BB}\right) + \tfrac{f}{\Sigma_5} n_{aB}\left(\tfrac{1}{2}n_{AB} + n_{AA}\right), \tag{III.98}$$

$$\dot{n}_{AA} \geq \dot{n}^{up}_{AA} - (n_{aa} + n_{aA} + n_{aB}) \left(\frac{f\left(\frac{1}{2}n_{aA}+\frac{1}{2}n_{AB}+n_{AA}\right)^2}{(n_{AA}+n_{AB}+n_{BB})2} + c n_{AA} \right), \tag{III.99}$$

$$\dot{n}_{AA} \leq \dot{n}^{up}_{AA} + \tfrac{f}{2\Sigma_5} n_{aA}\left(\tfrac{1}{2}n_{aA} + n_{AB} + n_{AA}\right). \tag{III.100}$$

Proof. We consider the rates of AA, AB and BB under the perturbation of aa, aA and aB:

$$\begin{aligned} b_{BB} =& \frac{f}{\Sigma_5}\left(\tfrac{1}{2}n_{AB} + n_{BB}\right)^2 + \frac{f n_{aB}\left(\frac{1}{4}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}\right)}{\Sigma_5} \\ =& b^{up}_{BB} - \frac{f\left(\left(\frac{1}{2}n_{AB} + n_{BB}\right)^2 (n_{aA} + n_{aB})\right)}{\Sigma_5(n_{AA} + n_{AB} + n_{BB})} + \frac{f n_{aB}\left(\frac{1}{4}n_{aB} + \frac{1}{2}n_{AB} + n_{BB}\right)}{\Sigma_5}, \end{aligned} \tag{III.101}$$

$$d_{BB} = d^{up}_{BB} + c n_{BB}(n_{aB} + n_{aA}) - \eta n_{aA} n_{BB}. \tag{III.102}$$

Thus,

$$\dot{n}_{BB} \leq \dot{n}^{up}_{BB} + \tfrac{f}{\Sigma_5} n_{aB}\left(\tfrac{1}{4}n_{aB} + \tfrac{1}{2}n_{AB} + n_{BB}\right) + \eta n_{aA} n_{BB}, \tag{III.103}$$

$$\dot{n}_{BB} \geq \dot{n}^{up}_{BB} - \frac{f(n_{aA} + n_{aB})\left(\frac{1}{2}n_{AB} + n_{BB}\right)^2}{\Sigma_5(n_{AA} + n_{AB} + n_{BB})} - c n_{BB}(n_{aA} + n_{aB}). \tag{III.104}$$

For the AB population we get:

$$\begin{aligned} b_{AB} =& \frac{2f\left(\frac{1}{2}n_{AB}+n_{BB}\right)\left(\frac{1}{2}n_{AB}+n_{AA}\right)}{\Sigma_5} - \frac{fn_{aa}n_{AA}\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{\Sigma_5\Sigma_6} + \frac{fn_{aB}n_{AA}}{2\Sigma_6} \\ &+ \frac{fn_{aB}(n_{AA}+n_{AB})}{2\Sigma_5} + \frac{fn_{aA}\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{2\Sigma_5} + \frac{fn_{aA}\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{2\Sigma_6} \\ =& b_{AB}^{up} + n_{aA}\left(\tfrac{1}{2}n_{aB}+\tfrac{1}{2}n_{AB}+n_{BB}\right)\left(\tfrac{f}{2\Sigma_5}+\tfrac{f}{2\Sigma_6}\right) + n_{aB}n_{AA}\left(\tfrac{f}{2\Sigma_5}+\tfrac{f}{2\Sigma_6}\right) + \frac{fn_{aB}n_{AB}}{2\Sigma_5} \\ &- \frac{fn_{aa}n_{AA}\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{\Sigma_5\Sigma_6} - \frac{2f\left(\frac{1}{2}n_{AB}+n_{AA}\right)\left(\frac{1}{2}n_{AB}+n_{BB}\right)(n_{aA}+n_{aB})}{\Sigma_5(n_{AA}+n_{AB}+n_{BB})}, \end{aligned} \tag{III.105}$$

$$d_{AB} = d_{AB}^{up} + cn_{AB}(n_{aB}+n_{aA}), \tag{III.106}$$

$$\dot{n}_{AB} \leq \dot{n}_{AB}^{up} + \tfrac{f}{\Sigma_5}n_{aA}\left(\tfrac{1}{2}n_{aB}+\tfrac{1}{2}n_{AB}+n_{BB}\right) + \tfrac{f}{\Sigma_5}n_{aB}n_{AA} + \tfrac{f}{2\Sigma_5}n_{aB}n_{AB}, \tag{III.107}$$

$$\begin{aligned} \dot{n}_{AB} \geq& \dot{n}_{AB}^{up} - \frac{2f\left(\frac{1}{2}n_{AB}+n_{AA}\right)\left(\frac{1}{2}n_{AB}+n_{BB}\right)}{\Sigma_5(n_{AA}+n_{AB}+n_{BB})}(n_{aA}+n_{aB}) \\ &- \frac{f\left(\frac{1}{2}n_{aB}+\frac{1}{2}n_{AB}+n_{BB}\right)}{\Sigma_5\Sigma_6}n_{aa}n_{AA} - cn_{AB}(n_{aB}+n_{aA}). \end{aligned} \tag{III.108}$$

And finally for the AA population:

$$\begin{aligned} b_{AA} =& \frac{f\left(\frac{1}{2}n_{AB}+n_{AA}\right)^2}{\Sigma_5} + \frac{fn_{aA}n_{AB}}{4\Sigma_5} + \frac{fn_{aA}\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{2\Sigma_6} \\ &- \frac{fn_{aa}n_{AA}\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{\Sigma_5\Sigma_6} \end{aligned} \tag{III.109}$$

$$\begin{aligned} =& b_{AA}^{up} - \frac{f\left(\frac{1}{2}n_{AB}+n_{AA}\right)^2(n_{aA}+n_{aB})}{\Sigma_5(n_{AA}+n_{AB}+n_{BB})} + \frac{fn_{aA}n_{AB}}{4\Sigma_5} + \frac{fn_{aA}\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{2\Sigma_6} \\ &- \frac{fn_{aa}n_{AA}\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{\Sigma_5\Sigma_6}, \end{aligned} \tag{III.110}$$

$$d_{AA} = d_{AA}^{up} + cn_{AA}(n_{aa}+n_{aA}+n_{aB}), \tag{III.111}$$

$$\dot{n}_{AA} \leq \dot{n}_{AA}^{up} + \frac{fn_{aA}n_{AB}}{4\Sigma_5} + \frac{fn_{aA}\left(\frac{1}{2}n_{aA}+n_{AA}+\frac{1}{2}n_{AB}\right)}{2\Sigma_6}, \tag{III.112}$$

$$\dot{n}_{AA} \geq \dot{n}_{AA}^{up} - \frac{f\left(\frac{1}{2}n_{aA}+\frac{1}{2}n_{AB}+n_{AA}\right)^2(n_{aa}+n_{aA}+n_{aB})}{\Sigma_5(n_{AA}+n_{AB}+n_{BB})} - cn_{AA}(n_{aa}+n_{aA}+n_{aB}). \tag{III.113}$$

□

As solutions of a dynamical system are continuous with respect to its parameters (in particular with respect to δ), the latter theorem shows that until T_2, the 3-system (AA, AB, BB) is at most perturbed by $\Theta(\delta)$. We will show that T_2 diverges with ε. Thus, for small enough δ, AB will have time to reach the small fixed value $\sqrt{\varepsilon_0} > 0$ in this phase, and we can use the asymptotic decay of the AB and AA populations which is proved in [84]. We now start to analyse the growth of the small aa-, aA- and aB populations. The sum-process Σ_5 plays a crucial role for the behaviour of the system in this phase and we need finer bounds on it:

Proposition III.4. *The sum-process* $\Sigma_5 = n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}$ *satisfies for all* $t \in [T_1, T_2]$*:*

$$\bar{n}_B - \frac{\Delta}{c\bar{n}_B} n_{AA} - \frac{\Delta^2}{c\bar{n}_B} n_{AA} \leq \Sigma_5 \leq \bar{n}_B - \frac{\Delta}{c\bar{n}_B} n_{AA} + \frac{\Delta^2}{c\bar{n}_B} n_{AA}. \tag{III.114}$$

Proof. We estimate a minorising process and a majorising process on Σ_5:

$$\begin{aligned} b_{\Sigma_5} \leq & f \frac{(n_{AA} + n_{AB} + n_{BB})(n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB})}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} \\ & + f \frac{(n_{aA} + n_{aB})(\frac{3}{4} n_{aA} + n_{AA} + \frac{3}{4} n_{aB} + n_{AB} + n_{BB})}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} + \Theta(\delta) \leq f\Sigma_5 + \Theta(\delta), \end{aligned} \tag{III.115}$$

$$\begin{aligned} b_{\Sigma_5} \geq & f \frac{(n_{AA} + n_{AB} + n_{BB})(n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB})}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} \\ & + f \frac{(n_{aA} + n_{aB})(\frac{3}{4} n_{aA} + n_{AA} + \frac{3}{4} n_{aB} + n_{AB} + n_{BB})}{n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}} - \Theta(\delta) \geq f\Sigma_5 - \Theta(\delta), \end{aligned} \tag{III.116}$$

$$d_{\Sigma_5} \leq \Sigma_5(D - \Delta + c\Sigma_5) + \Delta(n_{AA} + n_{aA}) - 2\eta n_{aA} n_{BB} + \Theta(\delta), \tag{III.117}$$

$$d_{\Sigma_5} \geq \Sigma_5(D - \Delta + c\Sigma_5) + \Delta(n_{AA} + n_{aA}) - 2\eta n_{aA} n_{BB}. \tag{III.118}$$

We get

$$\dot{\Sigma}_5 \leq -c\Sigma_5^2 + \Sigma_5(f - D + \Delta) - \Delta n_{AA} + \Theta(\delta), \tag{III.119}$$

$$\dot{\Sigma}_5 \geq -c\Sigma_5^2 + \Sigma_5(f - D + \Delta) - \Delta n_{AA} - \Theta(\delta). \tag{III.120}$$

We start with the proof of the upper bound. We use Lemma III.1 and show that when Σ_5 reaches the upper-bound, it decays faster than the latter. Using (III.119) we compute $\dot{\Sigma}_5$ at the bound. Note that if $\Sigma_5 \leq \bar{n}_B - \frac{\Delta}{c\bar{n}_B} n_{AA} + \frac{\Delta^2}{c\bar{n}_B} n_{AA}$, then $\Sigma_5^2 \leq \bar{n}_B^2 - \frac{2\Delta}{c} n_{AA} + \frac{\Delta^2}{c^2\bar{n}_B^2} n_{AA}^2 + \frac{2\Delta^2}{c} n_{AA} + \Theta(\Delta^4) n_{AA}^2$, thus

$$\dot{\Sigma}_5 \leq -\Delta^2 n_{AA} - \frac{\Delta^2}{c\bar{n}_B^2} n_{AA}^2 + \Theta(\delta) < 0. \tag{III.121}$$

It is left to show that $\dot{\Sigma}_5 \leq -\frac{\Delta}{c\bar{n}_B} \dot{n}_{AA} + \frac{\Delta^2}{c\bar{n}_B} \dot{n}_{AA}$. Since we already know (cf. Lemma III.2) that (AA, AB, BB) behaves like a 3-system with $\Theta(\delta)$ perturbations, then AA is decreasing, $\dot{n}_{AA} \leq 0$, this finishes the proof of the upper bound.

Now we check the lower bound. If $\Sigma_5 \geq \bar{n}_B - \frac{\Delta}{c\bar{n}_B} n_{AA} - \frac{\Delta^2}{c\bar{n}_B} n_{AA}$ then $\Sigma_5^2 \geq \bar{n}_B^2 - \frac{2\Delta}{c} n_{AA} - \frac{\Delta^2}{c^2\bar{n}_B^2} n_{AA} - \frac{2\Delta^2}{c} n_{AA}$. Using (III.120), the derivative of Σ_5 at the lower bound is thus lower bounded by

$$\dot{\Sigma}_5 \geq \Delta^2 n_{AA} - \frac{\Delta^2}{c\bar{n}_B^2} n_{AA} - \Theta(\delta) \geq \Delta^2 n_{AA} \left(1 - \frac{1}{c\bar{n}_B}\right) - \Theta(\delta) > 0. \tag{III.122}$$

By Lemma III.1, it is enough to show that at the lower bound $\dot{\Sigma}_5 \geq -\frac{\Delta}{c\bar{n}_B} \dot{n}_{AA}$. For this we calculate a majorising process on AA:

$$b_{AA} \leq \frac{f}{\Sigma_5} n_{AA}(n_{AA} + n_{AB}) + \frac{f}{4\Sigma_5} n_{AB}^2 + \Theta(\delta), \tag{III.123}$$

$$d_{AA} \geq f n_{AA}, \tag{III.124}$$

$$\dot{n}_{AA} \leq -\frac{f}{\Sigma_5} n_{AA} n_{BB} + \frac{f}{4\Sigma_5} n_{AB}^2 + \Theta(\delta). \tag{III.125}$$

Hence we have to show that $\Delta^2 n_{AA}\left(1-\frac{1}{c\bar{n}_B}\right)-\Theta(\delta)\geq\frac{\Delta f}{c\bar{n}_B\bar{n}_A}\left(n_{AA}n_{BB}-\frac{1}{4}n_{AB}^2\right)-\Theta(\delta\Delta)$, in the case $n_{AA}n_{BB}>\frac{1}{4}n_{AB}^2$. This is equivalent to show that $\chi:=n_{AA}n_{BB}-\frac{1}{4}n_{AB}^2\leq\frac{\Delta\bar{n}_A}{f}\left(c\bar{n}_B-1\right)n_{AA}$. For this we use once again Lemma III.1 and estimate the derivative of χ from above with the help of minorising processes on AA and BB and a majorising process on AB:

$$b_{AA}\geq\tfrac{f}{\Sigma_5}n_{AA}(n_{AA}+n_{AB})+\tfrac{f}{4\Sigma_5}n_{AB}^2-\Theta(\delta), \tag{III.126}$$
$$d_{AA}\leq(f+\Delta)n_{AA}+\Theta(\delta), \tag{III.127}$$
$$\dot{n}_{AA}\geq-\tfrac{f}{\Sigma_5}n_{AA}n_{BB}-\Delta n_{AA}+\tfrac{f}{4\Sigma_5}n_{AB}^2-\Theta(\delta). \tag{III.128}$$
$$b_{BB}\geq\tfrac{f}{\Sigma_5}n_{BB}(n_{AB}+n_{BB})+\tfrac{f}{4\Sigma_5}n_{AB}^2-\Theta(\delta), \tag{III.129}$$
$$d_{BB}\leq fn_{BB}, \tag{III.130}$$
$$\dot{n}_{BB}\geq-\tfrac{f}{\Sigma_5}n_{AA}n_{BB}+\tfrac{f}{4\Sigma_5}n_{AB}^2+\Theta(\delta). \tag{III.131}$$
$$b_{AB}\leq\tfrac{f}{\Sigma_5}n_{AB}\left(n_{AA}+\tfrac{1}{2}n_{AB}+n_{BB}\right)+\tfrac{2f}{\Sigma_5}n_{AA}n_{BB}+\Theta(\delta), \tag{III.132}$$
$$d_{AB}\geq(f-\Delta)n_{AB}, \tag{III.133}$$
$$\dot{n}_{AB}\leq\tfrac{2f}{\Sigma_5}n_{AA}n_{BB}-\tfrac{f}{2\Sigma_5}n_{AB}^2+\Delta n_{AB}+\Theta(\delta). \tag{III.134}$$

The derivative is given by:

$$\begin{aligned}\dot{\chi}&=\dot{n}_{AA}n_{BB}+n_{AA}\dot{n}_{BB}-\tfrac{1}{2}\dot{n}_{AB}n_{AB}\\&\leq-f\chi+\Theta(\delta).\end{aligned} \tag{III.135}$$

At the upper bound we get:

$$\dot{\chi}\leq-\Delta\bar{n}_A(c\bar{n}_B-1)n_{AA}+\Theta(\delta)<0. \tag{III.136}$$

It is left to show that $\dot{\chi}\leq\frac{\Delta\bar{n}_A}{f}(c\bar{n}_B-1)\dot{n}_{AA}$. Using the minorising process $\dot{n}_{AA}\geq-\Delta n_{AA}-\frac{f}{\bar{n}_A}\chi-\Theta(\delta)$ we show that

$$0\leq(f-2\Delta)\chi-\tfrac{\Delta\bar{n}_A}{\bar{n}_B}(c\bar{n}_B-1)\chi-\tfrac{\Delta^2\bar{n}_A}{f}(c\bar{n}_B-1)n_{AA}-\Theta(\delta). \tag{III.137}$$

An easy calculation proves this fact and finishes the proof of the lower bound.

□

Lemma III.3. *For $t\in[T_1,T_2]$ and for Δ sufficiently small it holds,*

$$\dot{\Sigma}_{aA,aB}\geq-\Theta(\Delta)\Sigma_{aA,aB}. \tag{III.138}$$

Proof. Using Propositions III.4, we have the following bound on the process:

$$\begin{aligned}b_{\Sigma_{aA,aB}}&\geq f\frac{n_{aA}(\frac{1}{2}n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB})+n_{aB}(n_{AA}+\frac{1}{2}n_{aB}+n_{AB}+n_{BB})}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}-\Theta(\delta n_{aA})\\&\geq f\Sigma_{aA,aB}-\Theta(\delta n_{aA}),\end{aligned} \tag{III.139}$$
$$\begin{aligned}d_{\Sigma_{aA,aB}}&=\Sigma_{aA,aB}(D-\Delta+c\Sigma_5)-\eta n_{aA}n_{BB}+\Delta n_{aA}+cn_{aA}n_{aa}\\&\leq f\Sigma_{aA,aB}-n_{aA}(\eta n_{BB}-\Delta)+\Theta(\Delta^2n_{AA})\Sigma_{aA,aB}^2,\end{aligned} \tag{III.140}$$
$$\begin{aligned}\dot{\Sigma}_{aA,aB}&\geq n_{aA}(\eta n_{BB}-\Delta-\Theta(\delta))\Theta(\Delta^2n_{AA})\Sigma_{aA,aB}^2\\&\geq n_{aA}(-\Delta-\Theta(\delta))-\Theta(\delta\Delta^2n_{AA})\Sigma_{aA,aB}\\&\geq\Sigma_{aA,aB}(-\Delta-\Theta(\delta)).\end{aligned} \tag{III.141}$$

□

Lemma III.4. *For all $t \in [T_1, T_2]$ the aa population is bounded by*

$$\frac{f}{4\bar{n}_B(f+\Delta)}\Sigma^2_{aA,aB} \leq n_{aa} \leq \frac{f}{\bar{n}_A(D+\Delta)}\Sigma^2_{aA,aB}. \tag{III.142}$$

Observe that this implies $T_2 = T^{aA=\delta AA} \wedge T^{aB=\delta AB}$.

Proof. First observe that the inequality is satisfied at $t = T_1$. We start with the upper bound and show that n_{aa} would decrease at this bound. For this we estimate a majorising process on aa:

$$b_{aa} \leq \tfrac{f}{n_{aa}+n_{aA}+n_{AA}} n_{aa}\left(\tfrac{1}{2}n_{aA} + n_{aa}\right) + \tfrac{f}{4\Sigma_5}\Sigma^2_{aA,aB} + \tfrac{f}{2\Sigma_5} n_{aA}n_{aa}, \tag{III.143}$$

$$d_{aa} \geq n_{aa}(D+\Delta), \tag{III.144}$$

$$\dot{n}_{aa} \leq \tfrac{f}{n_{aa}+n_{aA}+n_{AA}} n_{aa}^2 + \tfrac{f}{n_{aa}+n_{aA}+n_{AA}} n_{aa}n_{aA} + \tfrac{f}{4\Sigma_5}\Sigma^2_{aA,aB} - n_{aa}(D+\Delta). \tag{III.145}$$

We calculate the slope of this process at the upper bound:

$$\dot{n}_{aa} \leq \tfrac{f}{4\Sigma_5}\Sigma^2_{aA,aB} - \tfrac{f}{\bar{n}_A}\Sigma^2_{aA,aB} + \Theta(\Sigma^2_{aA,aB}n_{aA}) \leq -\tfrac{3f-\Theta(\delta)}{4\bar{n}_A}\Sigma^2_{aA,aB} < 0. \tag{III.146}$$

By Lemma III.1, to ensure that (III.142) stays an upper bound it is enough to show that

$$-\tfrac{3f-\Theta(\delta)}{4\bar{n}_A}\Sigma^2_{aA,aB} \leq \tfrac{2f}{\bar{n}_A(D+\Delta)}\dot{\Sigma}_{aA,aB}\Sigma_{aA,aB}. \tag{III.147}$$

This is a consequence of Lemma III.3.
For the lower bound we proceed similarly. This time, with the knowledge of the upper bound, we estimate a minorising process on aa:

$$b_{aa} \geq \tfrac{f}{\Sigma_5}\Sigma^2_{aA,aB} - \Theta\left(\Sigma^3_{aA,aB}\right), \tag{III.148}$$

$$d_{aa} \leq n_{aa}(f+\Delta), \tag{III.149}$$

$$\dot{n}_{aa} \geq \tfrac{f}{\bar{n}_B}\Sigma^2_{aA,aB} - n_{aa}(f+\Delta) - \Theta\left(\Sigma^3_{aA,aB}\right). \tag{III.150}$$

At the lower bound the process increases:

$$\dot{n}_{aa} \geq \left(\tfrac{f}{\bar{n}_B} - \tfrac{f}{4\bar{n}_B}\right)\Sigma^2_{aA,aB} - \Theta(\Sigma^3_{aA,aB}) = \tfrac{3f}{4\bar{n}_B}\Sigma^2_{aA,aB} - \Theta\left(\Sigma^3_{aA,aB}\right) > 0. \tag{III.151}$$

By Lemma III.1, it is left to show that $\dot{n}_{aa} \geq \frac{f}{2\bar{n}_B(f+\Delta)}\dot{\Sigma}_{aA,aB}\Sigma_{aA,aB}$. Thus we have to calculate a majorising process on $\Sigma_{aA,aB}$:

$$b_{\Sigma_{aA,aB}} \leq f\Sigma_{aA,aB} + \Theta\left(\Sigma^2_{aA,aB}\right), \tag{III.152}$$

$$\begin{aligned} d_{\Sigma_{aA,aB}} &\geq (f-\Delta)\Sigma_{aA,aB} + n_{aA}(\Delta - \eta n_{BB}) \\ &\geq (f-\Delta)\Sigma_{aA,aB} - (f-D)\Sigma_{aA,aB} \\ &= (D-\Delta)\Sigma_{aA,aB}, \end{aligned} \tag{III.153}$$

$$\dot{\Sigma}_{aA,aB} \leq (f-D+\Delta)\Sigma_{aA,aB} + \Theta\left(\Sigma^2_{aA,aB}\right). \tag{III.154}$$

Thus we get

$$\begin{aligned} \tfrac{f(f-D+\Delta)}{2\bar{n}_B(f+\Delta)}\Sigma^2_{aA,aB} - \tfrac{3f}{4\bar{n}_B}\Sigma^2_{aA,aB} + \Theta(\Sigma^3_{aA,aB}) &= -\tfrac{f}{2\bar{n}_B}\Sigma^2_{aA,aB}\left(\tfrac{3}{2} - \tfrac{f-D+\Delta}{f+\Delta}\right) + \Theta\left(\Sigma^3_{aA,aB}\right) \\ &= -\tfrac{f}{2\bar{n}_B}\Sigma^2_{aA,aB}\tfrac{f+2D+\Delta}{2(f+\Delta)} + \Theta\left(\Sigma^3_{aA,aB}\right) < 0. \end{aligned} \tag{III.155}$$

This finishes the proof of the lower bound. □

Let

$$T_{\equiv} = \inf\{t > T_1 : n_{aA}(t) = n_{aB}(t)\}. \tag{III.156}$$

Proposition III.5. *For all $t \in [T_1, T_=]$ it holds*

$$n_{aB} \leq n_{aA} = \Theta(\varepsilon). \tag{III.157}$$

Proof. In this time interval the newborns of genotype aA are in majority produced by reproductions of a population of order one, namely AB or AA, with the population aA. Since n_{aA} feels competition from a macroscopic population (AA, AB or BB) the aA population stays of order $\Theta(\varepsilon)$. We make this more rigorous. To show this we consider a majorising process on aA and use Proposition III.4, and Lemma III.4:

$$b_{aA} \leq f n_{aA} - \tfrac{f}{\Sigma_5} n_{aA}(n_{BB} + \tfrac{1}{2} n_{AB}) + \tfrac{f}{2\Sigma_5} n_{aB}(2n_{AA} + n_{AB}) + \Theta\left(\Sigma^2_{aA,aB}\right), \tag{III.158}$$

$$d_{aA} \geq n_{aA}(f + \Delta - \tfrac{\Delta}{\bar{n}_B} n_{AA} - \eta n_{BB} - \Theta\left(\Delta^2 n_{AA}\right)), \tag{III.159}$$

$$\begin{aligned}
\dot{n}_{aA} \leq & -n_{aA}\Big(n_{BB}\tfrac{f-\eta\Sigma_5}{\Sigma_5} + \tfrac{f}{2\Sigma_5} n_{AB} + \Delta\Big(1 - \tfrac{n_{AA}}{\bar{n}_B}\Big) - \Theta\big(\Delta^2 n_{AA}\big)\Big) + \tfrac{f}{\Sigma_5} n_{aB}\big(\tfrac{1}{2} n_{AB} + n_{AA} + \Theta(\delta)\big) \\
\leq & -n_{aA}\Big(n_{BB}\tfrac{D+\Delta}{\Sigma_5} + \tfrac{f}{2\Sigma_5} n_{AB} + \Delta\Big(1 - \tfrac{n_{AA}}{\bar{n}_B}\Big) - \Theta\big(\Delta^2 n_{AA}\big)\Big) + \tfrac{f}{\Sigma_5} n_{aB}\big(\tfrac{1}{2} n_{AB} + n_{AA} + \Theta(\delta)\big) \\
\leq & -n_{aA}\Big(\tfrac{f}{\Sigma_5}\Big(\tfrac{D+\Delta}{f} n_{BB} + \tfrac{1}{2} n_{AB}\Big) + \Delta\Big(1 - \tfrac{n_{AA}}{\bar{n}_B}\Big) - \Theta\big(\Delta^2 n_{AA}\big)\Big) + \tfrac{f}{\Sigma_5} n_{aB}\big(\tfrac{1}{2} n_{AB} + n_{AA} + \Theta(\delta)\big).
\end{aligned} \tag{III.160}$$

By Proposition III.2 and [84] there exists a time $t_0 = \Theta(1)$ such that the expression in the first bracket becomes bigger than the expression in the second bracket. Thus n_{aA} decreases after t_0 and since aA does not exceed $\Theta(\varepsilon)$ until t_0 it will stay smaller or equal to $\Theta(\varepsilon)$ until $T_=$. □

We show that as soon as aB crosses aA the BB population is already bigger than or equal to the AA population. First we estimate a upper bound for aB:

Lemma III.5. *For all $t \in [T_1, T_2]$ the aB population is upper bounded by*

$$n_{aB} \leq \frac{n_{AB} + 2n_{BB} + \frac{2\Delta}{c}}{n_{AB} + 2n_{AA}} n_{aA} \equiv C(t) n_{aA}. \tag{III.161}$$

Proof. First observe that the bound is fulfilled at $t = T_1$. Similarly to the proof of Lemma III.4 we estimate a majorising process on aB given by:

$$\dot{n}_{aB} \leq -n_{aB}\left(\tfrac{f}{2\Sigma_5}(n_{AB} + 2n_{AA}) - \tfrac{\Delta}{\bar{n}_B} n_{AA} - \Theta(\Delta^2 n_{AA})\right) + n_{aA}\tfrac{f}{2\Sigma_5}(n_{AB} + 2n_{BB} + \Theta(\delta)). \tag{III.162}$$

By Lemma III.1, we have to show that as soon as aB reaches the upper bound it decreases faster than the bound, thus we calculate the slope of the majorising process at this value:

$$\begin{aligned}
\dot{n}_{aB} \leq & -\tfrac{f}{2\Sigma_5}\left(n_{AB} + 2n_{BB} + \tfrac{2\Delta}{c} - \Theta(\Delta^2 n_{AA})\right) n_{aA} + \tfrac{\Delta(n_{AB} + 2n_{BB} + 2\Delta/c)}{\bar{n}_B(n_{AB} + 2n_{AA})} n_{AA} n_{aA} \\
& + \tfrac{f}{2\Sigma_5}(n_{AB} + 2n_{BB}) n_{aA} \\
\leq & -\tfrac{\Delta f - \Theta(\Delta^2 n_{AA})}{c\Sigma_5} n_{aA} + \tfrac{\Delta}{\Sigma_5}\left(\tfrac{1}{2} n_{AB} + n_{BB} + \tfrac{\Delta}{c}\right) n_{aA} \\
\leq & \tfrac{\Delta + \Theta(\Delta^2 n_{AA})}{\Sigma_5} n_{aA}\left(\bar{n}_B + \tfrac{\Delta}{c} - \tfrac{f}{c}\right) \\
\leq & -\tfrac{\Delta + \Theta(\Delta^2 n_{AA})}{c\Sigma_5}(D - 2\Delta) n_{aA} \leq 0.
\end{aligned} \tag{III.163}$$

We have to show that $\dot{n}_{aB} \leq C(t)\dot{n}_{aA} + \dot{C}(t)n_{aA}$. Since the 3-system converges towards $(0, 0, \bar{n}_B)$, $C(t)$ is a monotone increasing function and hence $\dot{C}(t) \geq 0$. Thus if we can show that $\dot{n}_{aB} \leq C(t)\dot{n}_{aA}$ we are done. For this we have to calculate the slope of the minorising process on aA when aB would reach the upper bound. This process is given by:

$$\dot{n}_{aA} \geq -n_{aA}\left(\tfrac{f}{2} + \Delta - \eta n_{BB} + \tfrac{f}{2\Sigma_5}(n_{BB} - n_{AA}) + \Theta(\delta)\right) + n_{aB}\tfrac{f}{2\Sigma_5}(n_{AB} + 2n_{AA}). \tag{III.164}$$

The slope at the upper bound is:

$$\begin{aligned}\dot{n}_{aA} &\geq -n_{aA}\left(\tfrac{f}{2} + \Delta - \eta n_{BB} + \tfrac{f}{2\Sigma_5}(n_{BB} - n_{AA}) - \tfrac{f}{2\Sigma_5}\left(n_{AB} + 2n_{BB} + \tfrac{2\Delta}{c} + \Theta(\delta)\right)\right) \\ &\geq -n_{aA}\left(\Delta - \eta n_{BB} - \tfrac{\Delta f}{c\Sigma_5} + \Theta(\delta)\right) \\ &\geq n_{aA}\left(\Delta\tfrac{D-\Delta}{c\Sigma_5} + \eta n_{BB} - \Theta(\delta)\right) \geq 0. \end{aligned} \tag{III.165}$$

Since $C(t) > 0$ this finishes the proof. □

Lemma III.6. *We have $T_= \leq T_2$. Moreover it holds,*

$$n_{AA}(T_=) \leq n_{BB}(T_=) + \Theta(\Delta). \tag{III.166}$$

Proof. We first show that $T_= < T_2$. Using Proposition III.4 we construct two processes that provide an upper bound and a lower bound on n_{aB}, respectively:

$$b_{aB} \geq f n_{aB} - \tfrac{f}{\Sigma_5} n_{aB}(\tfrac{1}{2}n_{AB} + n_{AA}) + \tfrac{f}{\Sigma_5} n_{aA}(\tfrac{1}{2}n_{AB} + n_{BB} - \Theta(\delta^2)), \tag{III.167}$$

$$b_{aB} \leq f n_{aB} - \tfrac{f}{\Sigma_5} n_{aB}(\tfrac{1}{2}n_{AB} + n_{AA}) + \tfrac{f}{\Sigma_5} n_{aA}(\tfrac{1}{2}n_{AB} + n_{BB} + \Theta(\delta)), \tag{III.168}$$

$$d_{aB} \leq n_{aB} f, \tag{III.169}$$

$$d_{aB} \geq n_{aB}(f - \tfrac{\Delta}{\bar{n}_B} n_{AA} - \Theta(\Delta^2 n_{AA})), \tag{III.170}$$

$$\dot{n}_{aB} \leq -n_{aB}\left(\frac{f(\frac{1}{2}n_{AB} + n_{AA})}{\Sigma_5} - \tfrac{\Delta}{\bar{n}_B} n_{AA} - \Theta(\Delta^2 n_{AA})\right) + n_{aA}\frac{f(\frac{1}{2}n_{AB} + n_{BB} + \Theta(\delta))}{\Sigma_5}, \tag{III.171}$$

$$\dot{n}_{aB} \geq -n_{aB}\frac{f(\frac{1}{2}n_{AB} + n_{AA})}{\Sigma_5} + n_{aA}\frac{f(\frac{1}{2}n_{AB} + n_{BB} - \Theta(\delta^2))}{\Sigma_5}. \tag{III.172}$$

We first show that $T_= < \infty$. We know that the 3-system (AA, AB, BB) converges to $(0, 0, \bar{n}_B)$ and that $n_{aB} \leq n_{aA} = \Theta(\varepsilon)$ (Proposition III.5), for $t \leq T_=$. We consider the worst case and assume that $n_{aB} < n_{aA}$ then we get from (III.172) that at some time t_0, where $n_{AB} + 2n_{BB}$ is already macroscopic,

$$\dot{n}_{aB} \geq \Theta(\varepsilon), \quad n_{aB} \geq \Theta(\varepsilon)t. \tag{III.173}$$

Thus the time aB needs to reach $n_{aA} = \Theta(\varepsilon)$ is of order $\Theta(1)$. This time is shorter than $T_{aA=\delta AA}$. Indeed, suppose the contrary, then by Proposition III.5 n_{aA} does not exceed $\Theta(\varepsilon)$ before T_2, and thus $T^{aA=\delta AA} \geq T^{AA}_{\Theta(\varepsilon/\delta)} = \Theta\left((\delta/\varepsilon)^2\right)$ which diverges with ε. A similar reasoning shows that $T_= < T^{aB=\delta AB}$. Hence $T_= < T_2$.

It is left to show that $n_{AA}(T_=) \leq n_{BB}(T_=) + \Theta(\Delta)$. From Lemma III.5 we deduce that at $T_=$ it holds

$$\begin{aligned} \tfrac{1}{2}n_{AB} + n_{AA} &\leq \tfrac{1}{2}n_{AB} + n_{BB} + \tfrac{\Delta}{c} \\ n_{AA} &\leq n_{BB} + \Theta(\Delta). \end{aligned} \tag{III.174}$$

□

Lemma III.7. *For all $t \in [T_1, T_2]$ the AB population is bounded by*

1. $n_{AB} \geq 2\sqrt{\bar{n}_B n_{AA}} - 2n_{AA}\left(1 + \frac{\Delta}{c\bar{n}_B}\right)$,
2. $n_{AB} \leq 2\sqrt{\bar{n}_B n_{AA}\left(1 + \frac{\Delta}{f}\right)} - 2n_{AA}$.

Proof. 1. The proof works like the one of Lemma III.4. First observe that the bound holds at $t = T_1$. Then we calculate a minorising process on AB:

$$b_{AB} \geq f(2n_{AA} + n_{AB}) - \tfrac{f}{\Sigma_5}(2n_{AA} + n_{AB})(n_{AA} + \tfrac{1}{2}n_{AB} + \Theta(\delta^2)), \tag{III.175}$$

$$d_{AB} \leq f n_{AB}, \tag{III.176}$$

$$\dot{n}_{AB} \geq -n_{AB}\tfrac{f}{\Sigma_5}\left(\tfrac{1}{2}n_{AB} + n_{AA} + \Theta(\delta^2)\right) + 2f n_{AA} - \tfrac{2f}{\Sigma_5}n_{AA}\left(\tfrac{1}{2}n_{AB} + n_{AA} + \Theta(\delta^2)\right). \tag{III.177}$$

We use Proposition III.4 and show that this minorising process would increase quicker than the lower-bound if AB reaches it:

$$\begin{aligned} \dot{n}_{AB} \geq& -\tfrac{2f}{\Sigma_5}\left(\sqrt{\bar{n}_B n_{AA}} - n_{AA}\left(1 + \tfrac{\Delta}{c\bar{n}_B}\right)\right)\left(\sqrt{\bar{n}_B n_{AA}} - \tfrac{\Delta}{c\bar{n}_B}n_{AA} + \Theta(\delta^2)\right) \\ &+ 2f n_{AA} - \tfrac{2f}{\Sigma_5}n_{AA}\left(\sqrt{\bar{n}_B n_{AA}} - \tfrac{\Delta}{c\bar{n}_B}n_{AA}\right) \\ \geq& \tfrac{2f}{\Sigma_5}\tfrac{\Delta}{c\bar{n}_B}n_{AA}(2\sqrt{\bar{n}_B n_{AA}} - n_{AA}) - \Theta(\Delta^2) > 0. \end{aligned} \tag{III.178}$$

It is left to show that at the lower bound,

$$\dot{n}_{AB} \geq \frac{\bar{n}_B \dot{n}_{AA}}{\sqrt{\bar{n}_B n_{AA}}} - 2\dot{n}_{AA}\left(1 + \tfrac{\Delta}{c\bar{n}_B}\right). \tag{III.179}$$

For this we calculate a majorising process on AA:

$$b_{AA} \leq \tfrac{f}{\Sigma_5}n_{AA}(n_{AA} + n_{AB}) + \tfrac{f}{4\Sigma_5}n_{AB}^2 + \Theta(\delta), \tag{III.180}$$

$$d_{AA} \geq f n_{AA}, \tag{III.181}$$

$$\dot{n}_{AA} \leq -n_{AA}\left(f - \tfrac{f}{\Sigma_5}(n_{AA} + n_{AB})\right) + \tfrac{f}{4\Sigma_5}n_{AB}^2 + \Theta(\delta). \tag{III.182}$$

If we now insert the lower bound and use Proposition III.4 we get

$$\dot{n}_{AA} \leq -\tfrac{f}{\Sigma_5}\tfrac{\Delta}{c\bar{n}_B}n_{AA}(\sqrt{\bar{n}_B n_{AA}} - n_{AA}) + \Theta(\Delta^2) < 0. \tag{III.183}$$

Thus (III.179) is fulfilled.

2. First, observe that the upper bound is fullfiled at $t = T_1$. We then have to estimate a majorising process on AB:

$$b_{AB} \leq f(2n_{AA} + n_{AB}) - f(2n_{AA} + n_{AB})\frac{n_{AA} + \frac{1}{2}n_{AB}}{\Sigma_5} + \Theta(\delta), \tag{III.184}$$

$$d_{AB} \geq n_{AB}(D - \Delta + c\bar{n}_B - \tfrac{\Delta}{\bar{n}_B}n_{AA} - \Theta(\Delta^2 n_{AA})) \tag{III.185}$$

$$\geq n_{AB}(f - \tfrac{\Delta}{\bar{n}_B}n_{AA} - \Theta(\Delta^2 n_{AA})), \tag{III.186}$$

$$\dot{n}_{AB} \leq -\tfrac{f}{2\bar{n}_B}n_{AB}^2 - n_{AB}\tfrac{2f-\Delta}{\bar{n}_B}n_{AA} + 2fn_{AA} - \tfrac{2f}{\bar{n}_B}n_{AA}^2 + \Theta(\Delta^2 n_{AA}). \tag{III.187}$$

As before we calculate the slope of this majorising process if it would reach the upper bound:

$$\dot{n}_{AB} \leq -\tfrac{2\Delta}{\bar{n}_B}n_{AA}^2 + \Theta(\Delta^2 n_{AA}) < 0. \tag{III.188}$$

By Lemma III.1 we have to show that

$$\dot{n}_{AB} \leq \dot{n}_{AA}\left(\tfrac{\bar{n}_B(1+\Delta/f)}{\sqrt{\bar{n}_B n_{AA}(1+\Delta/f)}} - 2\right). \tag{III.189}$$

For this we calculate the slope of a minorising process on AA given by

$$\dot{n}_{AA} \geq -n_{AA}\left(f - \tfrac{f}{\Sigma_5}(n_{AA} + n_{AB}) + \Delta + \Theta(\delta^2)\right) + \tfrac{f}{4\Sigma_5}n_{AB}^2. \tag{III.190}$$

At the upper bound AA would start to increases:

$$\dot{n}_{AA} \geq \tfrac{\Delta}{\bar{n}_B}n_{AA}^2 - \Theta(\delta^2) > 0. \tag{III.191}$$

Thus we get

$$\dot{n}_{AA}\left(\tfrac{\bar{n}_B(1+\Delta/f)}{\sqrt{\bar{n}_B n_{AA}(1+\Delta/f)}} - 2\right) - \dot{n}_{AB} \geq \tfrac{\Delta(1+\Delta/f)}{\sqrt{\bar{n}_B n_{AA}(1+\Delta/f)}}n_{AA}^2 - \Theta(\Delta^2 n_{AA}) > 0. \tag{III.192}$$

This finishes the proof.

□

The following Proposition is a statement for the 3-system (AA, AB, BB) but it holds also true until T_2 in the 6-system (aa, aA, AA, aB, AB, BB) for $\delta < \Delta$.

Proposition III.6. *The maximal value n_{AB}^{max} of n_{AB} in $[T_1, T_2]$ is bounded by*

$$\frac{\bar{n}_B}{2} - \Theta(\Delta) \leq n_{AB}^{max} \leq \frac{\bar{n}_B}{2} + \Theta(\Delta). \tag{III.193}$$

Moreover, let T_{AB}^{max} be the time when n_{AB} takes on its maximum, then n_{AA} and n_{BB} are bounded by

$$\frac{\bar{n}_B}{4} - \Theta(\Delta) \leq n_{AA}(T_{AB}^{max}) \leq \frac{\bar{n}_B}{4} + \Theta(\Delta), \tag{III.194}$$

$$\frac{\bar{n}_B}{4} - \Theta(\Delta) \leq n_{BB}(T_{AB}^{max}) \leq \frac{\bar{n}_B}{4} + \Theta(\Delta). \tag{III.195}$$

Proof. From Lemma III.7 (1) we get that

$$n_{AB} \geq 2\sqrt{\bar{n}_B n_{AA}} - 2n_{AA}\left(1 + \tfrac{\Delta}{c\bar{n}_B}\right), \tag{III.196}$$

We look for the value of AA where the expression on the right hand side takes on its minimum, thus we have to derivate n_{AA} and set it to zero:

$$\frac{\bar{n}_B}{\sqrt{\bar{n}_B n_{AA}}} - \left(2 + \tfrac{\Delta}{c\bar{n}_B}\right) = 0 \tag{III.197}$$

$$\bar{n}_B^2 = \left(4 - 4\tfrac{\Delta}{c\bar{n}_B} + \Theta(\Delta^2)\right)\bar{n}_B n_{AA} \tag{III.198}$$

$$\frac{\bar{n}_B}{4} - \Theta(\Delta) = n_{AA}. \tag{III.199}$$

If we insert this in n_{AB} we get the lower bound:

$$n_{AB} \geq \tfrac{\bar{n}_B}{2} + \Theta(\Delta). \tag{III.200}$$

For the upper bound on n_{AB} we proceed similarly. Form Lemma III.7 (2) we get

$$n_{AB} \leq 2\sqrt{\bar{n}_B n_{AA}\left(1 + \tfrac{\Delta}{f}\right)} - 2n_{AA}. \tag{III.201}$$

Setting the derivation of the rhs to zero gives:

$$0 = \frac{\bar{n}_B\left(1+\frac{\Delta}{f}\right)}{\sqrt{\bar{n}_B n_{AA}\left(1+\frac{\Delta}{f}\right)}} - 2 \tag{III.202}$$

$$n_{AA} = \frac{\bar{n}_B}{4} + \Theta(\Delta). \tag{III.203}$$

Finally we get

$$n_{AB} \leq \tfrac{\bar{n}_B}{2} - \Theta(\Delta) \quad \text{and} \quad n_{AA} = \tfrac{\bar{n}_B}{4} - \Theta(\Delta). \tag{III.204}$$

□

Remark III.4. *Note that $n_{AA} = n_{BB} \pm \Theta(\Delta) = \frac{\bar{n}_B}{4} \pm \Theta(\Delta)$ as soon as n_{AB} reaches its maximal value.*

Proposition III.7. *For all $t \in [T_1, T_2]$,*

$$n_{aA} \leq \Theta(\varepsilon) \vee n_{aB}. \tag{III.205}$$

Proof. For $t \leq T_=$ this follows from Proposition III.5. For $t > T_=$ we show this by constructing a majorising process on $n_{aA}(t)$:

$$\begin{aligned} b_{aA} \leq & f\frac{(n_{aA}+n_{aB})(2n_{AA}+n_{AB}+\Theta(\delta))}{2\Sigma_5} \\ \leq & \tfrac{f+\Theta(\delta)}{2}(n_{aA}+n_{aB}) + \frac{f(n_{AA}-n_{BB})}{2\bar{n}_A}(n_{aA}+n_{aB}), \end{aligned} \tag{III.206}$$

$$\begin{aligned} d_{aA} \geq & n_{aA}\left(D + c\bar{n}_B - \tfrac{\Delta}{\bar{n}_B}n_{AA} - \eta n_{BB} - \Theta(\Delta^2 n_{AA})\right) \\ \geq & n_{aA}(f - \eta n_{BB}), \end{aligned} \tag{III.207}$$

$$\dot{n}_{aA} \leq - n_{aA}\left(\tfrac{f}{2} - \tfrac{f(n_{AA}-n_{BB})}{2\bar{n}_A} - \eta n_{BB} - \Theta(\delta)\right) + n_{aB}\left(\tfrac{f}{2} + \tfrac{f(n_{AA}-n_{BB}+\Theta(\delta))}{2\bar{n}_A}\right). \tag{III.208}$$

By Lemma III.1, it is left to show that $\dot{n}_{aA} \leq \dot{n}_{aB}$ whenever $n_{aA} = n_{aB}$. At this upper bound we have $\dot{n}_{aA} \leq n_{aB}(\frac{f}{\bar{n}_A}(n_{AA} - n_{BB}) + \eta n_{BB} + \Theta(\delta))$. We now calculate a minorising process on n_{aB}:

$$b_{aB} \geq \tfrac{f}{2\Sigma_5}(n_{aA} + n_{aB})(n_{aB} + n_{AB} + 2n_{BB}), \tag{III.209}$$

$$d_{aB} \leq n_{aB}(D - \Delta + c\bar{n}_B) = f n_{aB}, \tag{III.210}$$

$$\dot{n}_{aB} \geq \tfrac{f}{2\Sigma_5} n_{aA}(n_{aB} + n_{AB} + 2n_{BB}) - \tfrac{f}{2\Sigma_5} n_{aB}(2n_{AA} + 2n_{aA} - n_{aB} - n_{AB}). \tag{III.211}$$

Thus $\dot{n}_{aB} \geq \frac{f}{\Sigma_5} n_{aB}(n_{BB} - n_{AA} + n_{AB})$ whenever $n_{aA} = n_{aB}$, and hence $\dot{n}_{aB} - \dot{n}_{aA} \geq \frac{f}{\bar{n}_A} n_{aB}(2n_{BB} - 2n_{AA} + \eta n_{BB} - \Theta(\Delta)) > 0$ by Proposition III.6. This finishes the proof. □

Now we show that the time $T^{aA=\delta AA}$ is finite and prove that it is smaller than or equal to $T^{aB=\delta BB}$. To estimate the order of magnitude of the time T_2 we need bounds on n_{aA} which depends on $\Sigma_{aA,aB}$.

Lemma III.8. *For all $t \in [T_1, T_2]$ the aA population is bounded by*

$$\frac{f(n_{AB} + 2n_{AA})}{4\bar{n}_B(f + \Delta)} \Sigma_{aA,aB} \leq n_{aA} \leq \frac{f(n_{AB} + 2n_{AA})}{\bar{n}_A(D - 2\Delta)} \Sigma_{aA,aB}. \tag{III.212}$$

Proof.

1. We start with the upper bound. First observe that it holds at $t = T_1$. By Lemma III.1 it is enough to show that if n_{aA} would reach the upper bound it would decrease faster than the bound. Using Proposition III.4 and that $\eta < c$ a majorising process on aA is given by

$$b_{aA} \leq \tfrac{f}{2\Sigma_5} \Sigma_{aA,aB}(n_{AB} + 2n_{AA} + \Theta(\delta)), \tag{III.213}$$

$$d_{aA} \geq n_{aA} \left(D + c\bar{n}_B - \tfrac{\Delta}{\bar{n}_B} n_{AA} - \eta n_{BB} - \Theta(\Delta^2 n_{AA}) \right) \geq n_{aA}(D - 2\Delta), \tag{III.214}$$

$$\dot{n}_{aA} \leq \frac{f(2n_{AA} + n_{AB} + \Theta(\delta))}{2\Sigma_5} \Sigma_{aA,aB} - n_{aA}(D - 2\Delta). \tag{III.215}$$

We calculate the slope of the majorising process at the upper bound:

$$\dot{n}_{aA} \leq f(2n_{AA} + n_{AB}) \Sigma_{aA,aB} \left(\tfrac{1}{2\Sigma_5} - \tfrac{1}{\bar{n}_A} + \Theta(\delta) \right) \leq -\tfrac{f}{2\bar{n}_A}(2n_{AA} + n_{AB} + \Theta(\delta)) \Sigma_{aA,aB}. \tag{III.216}$$

We have to show that at the upper bound,

$$\dot{n}_{aA} \leq \frac{f(\dot{n}_{AB} + 2\dot{n}_{AA})}{\bar{n}_A(D - 2\Delta)} \Sigma_{aA,aB} + \frac{f(n_{AB} + 2n_{AA})}{\bar{n}_A(D - 2\Delta)} \dot{\Sigma}_{aA,aB}. \tag{III.217}$$

To do this we calculate minorising processes on n_{AB} and n_{AA}:

$$b_{AB} \geq \tfrac{f}{\Sigma_5} n_{AB} \left(\tfrac{1}{2} n_{AB} + n_{AA} + n_{BB} \right) + \tfrac{2f}{\Sigma_5} n_{AA}(n_{BB} - \Theta(\delta^2)), \tag{III.218}$$

$$d_{AB} \leq n_{AB} f, \tag{III.219}$$

$$\dot{n}_{AB} \geq -\tfrac{f}{2\Sigma_5} n_{AB}^2 + \tfrac{2f}{\Sigma_5} n_{AA}(n_{BB} - \Theta(\delta^2)), \tag{III.220}$$

$$b_{AA} \geq \tfrac{f}{\Sigma_5} n_{AA} \left(n_{AB} + n_{AA} - \Theta(\delta^2) \right) + \tfrac{f}{4\Sigma_5} n_{AB}^2, \tag{III.221}$$

$$d_{AA} \leq n_{AA}(f + \Delta + \Theta(\delta^2)), \tag{III.222}$$

$$\dot{n}_{AA} \geq -n_{AA} \left(\tfrac{f}{\Sigma_5} n_{BB} + \Delta + \Theta(\delta^2) \right) + \tfrac{f}{4\Sigma_5} n_{AB}^2. \tag{III.223}$$

Hence we get that

$$\dot{n}_{AB} + 2\dot{n}_{AA} \geq n_{AA}\left(\tfrac{2f}{\bar{\Sigma}_5} n_{BB} - \tfrac{2f}{\bar{\Sigma}_5} n_{BB} - 2\Delta - \Theta(\delta^2)\right) = -2(\Delta + \Theta(\delta^2))n_{AA}. \quad \text{(III.224)}$$

By Lemma III.3, we know that $\dot{\Sigma}_{aA,aB} \geq -\Delta\Sigma_{aA,aB}$. Thus the right-hand side minus the left-hand side of (III.217) is bounded from below by

$$\begin{aligned} &- \frac{2f(\Delta + \Theta(\delta^2))n_{AA}\Sigma_{aA,aB}}{\bar{n}_A(D - 2\Delta)} - \frac{f\Delta(n_{AB} + 2n_{AA})\Sigma_{aA,aB}}{\bar{n}_A(D - 2\Delta)} + \frac{f(n_{AB} + 2n_{AA} + \Theta(\delta))\Sigma_{aA,aB}}{2\bar{n}_A} \\ &\geq \frac{f n_{AA}\Sigma_{aA,aB}}{\bar{n}_A}\left(1 - \frac{4\Delta}{D - 2\Delta}\right) + \frac{f n_{AB}\Sigma_{aA,aB}}{2\bar{n}_A}\left(1 - \frac{2\Delta}{D - 2\Delta}\right) + \Theta(\delta^2) > 0. \quad \text{(III.225)} \end{aligned}$$

This finishes the proof of the upper bound.

2. For the lower bound we proceed similarly (using Lemma III.1). This time we show that if n_{aA} would reach the lower bound it would start to increase faster than the bound. Using Proposition III.4 a minorising process on n_{aA} is given by

$$b_{aA} \geq \tfrac{f}{2\bar{\Sigma}_5}\Sigma_{aA,aB}(2n_{AA} + n_{AB} - \Theta(\delta)), \quad \text{(III.226)}$$

$$d_{aA} \leq n_{aA}(f + \Delta + \Theta(\delta^2)), \quad \text{(III.227)}$$

$$\dot{n}_{aA} \geq \frac{f(2n_{AA} + n_{AB} - \Theta(\delta))}{2\bar{n}_B}\Sigma_{aA,aB} - n_{aA}(f + \Delta). \quad \text{(III.228)}$$

We calculate the slope of the minorising process at the lower bound:

$$\begin{aligned} \dot{n}_{aA} &\geq \frac{f(2n_{AA} + n_{AB} - \Theta(\delta))}{2\bar{n}_B}\Sigma_{aA,aB} - \frac{f(2n_{AA} + n_{AB})}{4\bar{n}_B}\Sigma_{aA,aB} \\ &= \frac{f(2n_{AA} + n_{AB} - \Theta(\delta))}{4\bar{n}_B}\Sigma_{aA,aB} > 0. \quad \text{(III.229)} \end{aligned}$$

Thus the minorising process on n_{aA} would increase when the aA population would reach the lower bound. To ensure this lower bound we have to show

$$\dot{n}_{aA} \geq \frac{f(\dot{n}_{AB} + 2\dot{n}_{AA})}{4\bar{n}_B(f + \Delta)}\Sigma_{aA,aB} + \frac{f(n_{AB} + 2n_{AA})}{4\bar{n}_B(f + \Delta)}\dot{\Sigma}_{aA,aB}. \quad \text{(III.230)}$$

For this we consider a majorising process on $\Sigma_{aA,aB}$ given by:

$$\dot{\Sigma}_{aA,aB} \leq \tfrac{\Delta}{\bar{n}_B} n_{AA}\Sigma_{aA,aB} - n_{aA}(\Delta - \eta n_{BB}) + \Theta(\Delta^2 n_{AA}). \quad \text{(III.231)}$$

Using that $\eta < c$, the slope of this process if n_{aA} reaches the lower bound is estimated by

$$\begin{aligned} \dot{\Sigma}_{aA,aB} &\leq \tfrac{\Delta}{\bar{n}_B} n_{AA}\Sigma_{aA,aB} - \frac{f(2n_{AA} + n_{AB})}{4\bar{n}_B(f + \Delta)}(\Delta - \eta n_{BB})\Sigma_{aA,aB} + \Theta(\Delta^2 n_{AA}) \\ &\leq \frac{f(2n_{AA} + n_{AB})}{4\bar{n}_B}\frac{f - D}{f + \Delta}\Sigma_{aA,aB} + \tfrac{\Delta}{\bar{n}_B} n_{AA}\Sigma_{aA,aB} + \Theta(\Delta^2 n_{AA}). \quad \text{(III.232)} \end{aligned}$$

Moreover we need majorising processes on AA and AB:

$$b_{AB} \leq \tfrac{f}{\Sigma_5} n_{AB} \left(\tfrac{1}{2} n_{AB} + n_{AA} + n_{BB}\right) + \tfrac{2f}{\Sigma_5} n_{AA} n_{BB} + \Theta(\delta), \tag{III.233}$$

$$d_{AB} \geq n_{AB} \left(f - \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA}\right), \tag{III.234}$$

$$\dot{n}_{AB} \leq -\tfrac{f}{2\Sigma_5} n_{AB}^2 + \tfrac{2f}{\Sigma_5} n_{AA} n_{BB} + \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA} n_{AB} + \Theta(\delta), \tag{III.235}$$

$$b_{AA} \leq \tfrac{f}{\Sigma_5} n_{AA} \left(n_{AB} + n_{AA}\right) + \tfrac{f}{4\Sigma_5} n_{AB}^2 + \Theta(\delta), \tag{III.236}$$

$$d_{AA} \geq n_{AA} \left(f + \Delta - \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA}\right), \tag{III.237}$$

$$\dot{n}_{AA} \geq -n_{AA} \left(\tfrac{f}{\Sigma_5} n_{BB} + \Delta - \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA}\right) + \tfrac{f}{4\Sigma_5} n_{AB}^2 + \Theta(\delta). \tag{III.238}$$

Thus we have

$$\dot{n}_{AB} + 2\dot{n}_{AA} \leq -\Delta n_{AA} \left(2 - \tfrac{2n_{AA}+n_{AB}}{\bar{n}_B}\right) + \Theta\left(\Delta^2 n_{AA}\right) < \Theta\left(\Delta^2 n_{AA}\right). \tag{III.239}$$

It is enough to show that

$$\dot{n}_{aA} \geq \frac{f(2n_{AA} + n_{AB})}{4\bar{n}_B(f+\Delta)} \dot{\Sigma}_{aA,aB} + \Theta\left(\Delta^2 n_{AA}\right) \Sigma_{aA,aB}, \tag{III.240}$$

using that $\eta < c$ we have

$$\begin{aligned}
&\frac{f(2n_{AA} + n_{AB} - \Theta(\delta))}{4\bar{n}_B} \Sigma_{aA,aB} - \frac{f^2(n_{AB} + 2n_{AA})^2}{16\bar{n}_B^2(f+\Delta)} \frac{f-D}{f+\Delta} \Sigma_{aA,aB} \\
&\quad - \frac{f(n_{AB} + 2n_{AA})}{4\bar{n}_B(f+\Delta)} \frac{\Delta}{\bar{n}_B} n_{AA} \Sigma_{aA,aB} - \Theta\left(\Delta^2 n_{AA}\right) \Sigma_{aA,aB} \\
&\geq \frac{f(2n_{AA} + n_{AB} - \Theta(\delta))}{4\bar{n}_B} \Sigma_{aA,aB} - \frac{f(2n_{AA} + n_{AB})}{8\bar{n}_B} \Sigma_{aA,aB} \left(1 + \frac{2\Delta(1+\Delta) n_{AA}}{\bar{n}_B(f+\Delta)}\right) \\
&\quad - \Theta\left(\Delta^2 n_{AA}\right) \Sigma_{aA,aB} \\
&> 0.
\end{aligned} \tag{III.241}$$

This concludes the proof. □

Proposition III.8. *For all $t \in [T_1, T_2]$ the process $\Sigma_{aA,aB}$ is bounded by*

1. $\dot{\Sigma}_{aA,aB} \leq n_{aA} \left(\eta n_{BB} - \Delta \frac{n_{AB} + \Theta(\Delta n_{AA})}{n_{AB} + 2n_{AA}}\right)$.
2. $\dot{\Sigma}_{aA,aB} \geq n_{aA}(\eta n_{BB} - \Delta - \Theta(\delta))$.

Proof.

1. We construct a majorising process on $\Sigma_{aA,aB}$ and use Proposition III.4 and Lemma III.4:

$$\begin{aligned}
b_{\Sigma_{aA,aB}} &\leq n_{aA} \frac{f\left(\frac{1}{2} n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}\right)}{\Sigma_5} + n_{aB} \frac{f\left(n_{AA} + \frac{1}{2} n_{aB} + n_{AB} + n_{BB}\right)}{\Sigma_5} + \Theta\left(\Sigma_{aA,aB}^2\right) \\
&\leq f \Sigma_{aA,aB} + \Theta\left(\Sigma_{aA,aB}^2\right),
\end{aligned} \tag{III.242}$$

$$\begin{aligned}
d_{\Sigma_{aA,aB}} &\geq \Sigma_{aA,aB}(D - \Delta + c\Sigma_5) + \Delta n_{aA} - \eta n_{aA} n_{BB} \\
&\geq \Sigma_{aA,aB}(f - \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA}) + \Delta n_{aA} - \eta n_{aA} n_{BB},
\end{aligned} \tag{III.243}$$

$$\dot{\Sigma}_{aA,aB} \leq \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA} n_{aB} - n_{aA}(\Delta - \tfrac{\Delta(1+\Delta)}{\bar{n}_B} n_{AA} - \eta n_{BB}) + \Theta\left(\Sigma_{aA,aB}^2\right). \tag{III.244}$$

To bound n_{aB} we use Lemma III.5:

$$\begin{aligned}\dot{\Sigma}_{aA,aB} &\leq n_{aA}\left(\frac{\Delta(1+\Delta)n_{AA}}{n_{AB}+2n_{AA}}\frac{n_{AB}+2n_{BB}+\frac{2\Delta}{c}}{\bar{n}_B}-\Delta+\frac{\Delta(1+\Delta)}{\bar{n}_B}n_{AA}+\eta n_{BB}+\Theta(\delta)\right)\\ &\leq n_{aA}\left(\eta n_{BB}+\frac{\Delta(n_{AA}(n_{AB}+2n_{BB})+n_{AA}(n_{AB}+2n_{AA})-\bar{n}_B(n_{AB}+2n_{AA}))+\Theta\left(\Delta^2 n_{AA}\right)}{\bar{n}_B(n_{AB}+2n_{AA})}\right)\\ &\leq n_{aA}\left(\eta n_{BB}-\Delta\frac{n_{AB}+\Theta(\Delta n_{AA})}{n_{AB}+2n_{AA}}\right). \qquad \text{(III.245)}\end{aligned}$$

2. This time we construct a minorising process on $\Sigma_{aA,aB}$ by using Proposition III.4 and Lemma III.4:

$$\begin{aligned}b_{\Sigma_{aA,aB}} &\geq f\frac{n_{aA}\left(\frac{1}{2}n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}\right)+n_{aB}\left(n_{AA}+\frac{1}{2}n_{aB}+n_{AB}+n_{BB}\right)}{n_{aA}+n_{AA}+n_{aB}+n_{AB}+n_{BB}}-\Theta\left(\delta^2\right)\\ &\geq f\Sigma_{aA,aB}-\Theta\left(\delta^2\right), \qquad \text{(III.246)}\\ d_{\Sigma_{aA,aB}} &\leq \Sigma_{aA,aB}(D-\Delta+c\Sigma_5)-\eta n_{aA}n_{BB}+\left(\Delta+\Theta\left(\delta^2\right)\right)n_{aA}\\ &\leq f\Sigma_{aA,aB}-n_{aA}\left(\eta n_{BB}-\Delta-\Theta\left(\delta^2\right)\right), \qquad \text{(III.247)}\\ \dot{\Sigma}_{aA,aB} &\geq n_{aA}(\eta n_{BB}-\Delta-\Theta(\delta)). \qquad \text{(III.248)}\end{aligned}$$

□

From this Proposition we can deduce

Corollary III.1. *There exists a $t^* \in [T_1, T_2]$, such that for all $t \in [t^*, T_2]$ and $\eta > \frac{4\Delta}{\bar{n}_B} =: \eta^\star$, it holds*

$$\dot{\Sigma}_{aA,aB}(t) > 0. \qquad \text{(III.249)}$$

Proof. A fine calculation will show that the competition $c-\eta$ felt by an aA individual from a BB individual allow the sum $\Sigma_{aA,aB}$ to grow when η is large enough, whereas it decreases when $\eta = 0$. Note that we consider here the sum $\Sigma_{aA,aB}$ because the influence of η cannot be seen in the rates of the aB population alone. Heuristically, the growth of the aB population happens due to the indirect influence (source of a allele) of the less decaying aA population. We prove that the minorising process on $\Sigma_{aA,aB}$ estimated in the Proposition III.8 starts to increase:

$$\dot{\Sigma}_{aA,aB} \geq n_{aA}(\eta n_{BB}-\Delta-\Theta(\delta)). \qquad \text{(III.250)}$$

As soon as $n_{BB} > \Delta/\eta$, the sum-process $\Sigma_{aA,aB}$ starts to increase. From Lemma III.6 and Proposition III.6 we know that, for $t \geq T_=$, we have $n_{BB} \geq \frac{\bar{n}_B}{4} - \Theta(\Delta)$. Hence, if we choose $\eta > \frac{4\Delta}{\bar{n}_B}$ the sum-process $\Sigma_{aA,aB}$ increases. □

Now we calculate the time $T^{aA=\delta AA} \wedge T^{aB=\delta AB}$ and we will see that $T^{aA=\delta AA} \wedge T^{aB=\delta AB} = T^{aA=\delta AA}$.

Theorem III.2. *The time $T_2 = \Theta(\varepsilon^{-1/(1+\eta\bar{n}_B-\Delta)})$.*

Proof. From Proposition III.8 (2) we have a lower bound on $\dot{\Sigma}_{aA,aB}$, and with Lemma III.8 (2) we can bound this further from below by:

$$\begin{aligned}\dot{\Sigma}_{aA,aB} &\geq n_{aA}(\eta n_{BB} - \Delta - \Theta(\delta)) \\ &\geq (\eta n_{BB}(t) - \Delta - \Theta(\delta))\frac{f(n_{AB}(t) + 2n_{AA}(t))}{4\bar{n}_B(f+\Delta)}\Sigma_{aA,aB}(t) \\ &\geq \frac{\Theta(\eta\bar{n}_B/4 - \Delta)}{\Theta(1) + \Theta(1)t}\Sigma_{aA,aB}(t), \end{aligned} \tag{III.251}$$

where the last estimation on n_{BB} and on n_{AB} comes from Proposition III.6 and from [84] since we know from there that the time until $n_{AB} = \Theta(\sqrt{n_{AA}})$, starts to decrease like $1/t$ is of order $\Theta(1)$. As $\Sigma_{aA,aB}(T_1) = \Theta(\varepsilon)$, the solution of the ODE that gives a lower bound is:

$$\Sigma_{aA,aB}(t) \geq \Theta(\varepsilon)(\Theta(1) + \Theta(1)t)^{\Theta(\eta\bar{n}_B/4 - \Delta)}. \tag{III.252}$$

By using Proposition III.8 (1), we get the same kind of solution as an upper bound on $\Sigma_{aA,aB}$ (note on the last step we can upper bound n_{BB} by $\bar{n}_B$):

$$\Sigma_{aA,aB}(t) \leq \Theta(\varepsilon)(\Theta(1) + \Theta(1)t)^{\Theta(\eta\bar{n}_B - \Delta)}. \tag{III.253}$$

Using (III.252) and Lemma III.8 we get a minorising process on aA:

$$n_{aA}(t) = \Theta(n_{AB}\Sigma_{aA,aB}) \geq \Theta(\varepsilon)(\Theta(1) + \Theta(1)t)^{\Theta(\eta\bar{n}_B/4 - \Delta)}/(\Theta(1) + \Theta(1)t). \tag{III.254}$$

The corresponding majorising process has an $\bar{n}_B$ instead of $\bar{n}_B/4$. By solving $n_{aA} = \delta n_{AA} = \Theta(n_{AB}^2)$ we get the order of magnitude of $T_{aa=\delta AA}$:

$$\Theta\left(\varepsilon^{-1/(1+\eta\bar{n}_B - \Delta)}\right) \leq T_{aa=\delta AA} \leq \Theta\left(\varepsilon^{-1/(1+\eta\bar{n}_B/4 - \Delta)}\right). \tag{III.255}$$

Note that $1 + \eta\bar{n}_B - \Delta > 0$ for Δ small enough, and thus $T_{aa=\delta AA}$ diverges with ε and the order calculations above are justified.

It is left to ensure that aB does not exceed δn_{AB} in this time. It follows from Lemma III.8 that during the time interval $[T_1, T_2]$, we have $\Sigma_{aA,aB} = \Theta(n_{aB})$. Thus, solving $n_{aB} = \delta n_{AB}$ amounts to solving $\Theta(\Sigma_{aA,aB}) = \Theta(1)/(\Theta(1) + \Theta(1)t)$ which gives the very same order of magnitude as for $T_{aA=\delta AA}$. Thus the two times are of the same order.

Note that for $\eta = 0$, $\Sigma_{aA,aB}(T_2) = \Theta\left(\varepsilon^{1+\Delta/(1-\Delta)}\right)$. This proves point 1 of Theorem III.1.

□

Proposition III.9. $T_2 = T^{aA=\delta AA}$.

Proof. This follows from Theorem III.2 and Lemma III.4. □

Proposition III.10. *At time $t = T_2$ and if f is taken sufficiently large (Assumption C2), n_{aa} starts to grow out of itself: there exists some positive constant $c_{T_2} > 0$ such that*

$$\dot{n}_{aa} \geq c_{T_2} \cdot n_{aa}. \tag{III.256}$$

Proof. We have $n_{AA}(T_2) = \Theta\left(\varepsilon^{2/(1+\eta\bar{n}_B-\Delta)}\right)$. Thus, at the end of the second phase,

$$b_{aa} \geq f n_{aa} \frac{\frac{1}{2}\delta n_{AA}}{n_{AA}(1+\Theta(\delta))} = \frac{\delta f n_{aa}}{2(1+\Theta(\delta))}, \tag{III.257}$$

$$d_{aa} \leq n_{aa}(D+\Delta+n_{AA}(1+\Theta(\delta))) = n_{aa}(D+\Delta+\Theta\left(\varepsilon^{2/(1+\eta\bar{n}_B-\Delta)}\right), \tag{III.258}$$

$$\dot{n}_{aa} \geq n_{aa}\left(\frac{\delta f}{2} - D - \Delta - \Theta\left(\varepsilon^{2/(1+\eta\bar{n}_B-\Delta)}\right)\right), \tag{III.259}$$

the right-hand side is positive for f large enough. □

4.4. Phase 3: Exponential growth of aa until co-equilibrium with BB

Since aa is growing now also out of itself it will influence the sum-process $\Sigma_5 = n_{aA} + n_{AA} + n_{aB} + n_{AB} + n_{BB}$ and we need new lower bounds on Σ_5 in the following steps, the proof of this works similar to the one of Proposition III.4 by taking into account all contributing populations. Let us compute the ODE to which Σ_5 is the solution:

Proposition III.11. *The sum-process Σ_5 is the solution to*

$$\begin{aligned}\dot{\Sigma}_5 =& \Sigma_5\left(f - D - \Delta - c\Sigma_5\right) - \Delta\left(n_{aA}+n_{AA}\right) - c n_{aa}\left(n_{aA}+n_{AA}\right) + 2\eta n_{aA} n_{BB} \\ &+ \frac{f}{\Sigma_3} n_{aa}\left(\frac{1}{2}n_{aA}+n_{AA}\right) - \frac{f}{4\Sigma_5} n_{aB}(n_{aA}+n_{aB}) - \frac{f}{4\Sigma_6} n_{aA}\left(2n_{aa}+n_{aA}+n_{aB}\right).\end{aligned} \tag{III.260}$$

Proof. We calculate the birth- and the death-rate of Σ_5 under consideration of the aa

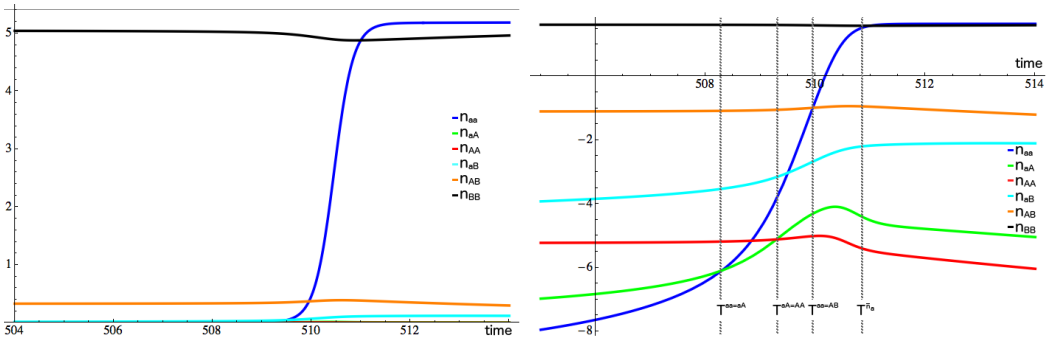

Figure III.6.: Zoom-in when aa recovers, general qualitative behaviour of $\{n_i(t), i \in \mathcal{G}\}$ (lhs) and log-plot (rhs).

population:

$$\begin{aligned} b_{\Sigma_5} =& \tfrac{f}{\Sigma_3} n_{aa}\left(\tfrac{1}{2}n_{aA}+n_{AA}\right)+\tfrac{f}{\Sigma_5}\left(\left(n_{aB}+n_{AB}+n_{BB}\right)\Sigma_5-\tfrac{1}{4}n_{aB}\left(n_{aA}+n_{aB}\right)\right) \\ &+\frac{f}{\Sigma_6}\left(\left(n_{aA}+n_{AA}\right)\Sigma_6-n_{aA}\left(\tfrac{1}{2}n_{aa}+\tfrac{1}{4}n_{aA}+\tfrac{1}{4}n_{aB}\right)\right) \\ =& f\Sigma_5+\tfrac{f}{\Sigma_3}n_{aa}\left(\tfrac{1}{2}n_{aA}+n_{AA}\right)-\tfrac{f}{4\Sigma_5}n_{aB}\left(n_{aA}+n_{aB}\right)-\tfrac{f}{4\Sigma_6}n_{aA}\left(2n_{aa}+n_{aA}+n_{aB}\right), \end{aligned} \tag{III.261}$$

$$d_{\Sigma_5} =\Sigma_5\left(D-\Delta+c\Sigma_5\right)+\left(cn_{aa}+\Delta\right)\left(n_{aA}+n_{AA}\right)-2\eta n_{aA}n_{BB}, \tag{III.262}$$

$$\dot{\Sigma}_5 =\Sigma_5\left(f-D-\Delta-c\Sigma_5\right)-\left(cn_{aa}+\Delta\right)\left(n_{aA}+n_{AA}\right)+2\eta n_{aA}n_{BB} \tag{III.263}$$

$$+\tfrac{f}{\Sigma_3}n_{aa}\left(\tfrac{1}{2}n_{aA}+n_{AA}\right)-\tfrac{f}{4\Sigma_5}n_{aB}(n_{aA}+n_{aB})-\tfrac{f}{4\Sigma_6}n_{aA}\left(2n_{aa}+n_{aA}+n_{aB}\right). \tag{III.264}$$

□

We introduce some notation for the order of magnitude of $n_{AA}(T_2)$. We write $n_{AA}(T_2)=\Theta(\varepsilon^\gamma)$ with

$$\gamma := 2/(1+\eta\bar{n}_B-\Delta). \tag{III.265}$$

Let

$$T_3 := T^{aa}_{\bar{n}_a-\varepsilon^{\gamma/2}} = \inf\left\{t>T_2 : n_{aa}(t)=\bar{n}_a-\varepsilon^{\gamma/2}\right\}. \tag{III.266}$$

We have to ensure that the aa population grows until a neighbourhood of its equilibrium. At the end of the second Phase it holds true that $n_{aA}, n_{AA}, n_{aB} \le n_{AB} < \Delta$. We start to bound Σ_5:

Lemma III.9. *The sum-process Σ_5 is bounded from above and below by*

$$\bar{n}_B-\tfrac{2(f+\Delta)}{c\bar{n}_B}n_{AB} \le \Sigma_5 \le \bar{n}_B+\tfrac{3(f+\Delta)}{c\bar{n}_B}n_{AB}. \tag{III.267}$$

Proof. We use Lemma III.1 and construct minorising and majorising processes on Σ_5 and n_{AB}. First observe that the bounds are satisfy at $t=T_2$ by Proposition III.4.

$$f\Sigma_5-fn_{AB} \le b_{\Sigma_5} \le f\Sigma_5+2fn_{AB}, \tag{III.268}$$

$$\Sigma_5(D-\Delta+c\Sigma_5)+2(f+\Delta)n_{AB} \ge d_{\Sigma_5} \ge \Sigma_5(D-\Delta+c\Sigma_5)-(f+\Delta)n_{AB}, \tag{III.269}$$

$$\Sigma_5(f-D+\Delta-c\Sigma_5)-(2f+\Delta)n_{AB} \le \dot{\Sigma}_5 \le \Sigma_5(f-D+\Delta-c\Sigma_5)+(3f+\Delta)n_{AB}. \tag{III.270}$$

At the lower bound we get that the sum-process would increase:

$$\dot{\Sigma}_5 \ge \Delta n_{AB}-\Theta\left(n_{AB}^2\right) > 0, \tag{III.271}$$

and at the upper bound it would decrease:

$$\dot{\Sigma}_5 \le -2\Delta n_{AB}+\Theta\left(n_{AB}^2\right) < 0. \tag{III.272}$$

It is left to show that

$$-\tfrac{2(f+\Delta)}{c\bar{n}_B}\dot{n}_{AB} \le \dot{\Sigma}_5 \le \tfrac{3(f+\Delta)}{c\bar{n}_B}\dot{n}_{AB}. \tag{III.273}$$

For this we construct a minorising process on AB:

$$b_{AB} \geq f n_{AB} - \tfrac{9f}{4\Sigma_5} n_{AB}^2, \tag{III.274}$$

$$d_{AB} \leq n_{AB}(D - \Delta + c\Sigma_5), \tag{III.275}$$

$$\dot{n}_{AB} \geq n_{AB}\left(f - D + \Delta - c\Sigma_5 - \tfrac{9f}{4\Sigma_5} n_{AB}^2\right). \tag{III.276}$$

At the lower bound we have $\dot{n}_{AB} \geq -\frac{f+8\Delta}{4\bar{n}_B} n_{AB}^2 + \Theta\left(n_{AB}^3\right)$ and the lhs of (III.273) $\leq \frac{(f+8\Delta)(f+\Delta)}{2c\bar{n}_B^2} n_{AB}^2 + \Theta\left(n_{AB}^3\right) < \Delta n_{AB} - \Theta\left(n_{AB}^2\right)$. At the upper bound we get $\dot{n}_{AB} \geq -\frac{21f+4\Delta}{4\bar{n}_B} n_{AB}^2 - \Theta\left(n_{AB}^3\right)$ and that the rhs of (III.273) is larger or equal to $-\frac{3(f+\Delta)(21f+4\Delta)}{4c\bar{n}_B^2} n_{AB}^2 - \Theta\left(n_{AB}^3\right) > -2\Delta n_{AB} - \Theta\left(n_{AB}^2\right)$. □

The following lemma ensure that the aA- and the AA populations stay smaller than $\Theta\left(n_{AB}^2\right)$:

Lemma III.10. *1. For all $t \in [T_2, T_3]$, the AA population is bounded from above by*

$$n_{AA} \leq \tfrac{2}{\bar{n}_B} n_{AB}^2, \tag{III.277}$$

2. For all $t \in [T_2, T_3]$ and $\eta < \frac{c}{2}\left(1 - \frac{\Theta(\Delta)}{f-D+\Delta}\right)$, the aA population is bounded from above by

$$n_{aA} \leq \tfrac{10f}{\bar{n}_B(D-\Delta)} n_{AB}^2. \tag{III.278}$$

Proof. The proof uses again Lemma III.1. First observe that by Lemma III.7 (1) and since $n_{aA}(T_2) = \delta n_{AA}(T_2)$ the bounds are satisfy at $t = T_2$.

1. We construct a majorising process on AA:

$$b_{AA} \leq \tfrac{f}{\Sigma_5} n_{AA}(n_{aA} + n_{AA} + n_{AB}) + \tfrac{f}{4\Sigma_5}(n_{aA} + n_{AB})^2, \tag{III.279}$$

$$d_{AA} \geq n_{AA}(f - \Theta(\Delta)), \tag{III.280}$$

$$\dot{n}_{AA} \leq -n_{AA}\left(\tfrac{f}{\Sigma_5}(n_{aB} + n_{AB}) - \Theta(\Delta)\right) + \tfrac{f}{\Sigma_5} n_{AB}^2. \tag{III.281}$$

This process decreases at the upper bound: $\dot{n}_{AA} \leq -\frac{f}{\bar{n}_B} n_{AB}^2 + \Theta\left(\Delta n_{AB}^2\right) < 0$. It is left to show that

$$\dot{n}_{AA} \leq \tfrac{4}{\bar{n}_B} n_{AB} \dot{n}_{AB}. \tag{III.282}$$

For this we construct a minorising process on AB, given by $\dot{n}_{AB} \geq -\frac{9f+\Theta(\Delta)}{2\bar{n}_B} n_{AB}^2$. Thus the rhs of (III.282) is larger or equal to $-\frac{18f+\Theta(\Delta)}{\bar{n}_B^2} n_{AB}^3 \geq -\frac{f}{\bar{n}_B} n_{AB}^2 + \Theta\left(\Delta n_{AB}^2\right)$.

2. Similarly to (1) we construct a majorising process on aA:

$$b_{aA} \leq \tfrac{f}{2\Sigma_3} n_{aa} n_{aA} + \tfrac{9f}{2\bar{n}_B} n_{AB}^2 + \tfrac{f}{2\Sigma_6} n_{aa} n_{aA} + \Theta(\Delta n_{AB}^2), \tag{III.283}$$

$$d_{aA} \geq n_{aA}\left(f + \Delta - \eta\bar{n}_B + c n_{aa} - \Theta(n_{AB})\right), \tag{III.284}$$

$$\dot{n}_{aA} \leq -n_{aA}\left(\tfrac{f}{2} - \eta\bar{n}_B + \tfrac{f-2(D-\Delta)}{2\Sigma_5} n_{aa}\right) + \tfrac{9f}{2\bar{n}_B} n_{AB}^2 + \Theta(\Delta n_{AB}^2), \tag{III.285}$$

$$\dot{n}_{aA} \leq -n_{aA} \tfrac{f-2\eta\bar{n}_B}{2} + \tfrac{9f+\Theta(\Delta)}{2\bar{n}_B} n_{AB}^2. \tag{III.286}$$

At the upper bound we get: $\dot{n}_{aA} \leq n_{AB}^2 \left(\frac{9f+\Theta(\Delta)}{2\bar{n}_B} - \frac{10f(f-2\eta\bar{n}_B)}{2\bar{n}_B(D-\Delta)}\right)$. It is to show that

$$\dot{n}_{aA} \leq \frac{20f}{\bar{n}_B(D-\Delta)} n_{AB}\dot{n}_{AB}. \tag{III.287}$$

For this we construct a minorising process on AB, given by $\dot{n}_{AB} \geq -\frac{9f+\Theta(\Delta)}{2\bar{n}_B} n_{AB}^2$. Thus the rhs of (III.287) is larger or equal to $-\frac{90f+\Theta(\Delta)}{\bar{n}_B(D-\Delta)} n_{AB}^3$. Thus we have to ensure that $\left(\frac{9f+\Theta(\Delta)}{2\bar{n}_B} - \frac{10f(f-2\eta\bar{n}_B)}{2\bar{n}_B(D-\Delta)}\right) < -\frac{90f+\Theta(\Delta)}{\bar{n}_B(D-\Delta)}\Delta$. This yields the condition on η:

$$\eta < \frac{f-D-\Theta(\Delta)}{2\bar{n}_B} = \frac{c}{2}\left(1 - \frac{\Theta(\Delta)}{f-D+\Delta}\right). \tag{III.288}$$

□

Remark III.5. *Observe that the condition $\eta < \frac{c}{2}\left(1 - \frac{\Theta(\Delta)}{f-D+\Delta}\right)$ in Lemma III.10 prevent aA to grow exponentially fast. Inequality* (III.285) *shows that for larger values for η, n_{aA} would start to grow out of itself and thus the system would converge towards the 6-point equilibrium as we checked numerically. Hence, the assumption $\eta < \frac{c}{2}\left(1 - \frac{\Theta(\Delta)}{f-D+\Delta}\right)$ is essential in this phase and propagates to the following lemmata since we need therein the n_{AB}^2-dependent bound on aA.*

Using Lemma III.10 we can also compute a lower bound for AA:

Lemma III.11. *For $\eta \leq \frac{c}{2} - \Theta(\Delta)$, the AA population is bounded from below by*

$$\frac{1}{8\bar{n}_B} n_{AB}^2 \leq n_{AA}. \tag{III.289}$$

Proof. By Lemma III.1 we construct a minorising process on AA. The bound is satisfy by Lemma III.7 2. at $t = T_2$.

$$b_{AA} \geq \frac{f}{4\Sigma_5} n_{AB}^2, \tag{III.290}$$

$$d_{AA} \leq n_{AA}(D + c\Sigma_5 + cn_{aa}), \tag{III.291}$$

$$\dot{n}_{AA} \geq -n_{AA}\left(f + \Delta + \frac{3(f+\Delta)}{\bar{n}_B} n_{AB} + cn_{aa}\right) + \frac{f}{4\Sigma_5} n_{AB}^2. \tag{III.292}$$

At the lower bound this process increases $\dot{n}_{AA} \geq \frac{D-\Theta(\Delta^2)}{8\bar{n}_B} n_{AB}^2 > 0$. It is left to show that

$$\dot{n}_{AA} \geq \frac{1}{4\bar{n}_B} n_{AB}\dot{n}_{AB}. \tag{III.293}$$

We construct a majorising process on AB, given by $\dot{n}_{AB} \leq \frac{2(f+\Delta+\Theta(\Delta^2))}{4\bar{n}_B} n_{AB}^2 + f n_{aA}$, using Lemma III.10 we get that the rhs of (III.293) is smaller or equal to $\frac{f}{4\bar{n}_B} n_{AB} n_{aA} + \Theta\left(n_{AB}^3\right) \leq \Theta\left(n_{AB}^3\right)$. □

With all these lemmata we are now able to show that n_{AB} stays of order $\Theta\left(\varepsilon^{\gamma/2}\right)$ until T_3 when n_{aa} reaches the neighbourhood of its equilibrium.

Lemma III.12. *For the time $T_{\varepsilon^{\gamma/2}}^{AB}$, until AB exceed the order $\varepsilon^{\gamma/2}$, it holds*

$$T_{\varepsilon^{\gamma/2}}^{AB} \geq \Theta\left(\varepsilon^{-\gamma/2}\right). \tag{III.294}$$

Proof. We construct a majorising process on AB:

$$b_{AB} \leq f n_{AB} + \Theta\left(n_{AB}^2\right), \tag{III.295}$$
$$d_{AB} \geq n_{AB}\left(f - \Theta\left(n_{AB}^2\right)\right), \tag{III.296}$$
$$\dot{n}_{AB} \leq \Theta\left(n_{AB}^2\right). \tag{III.297}$$

With the initial condition $n_{AB}(T_2) = \Theta\left(\varepsilon^{\gamma/2}\right)$, this ODE has the solution

$$n_{AB}(t) \leq \frac{\Theta(1)}{\Theta(1)t - \Theta\left(\varepsilon^{-\gamma/2}\right)}. \tag{III.298}$$

Thus $T^{AB}_{\varepsilon^{\gamma/2}} \geq \Theta\left(\varepsilon^{-\gamma/2}\right)$. □

Observe, it follows that $T^{AA}_{\varepsilon^{\gamma}} = \Theta(\varepsilon^{-\gamma/2})$.

To ensure the exponential growth of aa we need that the aA population reaches the order $\Theta(\epsilon^{\gamma})$ and stays there for a long enough time.

Lemma III.13. *For $\eta \leq \frac{c}{2} - \Theta(\Delta)$ and for all $t \leq T_2 + \Theta\left(\epsilon^{-\gamma/2}\right)$, if $n_{aa} \leq n_{aA} \leq n_{AA}$ or $n_{aA} \leq n_{aa} \leq n_{AA}$, it holds*

$$n_{aA} \geq \Theta\left(\delta n_{AB}^2\right). \tag{III.299}$$

Proof. We construct a minorising process on aA in a very rough way. The death rate can be estimate by:

$$d_{aA} \leq n_{aA}(f + \Delta + \Theta(n_{AB})). \tag{III.300}$$

At time T_2 we know that $n_{aA} = \delta n_{AA} = \Theta\left(\delta n_{AB}^2\right)$. Thus, using Lemma III.12, at T_2, n_{aA} is given by:

$$n_{aA}(t) \leq \frac{\Theta(1)}{\Theta(1)t - \Theta(\delta)\epsilon^{-\gamma}}. \tag{III.301}$$

This process would need time of order $\Theta\left(\delta\epsilon^{-\gamma-\alpha}\right)$, for $\alpha > 0$, to decrease under order $\Theta\left(\delta\epsilon^{\gamma}\right)$ which is larger than the time n_{AB} needs to leave order $\Theta\left(\epsilon^{\gamma/2}\right)$. This way we can ensure that n_{aA} does not decrease under order $\Theta\left(\delta n_{AB}^2\right)$ in time $\Theta\left(\epsilon^{-\gamma/2}\right)$. □

Now we show that n_{aa} increases to a neighbourhood of its equilibrium before time $T^{AB}_{\varepsilon^{\gamma/2}}$.

Lemma III.14. *For $\eta \leq \frac{c}{2} - \Theta(\Delta)$ and all $t \in [T_2, T_3]$ the aa population increases to a $\varepsilon^{\frac{\gamma}{2}}$-neighbourhood of its equilibrium $\bar{n}_a$ exponentially fast and it holds $T_3 < T^{AB}_{\varepsilon^{\gamma/2}}$.*

Proof. We construct a minorising process on aa and distinguishe some cases. First observe that by Lemma III.10 it holds that $d_{aa} \leq n_{aa}\left(D + \Delta + cn_{aa} + \Theta\left(n_{AB}^2\right)\right)$.

1. If $n_{aa} \leq n_{aA} \leq n_{AA}$ or $n_{aA} \leq n_{aa} \leq n_{AA}$
   In that case the birth-rate is given by $b_{aa} \geq f n_{aa} \frac{n_{aA}}{6n_{AA}}$. With Lemma III.10 and III.13 we get $b_{aa} \geq \Theta(\delta) f n_{aa}$ and $\dot{n}_{aa} \geq n_{aa}\left(\Theta(\delta)f - D - \Delta - \Theta\left(n_{AB}^2\right)\right)$. Hence the time $T^{aa}_{\varepsilon^{\gamma}} = \Theta\left(\ln(\varepsilon^{-\gamma})\right)$ until aa reaches $\Theta(\varepsilon^{\gamma})$.

2. If $n_{aa}, n_{AA} \leq n_{aA}$
   In that case, by Lemma III.10 we get $b_{aa} \geq \frac{f}{6} n_{aa}$ and $\dot{n}_{aa} \geq n_{aa}\left(\frac{f}{6} - D - \Delta - \Theta\left(n_{AB}^2\right)\right)$. Hence the time $T_{\varepsilon^\gamma}^{aa} = \Theta\left(\ln\left(\varepsilon^{-\gamma}\right)\right)$ until aa reaches $\Theta\left(\varepsilon^\gamma\right)$.

3. If $n_{aA} \leq n_{AA} \leq n_{aa} \leq n_{AB}$
   In that case, by Lemma III.10, $b_{aa} \geq \frac{f}{3} n_{aa}$ and $\dot{n}_{aa} \geq n_{aa}\left(\frac{f}{3} - D - \Delta - cn_{aa} - \Theta\left(n_{AB}^2\right)\right)$. Hence the time $T_{\varepsilon^{\gamma/2}}^{aa} = \Theta\left(\ln\left(\varepsilon^{-\gamma/2}\right)\right)$ until aa reaches $\Theta\left(\varepsilon^{\gamma/2}\right)$.

4. If $n_{AA} \leq n_{aA} \leq n_{aa} \leq n_{AB}$
   In that case, by Lemma III.10, $b_{aa} \geq \frac{f}{2} n_{aa}$ and $\dot{n}_{aa} \geq n_{aa}\left(\frac{f}{2} - D - \Delta - cn_{aa} - \Theta\left(n_{AB}^2\right)\right)$. Hence the time $T_{\varepsilon^{\gamma/2}}^{aa} = \Theta\left(\ln\left(\varepsilon^{-\gamma/2}\right)\right)$ until aa reaches $\Theta\left(\varepsilon^{\gamma/2}\right)$.

5. If $n_{aa} > n_{AB}$
   In that case, by Lemma III.10, $b_{aa} \geq n_{aa}\left(f - \Theta\left(\frac{n_{AB}^2}{n_{aa}}\right)\right)$ and $\dot{n}_{aa} \geq n_{aa}\Big(f - D - \Delta - cn_{aa} - \Theta\left(\frac{n_{AB}^2}{n_{aa}}\right)\Big)$. Hence the time $T_3 = \Theta(\varepsilon^\gamma \ln(\varepsilon^{-\gamma})) \leq T_{\varepsilon^{\gamma/2}}^{AB}$.

□

4.5. Phase 4: Convergence to $p_{aB} = (\bar{n}_a, 0, 0, 0, 0, \bar{n}_B)$

The Jacobian matrix of the field (III.34) at the fixed point p_{aB} has the 6 eigenvalues: 0 (double), and $-(2f-D), -(f-D+\Delta), -(f-D-\Delta), -((f-D)(5f-4D)+f\Delta)/(4(f-D)+\eta\bar{n}_B)$ which are strictly negative under Assumptions (C). Because of the zero eigenvalues, p_{aB} is a non-hyperbolic equilibrium point of the system and linearisation fails to determine its stability properties. Instead, we use the result of the center manifold theory (51, 88) that asserts that the qualitative behaviour of the dynamical system in a neighbourhood of the non-hyperbolic critical point p_{aB} is determined by its behaviour on the center manifold near p_{aB}.

Theorem III.3 (The Local Center Manifold Theorem 2.12.1 in 88)**.** *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*

$$\dot{x} = Cx + F(x, y), \tag{III.302}$$

$$\dot{y} = Py + G(x, y), \tag{III.303}$$

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 δ-neighbourhood of 0, that defines the local center manifold and satisfies:

$$Dh(x)[Cx + F(x, h(x))] - Ph(x) - G(x, h(x)) = 0, \tag{III.304}$$

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

$$\dot{x} = Cx + F(x, h(x)), \tag{III.305}$$

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

The Local Center Manifold Theorem shows that the non-hyperbolic critical point p_{aB} is indeed a stable fixed point and that the flow on the center manifold near the critical point approaches p_{aB} with speed $\frac{1}{t}$. This can be seen as follows:

By the affine transformation $(n_{aa}, n_{BB}) \mapsto (n_{aa} - \bar{n}_a, n_{BB} - \bar{n}_B)$ we get a translated system $\tilde{F}(n)$ which has a critical point at the origin. The two eigenvectors corresponding to 0 eigenvalues of the Jacobian matrix of $\tilde{F}$ at the fixed point $(0,0,0,0,0,0)$ are

$$EV_1 = (0,0,0,0,1,0,-1) \quad and \quad EV_2 = (0,0,0,0,-1,1,0). \tag{III.306}$$

We perform a new change of variable to work in the basis of eigenvectors of $\tilde{F}(n)$. Let us call the new coordinates $x_1, \ldots, x_6$. Let $h(x_1, x_2)$ be the local center manifold. We shall look at its local shape near $(0,0)$ and expand it up to second order:

$$h(x_1, x_2) = \begin{pmatrix} \lambda_3 x_1^2 + \nu_3 x_1 x_2 + \mu_3 x_2^2 \\ \lambda_4 x_1^2 + \nu_4 x_1 x_2 + \mu_4 x_2^2 \\ \lambda_5 x_1^2 + \nu_5 x_1 x_2 + \mu_5 x_2^2 \\ \lambda_6 x_1^2 + \nu_6 x_1 x_2 + \mu_6 x_2^2 \end{pmatrix} + O\left(x^3\right). \tag{III.307}$$

We then substitute the series expansions into the center manifold equation (III.304) which gives us 12 equations for the 12 unknowns $\lambda_3, \ldots, \mu_6$. Substitution of the explicit second order approximation of the center manifold equation into (III.305) yields the flow on the local center manifold:

$$\dot{x}_1 = \frac{A_1}{B_1} x_1 x_2 + \frac{C_1}{D_1} {x_2}^2 + \frac{E_1}{F_1} {x_1}^2 + O\left(x^3\right), \tag{III.308}$$

$$\dot{x}_2 = \frac{A_2}{B_2} x_1 x_2 + \frac{C_2}{D_2} {x_2}^2 + \frac{E_2}{F_2} {x_1}^2 + O\left(x^3\right), \tag{III.309}$$

where

$$A_1 = 3c^2 D f^2 - c^2 \Delta f^2 - 3c^2 f^3, \tag{III.310}$$

$$B_1 = (D - \Delta - f)\left(4cD^2 - 9cDf + c\Delta f + 5cf^2 - 4D^2\eta + 4D\Delta\eta + 8D\eta f - 4\Delta\eta f - 4\eta f^2\right), \tag{III.311}$$

$$\begin{aligned} C_1 = {} & 12c^2 D^3 f^2 - 4c^2 D^2 \Delta f^2 - 39c^2 D^2 f^3 + 12c^2 D\Delta f^3 + 42c^2 D f^4 - c^2 \Delta^2 f^3 - 8c^2 \Delta f^4 \\ & - 15c^2 f^5 + 12cD^3\eta f^2 - 16cD^2\Delta\eta f^2 - 36cD^2\eta f^3 + 4cD\Delta^2\eta f^2 + 32cD\Delta\eta f^3 \\ & + 36cD\eta f^4 - 4c\Delta^2\eta f^3 - 16c\Delta\eta f^4 - 12c\eta f^5, \end{aligned} \tag{III.312}$$

$$\begin{aligned} D_1 = {} & 8(D-2f)(D-f)(D-\Delta-f)\times \\ & \times \left(4cD^2 - 9cDf + c\Delta f + 5cf^2 - 4D^2\eta + 4D\Delta\eta + 8D\eta f - 4\Delta\eta f - 4\eta f^2\right), \end{aligned} \tag{III.313}$$

$$E_1 = cf, \quad F_1 = 2(-D + \Delta + f), \tag{III.314}$$

and

$$A_2 = 2c^2 D^2 f - 3c^2 D f^2 + c^2 f^3 - 2cD^2\eta f + 2cD\Delta\eta f + 4cD\eta f^2 - 2c\Delta\eta f^2 - 2c\eta f^3, \tag{III.315}$$

$$B_2 = (D - \Delta - f)\left(4cD^2 - 9cDf + c\Delta f + 5cf^2 - 4D^2\eta + 4D\Delta\eta + 8D\eta f - 4\Delta\eta f - 4\eta f^2\right), \tag{III.316}$$

$$C_2 = -3cD\eta f^2 + c\Delta\eta f^2 + 3c\eta f^3, \tag{III.317}$$

$$D_2 = 2(D-2f)\left(4cD^2 - 9cDf + c\Delta f + 5cf^2 - 4D^2\eta + 4D\Delta\eta + 8D\eta f - 4\Delta\eta f - 4\eta f^2\right), \tag{III.318}$$

$$E_2 = 0, \quad F_2 = 1. \tag{III.319}$$

It is left to show that the above system flows toward the origin, at least for η smaller than a certain constant. To do that, we perform another change of variables which allows us to work in the positive quadrant. We call the new coordinates (on the center manifold) y_1 and y_2, and the new field $\hat{F}$. Observe that it is sufficient to prove that the scalar product of the field with the position is negative. We thus consider the function

$$s(y_1, y_2) = \left(\hat{F}(y_1, y_2), (y_1, y_2)\right), \tag{III.320}$$

which is a quadratic form in y_1 and y_2. As the field $\hat{F}$ is homogeneous of degree 2 in its variables, it is enough to consider any direction given by $y_2 = \lambda y_1$, and prove that $s(y_1, \lambda y_1) < 0$ for all $\lambda > 0$. As the expressions are so ugly, we work perturbatively in f and consider it as large as needed. Observe that the numerator and the denominator of $s(y_1, \lambda y_1)$ are polynomials of degree 5 in f. We thus look at the coefficients in front of f^5:

$$s(y_1, \lambda y_1) = \frac{c{y_1}^3 \left(c\left(16\lambda^3 + 7\lambda^2 + 16\lambda + 40\right) - 4\eta\left(5\lambda^3 + 8\lambda^2 + 8\lambda + 8\right)\right)}{64\eta - 80c} f^5 + \Theta\left(f^4\right). \tag{III.321}$$

Observe that the denominator is always negative (because by the assumption that $\eta \leq c$). The minimal value of the ratio

$$r(\lambda) := \frac{16\lambda^3 + 7\lambda^2 + 16\lambda + 40}{4\left(5\lambda^3 + 8\lambda^2 + 8\lambda + 8\right)}, \tag{III.322}$$

is $r_{max} \simeq 0.593644$, thus, asymptotically as $f \to \infty$, the field is attractive for $\eta < c \cdot r_{max}$. Thus we see that p_{aB} is a stable fixed point which is approached with speed $\frac{1}{t}$ as long as $\eta < c \cdot r_{max}$.

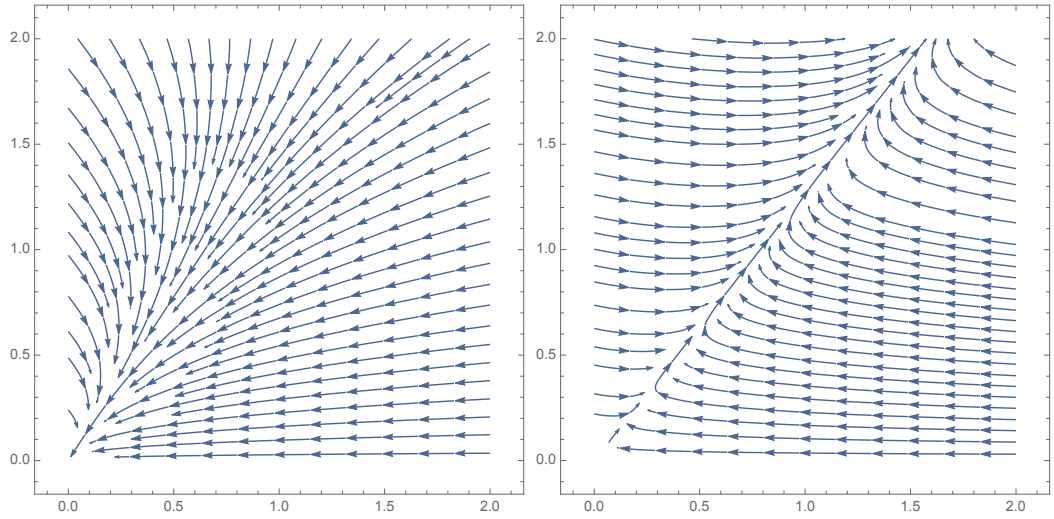

Figure III.7.: Flow of the dynamical system in the center manifold of the fixed point p_{aB}, for $\eta = 0.02$ (left) and $\eta = 0.6$ (right).

5. Discussion

In the rigorous results we have presented in this paper, we made some particular assumptions on the parameters of our model in order to simplify the analysis of the (already difficult) dynamical system. In this section we discuss which of these assumptions can be relaxed, based on heuristic considerations and numerical simulations.

The no-reproduction-small-competition model. In the model considered so far, we assume that the mutation to the B allele produces a new species different to the one of phenotype a. This is done by the *no reproduction* assumption between individuals of phenotype a and of phenotype B.

These requirements are not needed to observe the recovery of the aa population. In fact, what we require is that the invasion fitness of the aa population into a resident BB population is positive. Therefore, we can ease the no-competition assumption in our model and can add a small competition, c_{aB}, between aa individuals and BB individuals. This additional competition increases the time until aa can reinvade and also affects the two-population fixed point p_{aB} such that the two coexisting populations aa and BB will not reach their monomorphic equilibrium $\bar{n}_a$ and $\bar{n}_B$ anymore (see Figure III.8).

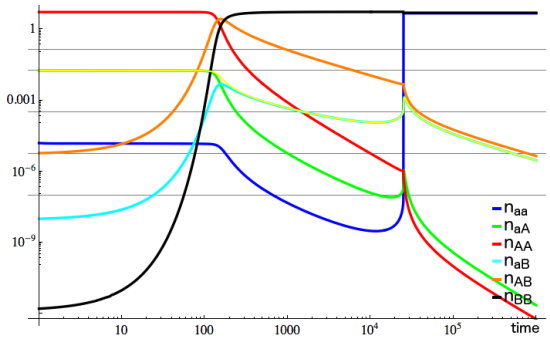

Figure III.8.: Numerical solution of the deterministic process, loglogplot for $\eta = 0$ and $c_{aB} = 0.1$.

Adding the factor η results in accelerating the process of recovery and consequently, allows to increase the competition c_{aB} (see Figure III.9).

For small η we end up in a aa-BB equilibrium but by accelerating (increasing η or decreasing c_{aB}) the process enough we can also end up in a 6-point equilibrium (all six population coexist) (see Figure III.10).

If we have competition between individuals of phenotype a and of phenotype B ($c_{aB} > 0$) the aa and BB populations have smaller equilibria as the no-competition equilibria $\bar{n}_a$ and $\bar{n}_B$, obtained when $c_{aB} = 0$. Thus the competition felt from aA by aa and BB is lower and a smaller η is enough to observe the 6-point equilibrium.

The all-with-all model. The assumption of no reproduction between individuals of phenotype a and of phenotype B is not really necessary in order to get the recovery of the aa population.

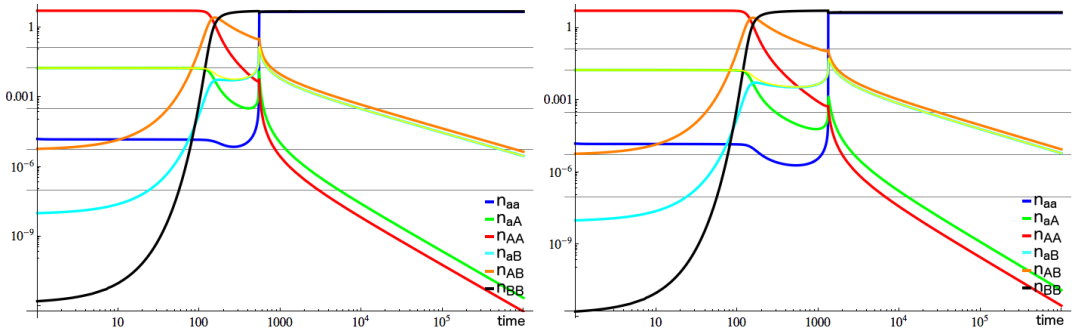

Figure III.9.: Numerical solution of the deterministic process, loglog-plot (left) for $\eta = 0.01$ and $c_{aB} = 0.1$, (right) for $\eta = 0.01$ and $c_{aB} = 0.2$.

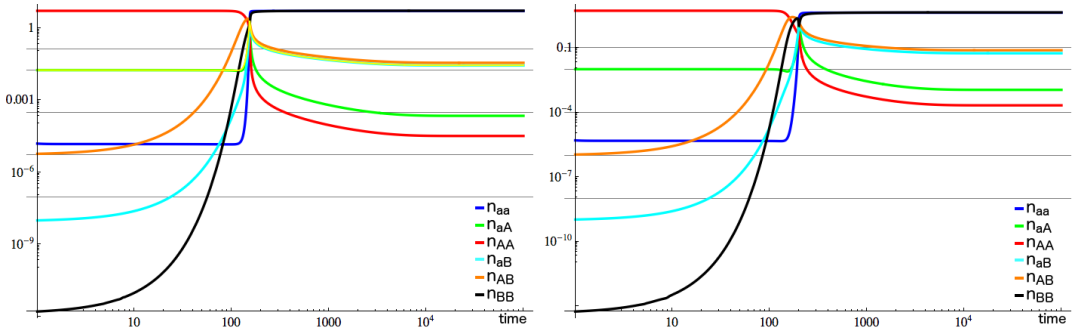

Figure III.10.: Numerical solution of the deterministic process, loglog-plot (left) for $\eta = 0.56$ and $c_{aB} = 0$, (right) for $\eta = 0.17$ and $c_{aB} = 0.2$.

Let us discuss the *all-with-all model* where all phenotypes can reproduce among themselves, that is where the reproductive compatibility is $R_i(j) = 1$, for all $i, j \in \mathcal{G}$. From numerical simulations we find that most of the results for the previous models also hold for this model but one main difference is that the 2-point equilibrium is replaced by a 3-point-equilibrium. The reason for this is that the reproduction between a and B individuals will always give birth to aB individuals and thus also an aB population survives (see Figure III.11). Also in the all-with-all model we can add small competition between individuals of phenotype a and individuals of phenotype B and get the 3-point equilibrium (see Figure III.11). As in the no-reproduction model, adding the factor η results in accelerating the process (see Figure III.12 (left)). Notice, compared to the previous models the fecundity f in the all-with-all model has to be much bigger to get the recovery of the aa population due to reproduction of all individuals among each other (see Figure III.12 (right)). Since in this case the whole population acts as potential partner for each individual the birth rate of aa scales with Σ_6 and thus f have to be big enough to compensate its death rate and to get a positive invasion fitness. With reasonable choices for η and c_{aB}, we end up in a 6-point-equilibrium where all populations coexists (see Figure III.13). Observe, the aB population can be bigger than the AB population, because it gets an additional birth factor from the reproduction of individuals of genotype aa with individuals of phenotype B which outcompetes the birth of AB individuals by reproduction of individuals of genotype AA and of phenotype B.

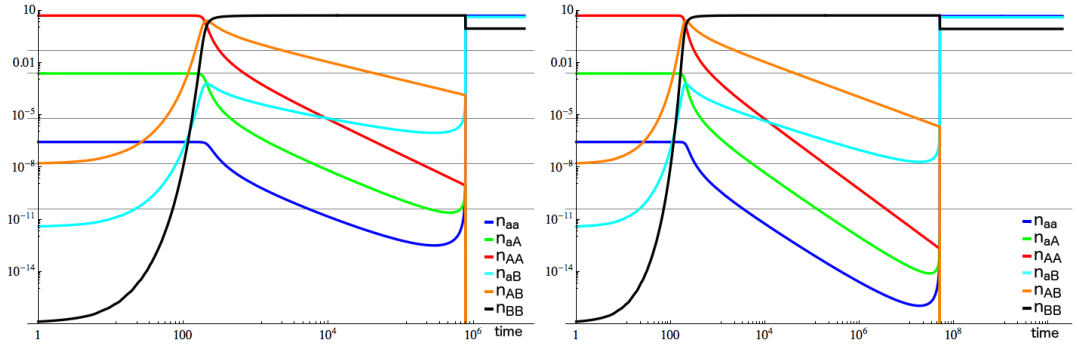

Figure III.11.: Numerical solution of the all-with-all deterministic process, loglog-plot (left) for $f = 6$, $\eta = 0$ and $c_{aB} = 0$, (right) for $f = 6$, $\eta = 0$ and $c_{aB} = 0.05$.

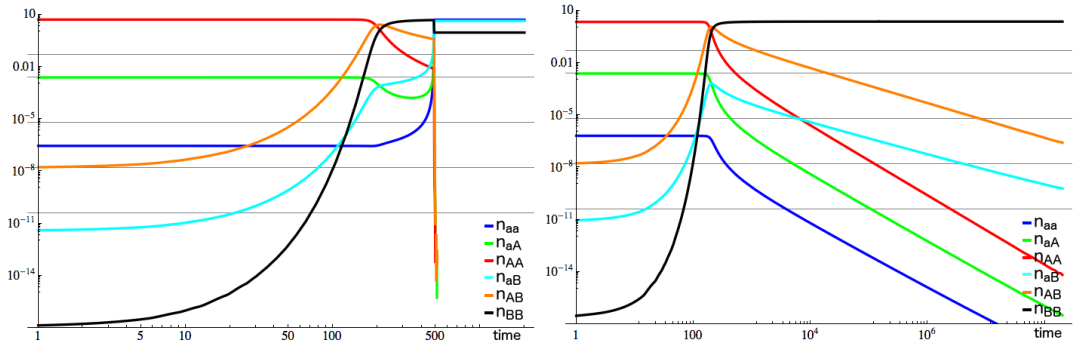

Figure III.12.: Numerical solution of the all-with-all deterministic process, loglog-plot (left) for $\eta = 0.02$ and $c_{aB} = 0$, (right) for $\eta = 0$, $c_{aB} = 0$ and $f = 3$.

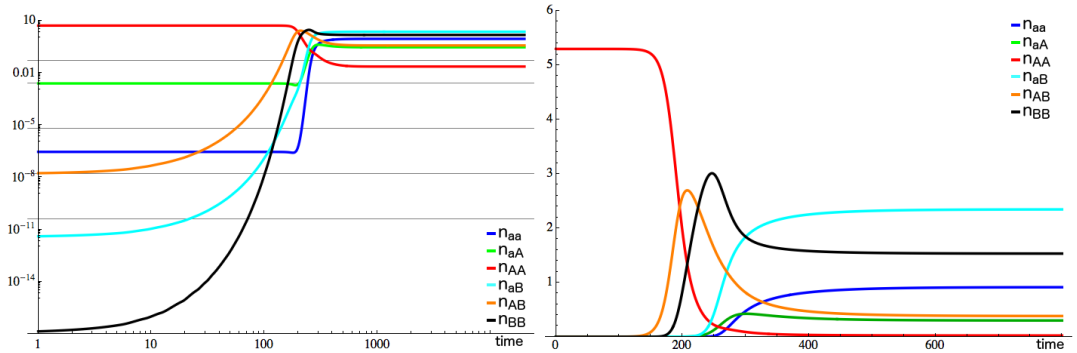

Figure III.13.: Numerical solution of the all-with-all deterministic process, loglog-plot for $\eta = 0.17$ and $c_{aB} = 0.925$ (left), rescaled individual plot for $\eta = 0.17$ and $c_{aB} = 0.925$ (left).

Bibliography

[1] K. B. Athreya and P. E. Ney. T. E. Harris and branching processes. *Ann. Probab.*, 39(2):429–434., 2011.

[2] M. Baar, A. Bovier, and N. Champagnat. From stochastic, individual-based models to the canonical equation of adaptive dynamics - In one step. *Ann. Appl. Probab.*, online first, 2016.

[3] V. Bansaye and S. Méléard. *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*, volume 1 of *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems*. Springer, Cham, 2015.

[4] N. H. Barton, D. E. G. Briggs, J. A. Eisen, D. B. Goldstein, and N. H. Patel. *Evolution*. Cold Spring Harbor Laboratory Press, 2007.

[5] G. Bell. *The Masterpiece of Nature:: The Evolution and Genetics of Sexuality*. CUP Archive, 1982.

[6] H. Bernstein, F. A. Hopf, and R. E. Michod. The molecular basis of the evolution of sex. *Advances in Genetics*, 24:323–370, 1987.

[7] S. Billiard and C. Smadi. The interplay of two mutations in a population of varying size: a stochastic eco-evolutionary model for clonal interference. *Stochastic Process. Appl.*, 127(3):701–748, 2017.

[8] B. Bolker and S. W. Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theor. Popul. Biol.*, 52(3):179–197, 1997.

[9] A. Bovier. Metastability: a potential theoretic approach. In *International Congress of Mathematicians. Vol. III*, pages 499–518. Eur. Math. Soc., Zürich, 2006.

[10] A. Bovier, L. Coquille, and R. Neukirch. The recovery of a recessive allele in a Mendelian diploid model. *arXiv preprint arXiv:1703.02459*, 2017.

[11] A. Bovier and F. den Hollander. *Metastability. A potential-theoretic approach*, volume 351 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2015.

[12] R. Bürger. *The mathematical theory of selection, recombination, and mutation*. Wiley Series in Mathematical and Computational Biology. John Wiley & Sons, Ltd., Chichester, 2000.

[13] A. Burt, M. R. Goddard, and D. Greig. Outcrossed sex allows a selfish gene to invade yeast populations. *Proc. R. Soc. Lond., B, Biol. Sci.*, 268(1485):2537–2542, 2001.

[14] N. Champagnat. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stoch. Proc. Appl.*, 116(8):1127–1160, 2006.

[15] N. Champagnat, R. Ferrière, and G. Ben Arous. The Canonical Equation of Adaptive Dynamics: A Mathematical View. *Selection*, 2:73–83., 2001.

[16] N. Champagnat, R. Ferrière, and S. Méléard. From individual stochastic processes to macroscopic models in adaptive evolution. *Stoch. Models*, 24(suppl. 1):2–44, 2008.

[17] N. Champagnat and S. Méléard. Polymorphic evolution sequence and evolutionary branching. *Prob. Theory Rel.*, 151(1-2):45–94, 2011.

[18] P. Collet, S. Méléard, and J. A. J. Metz. A rigorous model study of the adaptive dynamics of Mendelian diploids. *J. Math. Biol.*, 67(3):569–607, 2013.

[19] C. Coron. Stochastic modeling of density-dependent diploid populations and the extinction vortex. *Adv. in Appl. Probab.*, 46(2):446–477., 2014.

[20] C. Coron. Slow-fast stochastic diffusion dynamics and quasi-stationarity for diploid populations with varying size. *J. Math. Biol.*, 72(1-2):171–202., 2016.

[21] C. Coron, S. Méléard, E. Porcher, and A. Robert. Quantifying the mutational meltdown in diploid populations. *Am. Nat.*, 181(5):623–36, 2013.

[22] C. G. Correns. Mendel's Regel über das Verhalten der Nachkommenschaft der Rassenbastarde. *Berichte der Deutschen Botanischen Gesellschaft*, 18(4):158–168, 1900.

[23] J. F. Crow. An Advantage of Sexual Reproduction in a Rapidly Changing Environment. *J. Hered.*, 83(3):169, 1992.

[24] J. F. Crow and M. Kimura. *An introduction to population genetics theory*. Harper & Row, Publishers, New York-London, 1970.

[25] C. Darwin. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray, 1. edition, 1859.

[26] H. de Vries. Das Spaltungsgesetz der Bastarde. *Berichte der Deutschen Botanischen Gesellschaft*, 18(3):83–90, 1900.

[27] U. Dieckmann and M. Doebeli. On the origin of species by sympatric speciation. *Nature*, 400(6742):354–357, 1999.

[28] U. Dieckmann and R. Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.*, 34(5-6):579–612., 1996.

[29] O. Diekmann. A beginner's guide to adaptive dynamics. *Banach Center Publ.*, 63:47–86, 2004.

[30] A. Etheridge. *Some mathematical models from population genetics*, volume 2012 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011.

[31] S. N. Ethier and T. G. Kurtz. *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[32] Ewens, W. J. *Mathematical population genetics. I. Theoretical introduction*, volume 27 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, second edition, 2004.

[33] J. Felsenstein. The evolutionary advantage of recombination. *Genetics*, 78(2):737–756, 1974.

[34] R Fisher. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb.*, 42:399–433., 1918.

[35] R. A. Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, 1930.

[36] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.*, 14(4):1880–1919, 2004.

[37] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, third edition, 2012.

[38] S. A. H. Geritz and É. Kisdi. Adaptive dynamics in diploid, sexual populations and the evolution of reproductive isolation. *Proc. R. Soc. Lond., B, Biol. Sci.*, 267(1453):1671–1678, 2000.

[39] S. A. H. Geritz, É Kisdi, G. Meszéna, and J. A. J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.*, 12(1):35–57, 1998.

[40] S. A. H. Geritz, J. A. J. Metz, É. Kisdi, and G. Meszéna. Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.*, 78(10):2024, 1997.

[41] S. A. H. Geritz, E. van der Meijden, and J. A. J. Metz. Evolutionary dynamics of seed size and seedling competitive ability. *Theor. Popul. Biol.*, 55(3):324–343, 1999.

[42] J. H Gillespie. *Population genetics: a concise guide*. JHU Press, 2010.

[43] M. Golubitsky and V. Guillemin. *Stable Mappings and Their Singularities*, volume 14 of *Graduate Texts in Mathematics*. Springer US, New York, NY, 1973.

[44] J. Guckenheimer and P. J. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1990.

[45] J. B. S. Haldane. A mathematical theory of natural and artificial selection. Part I. *Trans. Camb. Phil. Soc., Biol, Sci.*, 23:19–41, 1924.

[46] J. B. S. Haldane. A mathematical theory of natural and artificial selection. Part II. *Trans. Camb. Phil. Soc., Biol, Sci.*, 1:158–163., 1924.

[47] W. D. Hamilton. Sex versus non-sex versus parasite. *Oikos*, pages 282–290, 1980.

[48] W. D. Hamilton, R. Axelrod, and R. Tanese. Sexual reproduction as an adaptation to resist parasites (a review). *Proc. Natl. Acad. Sci.*, 87(9):3566–3573, 1990.

[49] G. H. Hardy. Mendelian proportions in a mixed population. *Science*, 28(706):49–50, 1908.

[50] M. Hartfield and P. D. Keightley. Current hypotheses for the evolution of sex and recombination. *Integrative zoology*, 7(2):192–209, 2012.

[51] M. W. Hirsch, C. C. Pugh, and M. Shub. *Invariant Manifolds*, volume 583 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1977.

[52] J. Hofbauer and K. Sigmund. Adaptive dynamics and evolutionary stability. *Appl Math Lett*, 3(4):75–79, 1990.

[53] J. Jaenike. An hypothesis to account for the maintenance of sex within populations. *Evol. Theory*, 3:191–194, 1978.

[54] M. Kimura. Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626, 1968.

[55] M. Kimura. *The neutral theory of molecular evolution*. Cambridge University Press, 1983.

[56] M. Kimura. The neutral theory of molecular evolution: a review of recent evidence. *Jpn. J. Hum. Genet.*, 66(4):367–386, 1991.

[57] J. L. King and T. H. Jukes. Non-darwinian evolution. *Science*, 164(3881):788–798, 1969.

[58] M. Kirkpatrick and C. D. Jenkins. Genetic segregation and the maintenance of sexual reproduction. *Nature*, 339(6222):300–1, 1989.

[59] É. Kisdi and S. A. H. Geritz. Adaptive dynamics in allele space: Evolution of genetic polymorphism by small mutations in a heterogeneous environment. *Evolution*, 53:993–1008, 1999.

[60] É. Kisdi and S. A. H. Geritz. Evolutionary branching and sympatric speciation in diploid populations. 1999. Manuskript available at http://EconPapers.repec.org/RePEc:wop:iasawp:ir99048.

[61] É Kisdi, F. J. A. Jacobs, and S. A. H. Geritz. Red Queen evolution by cycles of evolutionary branching and extinction. *Selection*, 2(1-2):161–176, 2002.

[62] A. S. Kondrashov. Selection against harmful mutations in large sexual and asexual populations. *Genet. Res.*, 40(03):325–332, 1982.

[63] A. S. Kondrashov. Deleterious mutations and the evolution of sexual reproduction. *Nature*, 336(6198):435–440, 1988.

[64] A. S. Kondrashov. Classification of hypotheses on the advantage of amphimixis. *J. Hered.*, 84(5):372–387, 1993.

[65] E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, et al. Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921, 2001.

[66] R. Law and U. Dieckmann. *Moment Approximations of Individual based Models*. Cambridge Studies in Adaptive Dynamics. Cambridge University Press, 2000.

[67] H. Levene. Genetic equilibrium when more than one ecological niche is available. *Am. Nat.*, 87(836):331–333, 1953.

[68] W. M. Lewis. Interruption of synthesis as a cost of sex in small organisms. *Am. Nat.*, 121(6):825–833, 1983.

[69] W. M. Lewis. *The evolution of sex and its consequences: The cost of sex*, volume 4. Birkhauser Verlag, Basel, 1987.

[70] H. Lodish and S. L. Zipursky. Molecular cell biology. *Biochem. Mol. Biol. Educ.*, 29:126–133, 2001.

[71] P. Marrow, U. Dieckmann, and R. Law. Evolutionary dynamics of predator-prey systems: an ecological perspective. *J. Math. Biol.*, 34(5):556–578, 1996.

[72] P. Marrow, R. Law, and C. Cannings. The Coevolution of Predator–Prey Interactions: ESSS and Red Queen Dynamics. *Proc. R. Soc. Lond., B, Biol. Sci.*, 250(1328):133–141., 1992.

[73] G. Mendel. Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines in Brünn*, Bd. IV:3–47, 1866.

[74] G. Meszéna, M. Gyllenberg, F. J. Jacobs, and J. A. J. Metz. Link between population dynamics and dynamics of Darwinian evolution. *Phys. Rev. Lett.*, 95(7):078105, 2005.

[75] J. A. J. Metz, S. A. H. Geritz, G. Meszéna, F. J. A. Jacobs, and J. S. van Heerwaarden. Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In *Stochastic and spatial structures of dynamical systems (Amsterdam, 1995)*, Koninklijke Nederlandse Akademie van Wetenschappen Eerste Reeks, 45, pages 183–231. North-Holland, Amsterdam, 1996.

[76] J.A.J. Metz, R.M. Nisbet, and S.A.H. Geritz. How should we define "fitness" for general ecological scenarios? *Trends Ecol. Evol.*, 7(6):198–202, jun 1992.

[77] F. V. Monaghan and A. F. Corcos. Reexamination of the fate of Mendel's paper. *J. Hered.*, 78(2):116, 1987.

[78] T. H. Morgan. *Heredity and sex*. Columbia University Press, 1913.

[79] T. H. Morgan, A. H. Sturtevant, H. J. Muller, and C. B. Bridges. *The Mechanism of Mendelian heredity*. H. Holt and Company, 1915.

[80] H. J. Muller. Some genetic aspects of sex. *Am. Nat.*, 66(703):118–138, 1932.

[81] H. J. Muller. The relation of recombination to mutational advance. *Mutat. Res.*, 1(1):2–9, 1964.

[82] T. Nagylaki. *Introduction to theoretical population genetics*, volume 21 of *Biomathematics*. Springer-Verlag, Berlin, 1992.

[83] M. Nei. Modification of linkage intensity by natural selection. *Genetics*, 57(3):625, 1967.

[84] R. Neukirch and A. Bovier. Survival of a recessive allele in a Mendelian diploid model. *J. Math. Biol.*, pages 1–43, 2016.

[85] T. Ohta. Slightly Deleterious Mutant Substitutions in Evolution. *Nature*, 246 (5428):96–98, 1973.

[86] S. P. Otto. The evolutionary enigma of sex. *Am. Nat.*, 174(1):1–14, 2009.

[87] S. P. Otto and T. Lenormand. Resolving the paradox of sex and recombination. *Nat. Rev. Genet.*, 3(7):252–261, 2002.

[88] L. Perko. *Differential equations and dynamical systems*, volume 7 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 2001.

[89] M. Pigliucci. Genotype–phenotype mapping and the end of the 'genes as blueprint' metaphor. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 365(1540):557–566, 2010.

[90] Stephen R Proulx, Patrick C Phillips, and L Katz. Allelic divergence precedes and promotes gene duplication. *Evolution*, 60(5):881–892, 2006.

[91] A. Pugliese. On the evolutionary coexistence of parasite strains. *Math. Biosci.*, 177:355–375, 2002.

[92] S. Rouhani and N. Barton. The probability of peak shifts in a founder population. *J. Theor. Biol.*, 126(1):51 – 62., 1987.

[93] J. M. Smith. *The Evolution of Sex*. Cambridge University Press, 1978.

[94] E. Tschermak. Ueber künstliche Kreuzung bei Pisum sativum. *Berichte der Deutschen Botanischen Gesellschaft*, 18(6):232–239, 1900.

[95] T. J. M. Van Dooren. The evolutionary ecology of dominance-recessivity. *J. Theor. Biol.*, 198(4):519–532, 1999.

[96] Tom JM Van Dooren. The evolutionary dynamics of direct phenotypic overdominance: emergence possible, loss probable. *Evolution*, 54(6):1899–1914, 2000.

[97] G Sander Van Doorn, Ulf Dieckmann, and T Hansen. The long-term evolution of multilocus traits under frequency-dependent disruptive selection. *Evolution*, 60(11):2226–2238, 2006.

[98] L. van Valen. *A new evolutionary law*, volume 1. Evolutionary Theory, 1973.

[99] R. C. Vrijenhoek. Animal Clones and Diversity. *BioScience*, 48(8):617–628, 1998.

[100] D. Waxman and S. Gavrilets. 20 questions on adaptive dynamics. *J. Evol. Biol.*, 18(5):1139–1154, 2005.

[101] W. Weinberg. On the detection of heredity in man (in German). *Jh. Ver. Vaterl. Naturk. Wurttemb.*, 64:368–382, 1908.

[102] S. A. West, C. M. Lively, and A. F. Read. A pluralist approach to sex and recombination. *J. Evol. Biol.*, 12(6):1003–1012, 1999.

[103] S. Wright. Evolution in Mendelian populations. *Genetics*, 16(2):97–159, 1931.

[104] S. Wright. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*, 1:356–366, 1932.

[105] S Wright. Random drift and the shifting balance theory of evolution. In *Mathematical topics in population genetics*, pages 1–31. Springer, 1970.

[106] S. Wright. *Evolution and the genetics of populations, volume 3: experimental results and evolutionary deductions*, volume 3. University of Chicago press, 1984.

[107] G. U. Yule. *On the theory of inheritance of quantitative compound characters on the basis of Mendel's laws: a preliminary note*. Spottiswoode & Company, Limited, 1907.