Dynamic Fitness and Horizontal Gene Transfer in Stochastic Evolutionary Dynamics

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doktor rerum naturalium"
der Georg-August-Universität Göttingen

vorgelegt von
Hinrich Arnoldt
aus Gifhorn

Göttingen 2012

Mitglieder des Betreuungsausschusses:

Prof. Dr. Marc Timme (Referent)
Network Dynamics Group
Max Planck Institute for Dynamics and Self-Organization
Prof. Dr. Annette Zippelius
Institute for Theoretical Physics
Georg-August-Universität Göttingen
Dr. Oskar Hallatschek
Biological Physics and Evolutionary Dynamics Group
Max Planck Institute for Dynamics and Self-Organization

Weitere Mitglieder der Prüfungskommision:

Prof. Dr. Theo Geisel (Referent)
Institute for Nonlinear Dynamics
Georg-August-Universität Göttingen
Dr. Arne Traulsen
Research Group Evolutionary Theory
Max-Planck-Institute for Evolutionary Biology
Dr. Stefan Grosskinsky
Centre for Complexity Science
University of Warwick

Tag der mündlichen Prüfung: 28.01.2013

I comfirm that I have written this thesis independently and with no other sources and aids than quoted.

Göttingen,

Contents

1. Introduction

Since the dawn of our species, humans have wondered about their origin and the origin of the surrounding life. First, the general belief was that God or some other supernatural force had created everything, including life. According to this belief, the form of life remains unchanged to the present day: The book of Genesis is a good example for this creation story [49]. According to this book, God created life within a few days in the forms that still surround us today. This became the accepted view for millenia in the European culture but, beginning hundreds of years ago, Christian and Jewish scholars proposed that the account in Genesis should be seen allegorically as it rather describes different facets of God's creation than the real temporal sequence of events (cf. e.g. [101]). Furthermore, fossils were identified as petrified remnants of dead organisms. As more and more of them were found all over the world this added to the accumulating evidence that the forms of life changed over the ages and were not fixed for all time [75].

However, only after Darwin's work "On the Origin of Species" published in 1859 [15] the idea that life is in an ever changing process – nowadays called evolution – received widespread acceptance. He proposed that natural selection favours those variants of a species which are best adapted to their environment. He called these best adapted variants the "fittest". He further proposed that small variations in the species' forms arise across successive generations. Only the fittest of these variants would then survive the process of natural selection. These fittest variants could, however, have very different forms. Thus, the abundant variety of today's life could have evolved from one simple original form of life – the so-called first common ancestor. The origin of the variation across individuals was unknown to Darwin; the fact that individuals inherit genes from their ancestors was first noted by Mendel in the 1860s [54], but only became widely accepted in the early 20th century. As de Vries first noted [16], mutations may change these genes slightly in the process of inheritance and ultimately give rise to the emergence of new variants of species.

With the acceptance of the idea that evolution shapes life on earth, scientists tried to develop models to study evolutionary processes and explain the diversity of today's life. These models used to be and still are strong simplifications usually describing the evolutionary system with a small set of state variables. Although the forms of life themselves are highly complex and complex interactions in the ecosystems are omnipresent, the utilization of simplified models allows for general propositions about distinct aspects of evolutionary dynamics and thus helps to grasp the basic mechanisms of evolution. Existing models may be roughly categorized into three different types [6]. First, classic population-level models are formulated using deterministic differential equations. They describe evolution

through a view of entire populations, assuming that these populations are infinitely large so that stochastic effects of the single individuals' reproduction and death processes are neglected. Secondly, agent-based models take these stochastic effects into account by modelling the dynamics of a finite number of individuals. These individuals are assumed to be highly complex, i.e. each is described by a large number of attributes or degrees of freedom. Thirdly, individual-based models assume a finite number of very simple individuals which have very few attributes, e.g. their fitness and mutation probability. As Black and McKane proposed recently [6], individual-based models – which we use throughout this thesis – are best apt to study evolutionary dynamics. The reason, they argue, is that these models include the stochastic effects missed by population-level models, but still allow for analytical calculations and general conclusions about evolutionary dynamics that are difficult to obtain for agent-based models. In addition to this categorization into population-level, agent-based and individual-based models, there are more features distinguishing different models. One of the most important is whether or not the model includes a spatial component. A spatial structure of the environment can have a strong impact on the evolutionary dynamics as it allows for different types of populations living in different regions of the environment [42, 43, 59]. The models we use throughout this thesis completely neglect any spatial component yielding valuable insight into basic mechanisms of the underlying evolutionary processes independent of spatial effects. Thus, although reducing the complexity of the evolutionary setup, these models can still qualitatively explain evolutionary dynamics observed in reality, as well as being fitted to experimental data, e.g. to determine mutation rates [8, 10, 20, 58].

The first models of evolutionary dynamics were population-level models using a set of differential equations for the evolving species' population sizes to study the effects of selection in the course of evolution. Examples for this are the predator-prey model [93] that explains fluctuations in population sizes of both predators and its prey, or the models in classical evolutionary game theory [35, 79, 85] that study how the fitness of individuals is influenced by their mutual interactions. In such models, mutations play no or only a minor role; they are assumed to occur only very rarely and therefore on time scales much longer than the time scale imposed on the evolutionary dynamics by selection. However, as Eigen observed [23] mutations can occur at a high rate, for example in viruses. The role of high mutation rates was then adressed by Eigen in the quasispecies model [23, 24] which takes into account both selection and significant mutation rates.

It was believed that selection and random mutations are the only underlying processes of evolution until S. Wright introduced the concept of genetic drift in the 1920s [98]. Genetic drift is a stochastic effect in finite populations caused by random sampling in the reproduction process. The number of offspring that one individual produces is not deterministic, but can rather be seen as a random variable. Thus, the number of copies of a gene transferred from one generation to the next may increase or decrease randomly from generation to generation. Think for example of a gene only present in one individual of an entire population. If this individual dies before it can produce offspring, the gene is lost in the next generation and the number of copies is reduced from one to zero. This

example shows that through genetic drift genes get lost in the process of reproduction and so the diversity of the population diminishes [20]. The smaller a population the stronger is the effect of the random sampling. Therefore, deterministic population-level models may be used as an approach to describe very large populations where genetic drift is very weak, while it is important to use stochastic models to properly analyze the dynamics of small populations where genetic drift can be an important factor [20].

The idea of a tree of life that describes the relatedness among species through time has been a key concept in the theory of evolution since it was introduced by Darwin [15]. If we draw time on a vertical axis, whenever a parent individual produces offspring the genes of the parent organism are transferred vertically along this axis to its offspring. We therefore call the transfer of genes through reproduction vertical gene transfer (VGT). Considering two organisms from two species, we follow the origin of their genes back through time through repeated VGT- and mutation events up to the first parent organism which they commonly share. We may then draw two lines connecting the two species at the point in time where they share their common ancestor. Applying this procedure repeatedly, one obtains a tree-like structure with all of todays species at the branches of the tree and one organism at the root of the tree from which all life descended. This organism is often referred to as the first common ancestor; the origin of this first common ancestor, however, remains unclear. Even now, scientists are trying to determine the exact shape of such a tree of life by statistically analyzing presently known gene sequences of life forms under the assumption that evolution proceeds under the influence of selection, mutation and genetic drift [11].

However, the existence of a unifying tree-like structure for the relationship between all life-forms was recently questioned [11, 14] because in the last few decades more and more evidence accumulated that evolution is not only driven by VGT, but horizontal gene transfer (HGT) may also play a role. In general, HGT refers to the transfer of genetic material among different living cells of one generation [81, 82, 86]. In the picture of the tree of life this means a transfer of genetic information between different branchfes of the tree. Thus, the different branches of the tree would become interconnected, so that the overall existence of a unifying tree-like structure for the relationship between all life-forms was questioned [14]. While the debate on the importance of HGT for modern organisms is still ongoing [11, 14, 52, 69], a consensus seems to have been reached that HGT played a prominent role in early evolution [45]. Based on the idea of prominent HGT in early evolution, it was proposed that before the emergence of distinct species, instead of one first common ancestor there was a so-called "reactive soup" in which HGT dominated evolution [96, 97]. In this reactive soup each individual had its own distinct set of genetic material which was frequently changed by HGT-events. Due to a lack of data on early evolution, the possible evolutionary dynamics in this reactive soup are unknown and it is completely unclear how the first distinct species could have evolved from an early life environment dominated by HGT [14].

Theoretical physicists have been studying evolutionary dynamics for a long time. They

often approach the field from the point of view that evolution may be modelled as a stochastic process to be analyzed with the tools of statistical physics [20]. As Drossel put it, "the theoretical approaches lag far behind the experimental findings. While existing theoretical models and mathematical calculations cover a certain range of phenomena, verbal arguments and plausible stories prevail in many other areas, creating the need for more theoretical efforts" [20, p. 212]. The approach to tackle evolutionary problems with the methods of theoretical physics has proven to be fruitful in recent years. Early developed models, for example the works by Wright and Fisher [26, 99], Moran [55] or Kimura [39], were refined and extended, for example by combining classical evolutionary game theory [35,79,85] with the original stochastic models for finite populations [65,87,89]. By applying stochastic methods to these models it was possible to gain valuable insights into the role of selection, mutation and genetic drift in evolutionary dynamics in general. Although the models used are highly idealized, they nonetheless have repeatedly been used to compare theoretical predictions with real data and fared surprisingly well (cf. e.g. [58] and citations therein).

Yet, the attention has been mainly focused on the effects caused by selection, mutation and the dynamics' stochasticity, i.e. in vertical gene transfer processes. The impact of HGT on evolutionary dynamics remains unclear and theoretical approaches to study how HGT influences evolutionary dynamics has only recently begun. Raz and Tannenbaum analytically showed in a very simple model that HGT has a deleterious effect in static environments [73], which was confirmed in simulations by Vogan and Higgs shortly thereafter [92]. Believing in the generality of this result and because HGT is still present in today's bacteria populations one may conclude that HGT must confer an advantage for populations in *changing* environments. However, to our knowledge this has not yet been confirmed by analytical calculations or simulations. Other works studied HGT in the context of evolutionary dynamics mainly driven by mutations, i.e. they introduced it to the quasispecies model [7, 36, 68]. However, as all these studies focus on single aspects of evolutionary dynamics with HGT, we still need to clarify what the general consequences of HGT for evolutionary dynamics are.

Even in the very simple theoretical models used to study evolutionary dynamics there is still much to be explored and understood. With this thesis, we aim at improving our knowledge of the basic mechanisms underlying evolutionary dynamics. A specific aim of the thesis is to provide a first explanation of how distinct life forms could evolve from a HGT-dominated reactive soup, in particular with respect to the question which dynamical properties would need to change for such a transition in evolution to emerge. Furthermore, we analyze the role of HGT in changing environments in selection-dominated bacterial evolution. Before we attempt to conceive the role of HGT in evolution, it is important to gain a thorough understanding of the basic evolutionary dynamics without HGT. Here, the theoretical work by Traulsen et al. [87,89] on evolutionary game theory in finite populations provides a good starting point. In the first part of the thesis, we thus extend and generalize results from Traulsen, gaining a better understanding how dynamic fitness, mutations and genetic drift in general act together to shape evolutionary dynamics. Only after having

grasped the role of these basic processes in evolutionary dynamics, we focus on the role of HGT in evolution.

The work presented here is based on the established models describing selection, mutation and genetic drift [20, 55, 99], along with those describing individuals' interactions [65, 87]. We newly introduce a process effectively modelling HGT. We then aim at conceiving the resulting system dynamics for dominating HGT or dominating selection and also analyze the effect of HGT on the fitness of a population in changing environments.

The thesis is structured as follows: After this introduction we provide the fundamentals for this thesis in Chapter 2. Therein, we first give a short overview of the biological background of evolution, followed by a description of the models that we use throughout the thesis. In this part we also introduce a new HGT model which captures the essential features of HGT. At the end of Chapter 2 we provide the mathematical foundations needed for the analytic calculations in the thesis.

In Chapter 3 we address the question of how dynamic fitness, mutations and genetic drift in general act together to shape the course of evolutionary dynamics. Here, we consider a general class of functions for the form of dynamic fitness that can arise through the interactions between the individuals. Previous studies have only focussed on special linear or simple quadratic instances of this class of functions [65, 87–89], although experimental studies suggest that the dependence may be of a more complex form [51]. Our analysis reveals that such complex interactions can cause the emergence of many stable states for the dynamics, so that the population dynamics will stochastically switch between these different states. Furthermore, our studies in Chapter 3 show that the impact of fitness differences and mutations on the evolutionary dynamics scales with the population size.

To analyze how the dynamics are affected by dynamically changing population sizes we consider a model with variable population sizes in Chapter 4. There we show that such variable population size implies a rapid extinction of the population with high probability after only relatively short periods of time. We then demonstrate that dynamic fitness can stabilize the population size so the population will persist over much longer periods of time. The resulting model seems promising to study the emergence of complex evolutionary dynamics which arise due to stochastic reproduction processes and interactions between individuals. We demonstrate this by developing an ecological model which exhibits complex dynamics including quasi-cycles and punctuated equilibria. To our knowledge there is no previously existing model system which exhibits both of these evolutionary features [4, 31, 53, 60, 67].

In Chapters 5 and 6 we study the impact of HGT on evolutionary dynamics. In Chapter 5 we show that HGT can give a population a fitness advantage in changing environments. Previous studies only suggested that HGT may be beneficial for adapting populations [82], but it was only explicitly shown that HGT yields no fitness advantage in fixed environments [73, 92]. Thus, our work now confirms that HGT can be beneficial for adapting populations and our analysis reveals under which conditions this fitness advantage due to HGT emerges.

In Chapter 6 we analyze the evolutionary dynamics with frequent HGT. It has been suggested that such frequent HGT produces a reactive soup state where no distinct species exist [14]. This could have been the dominant state in evolution before the first species evolved [96, 97]. However, it remained unclear how the first distinct species could emerge from this reactive soup. Our results show that a reactive soup state emerges at high HGT rates and that the dynamics may stochastically switch between this HGT-dominated state and a selection-dominated state. Our analysis of the dynamics reveals under which conditions the reactive soup is stable and how it vanishes when the individuals' competence for HGT decreases. Thus, our results indicate a possible mechanism for the emergence of the first species from a reactive soup which we discuss at the end of Chapter 6.

We summarize and discuss all results in Chapter 7, where we also point out possible directions of future research on the role of dynamic fitness and HGT in evolutionary dynamics. Some of the details of the calculations in Chapters 2, 3 and 4 are contained in the Appendices A-D.

2. Fundamentals

2.1. Biological Background

Any evolving system relies on the basic building blocks of evolution which are replication, selection and mutation or, more generally, the introduction of new variations. This does not only include biological life as we know it, but also other evolving systems such as languages [63, 64], ideas [80] and social networks [19]. The models we study in this thesis are thus not only applicable to the evolution of life, but may also help to describe other systems. However, throughout this thesis we concentrate on the evolution of biological organisms. In the following sections we shortly introduce and discuss the basic building blocks of biological evolution.

2.1.1. From genotype to phenotype

The persistence and thriving of every form of life on earth relies on the accumulation of information about the life form's environment, which includes the physical laws determining the life form's dynamics as well as the objects with whicht it interacts. This information is stored in the *genome* of the individual organisms and determines their development and functioning. The genome is encoded in Deoxyribonucleic acid (DNA) in most organisms or Ribonucleic acid (RNA) in many viruses. Both DNA and RNA are long polymers, essentially consisting of a backbone holding the molecule together on which information carrying units are attached, called nucleobases or simply *bases*. Actually, the DNA is built up of two such polymers, the strands, forming a double helix structure coupled together by the nucleobases. The genome codes for many different functions necessary for an organism to survive and is therefore divided into different coding segments, the *genes* [28, 78]. Between the genes there are also non-coding sequences whose function is not yet completely determined; they may serve as regulatory elements for the function of the genes between which they are situated. Many organisms carry two (possibly different) copies of each gene; these organisms are called *diploid* compared to the simpler *haploid* organisms carrying only one copy of each gene. Most bacteria are haploid while sexually reproducing organisms are diploid. The specific form of a gene is called an *allele* and in diploid organisms they may either be identical – then called homozygous – or different – then called heterozygous. Each gene codes for a specific trait of an organism, which is

determined by its allele carrying the building plan for a protein. For more information on genomes, genes and DNA see e.g. the textbooks by Singer and Berg [78] or Futuyma [28].

There are four different types of bases in the DNA, namely adenine, cytosine, guanine and thymine (or uracil in RNA), abbreviated as A, C, G and T (U). These bases make up base pairs as their chemical interactions only allow for adenine to couple to thymine and for guanine to couple to cytosine. Thus, one strand of the DNA determines the other strand of the DNA as they are coupled together via these base pairs; the strands are complementary, and so the information carried within each strand is redundant. We may understand the bases as letters and thus the genome as one long word consisting of these four letters. Hence, the information of an organism about its environment determining the organism's design and functioning is stored in a long sequence of a four-letter alphabet. A given sequence defines one *genotype,* so that each individual life form may be assigned to a certain genotype. Note that two different individuals may be of the same genotype as they can have identical sequences. Typical genome lengths range from the order of 10^4 bases in simple viruses to the order of 10^9 bases in higher life forms such as e.g. humans.

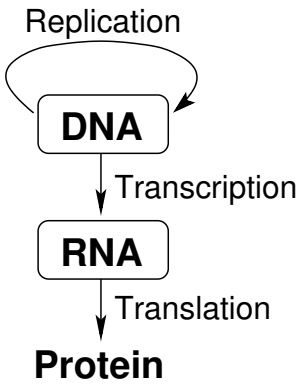

Figure 2.1.: A basic illustration of the relationship between DNA, RNA and protein, mediated by the replication, transcription and translation processes (cf. [78, p. 27]).

The DNA is in two ways an important basis for the functioning of life (cf. Figure 2.1). First, the DNA is involved in replication, where identical copies of the original DNA molecule are created. In the replication process the DNA is split up by enzymes into its two complementary strands. The complementary sequences of both strands are then recreated by an enzyme called DNA polymerase, which builds them up base by base. In this way, two complete DNA sequences are created and the storage of an organism's information is secured and handed over to new offspring. Secondly, the genome holds the construction plan for different proteins created through a complex machinery from the genetic information. We will only roughly describe the process here. For more details see for example [78]. First, in the *transcription process* a so-called messenger RNA (mRNA) is created as a copy of a sequence from the DNA. Then, the information contained in the mRNA is translated into a protein. In this *translation process* ribosomes – complex molecular machines themselves – attach to the mRNA and produce a protein while moving along the mRNA. Here, always a set of three letters – called a codon – from the sequence of the mRNA code for one amino acid which the ribosome binds together to form the resulting protein. There are 20 different amino acids that are these basic building blocks of the proteins. Each codon codes for one specific of these 20 different amino acids according to the *genetic code*. As there are $4^3 = 64$ different possibilities to combine the four different letters into a set of three and each of these 64 codons codes for one of the 20 amino

acids, the genetic code is redundant. It is universal for all life forms and has been shown to be highly optimal in correcting coding errors [27, 32]. It was therefore proposed that the genetic code evolved and outcompeted other genetic codes before life emerged in its current form [27,30]. The basic relationship between DNA, RNA and proteins through the processes of replication, transcription and translation is illustrated in Figure 2.1.

All the processes described above are highly complex and we thus refer to [28,78] for more details. However, the whole machinery involved may be essentially conceived as a self-replicating computing machine [63] because the translation process works similarly to a computing machine: The ribosome moves along a given sequence of letters and translates them into proteins according to a given code. As such a machine is deterministic, it is thus often assumed that the genotype alone defines the *phenotype* of an individual, i.e. its characteristics such as size and morphology, but also its behavior [20, 63]. However, this is not entirely true because the environment also has an influence on the phenotype of an individual, as certain constraints are put on its development [34]. Consider for example two individuals with identical genotype growing up in two very different environments, one where nutrients are in ample supply and one where it is hard to stay fed. Then, the individual living in the former environment may grow larger and the individuals will have different phenotypes. However, it is often assumed that in a given environment the genotype completely determines the phenotype of individuals [63]. This assumption is widely used in evolutionary theory [20] and works well in large enough populations. As the genotype determines the average phenotype, in large populations enough individuals of one genotype are born so that the average phenotype may well describe the entire population.

2.1.2. Natural Selection

Let us assume that the phenotype of an individual is completely determined by its genotype. Then individuals of different genotypes evolving in a given environment will exhibit differing phenotypes and thus may fare differently. Hence, their phenotype will influence the expected number of offspring they will have. We say that the individual having a higher number of expected offspring is better adapted to the environment. This is often quantified using a so-called fitness measure: The average number of offspring it will obtain compared to a reference genotype defines the *fitness* of an individual. As we assume that the genotype determines the phenotype and the phenotype defines the fitness, the fitness of an individual is thus directly determined by its genotype [25, 63, 76]. Note that the fitness of an individual may be time-dependent, because the environment may change over time. Actually, for any individual all other organisms are part of the environment: They can influence the number of offspring an individual can produce e.g. through predation or competetion for nutrients. Thus, in general the individuals mutually modify their fitness through their interactions. This is usually analyzed with evolutionary game theory [10, 35, 79, 85].

Through higher probability of reproduction the fitter genotypes in a population have a

higher probability to persist than the less fit genotypes. Therefore, on average (ignoring stochastic effects) the fitter genotypes outcompete the less fit genotypes in the long run; the fitter genotypes are selected for while the less fit genotypes vanish from the population. This process is called *natural selection*. Thus, the genotype stores the information about how to best survive and reproduce in a given environment. Generally, natural selection is a process reducing variability in a population by letting the fitter genotypes outcompete the less fit genotypes which vanish from the population. This was already realized by Darwin [15] who did not know from where the obviously wide-spread diversity in life comes. The main cause of variability are mutations which we will discuss in the following.

2.1.3. Mutations

DNA is essentially a highly complex macromolecule and as such it can change in its structure through external influences such as e.g. radiation or chemical influences [78]. Also, copying errors can occur in the process of replication when a copy of the DNA is made to be handed on to the newly created organism. In general, all these changes of the genomic sequence are called *mutations*. They introduce new genotypes to a population and therefore increase the variability in a population [20]. One of the basic features of mutations is their stochastic nature. They occur randomly and are thus unpredictable. However, some mutations are more probable than others as different parts of the DNA have different stability properties. We may thus define mutation rates from one genotype to another reflecting the frequency at which such mutations occur.

One type of mutation is the *point mutation*, meaning that only one base of the genomic sequence changes in one mutation event [28,78]. If such a mutation alters the corresponding codon in such a way that it still codes for the same amino acid, the mutation has no effect for the phenotype of the organism. Such a mutation is called *synonymous*. On the other hand, for all other, *nonsynonymous* mutations the small change in the genotype can have massive effects on the phenotype of the corresponding individual and thus strongly modify its fitness. As most populations are well adapted to their environment, mutations are typically either deleterious, i.e. they decrease the fitness of the concerned individual [25,76], or neutral, i.e. they do not affect the fitness of the individual [41]. However, for a population moving into a new environment or living in a changing environment, some mutations may be beneficial, increasing their fitness. In this way mutations are important for populations to adapt to new or changing environments [66]. Other types of mutations include frameshift mutations and sequence changes arising from recombination. They are discussed in detail for example in [28].

Typical mutation rates for the bases range from the order of 10^{-3} per base per generation in viruses to the order of 10^{-10} per base per generation in highly developed organisms such as humans (cf. [63, p. 40]). Taking into account the length of the genomes, the mutation rate per genome ranges from the order of 1 per genome in viruses to 10^{-3} in more complex

organisms. This means, that some viruses mutate about once per generation while the rate is much lower for higher developed organisms.

In general it is believed today that through mutation and natural selection new species emerge. Darwin proposed that therefore all species are related and ultimately come from one first organism, the first common ancestor [15, 20]. Looking back in time we can then establish a diagram showing the relationship of all existing species up to the first common ancestor. This diagram has a tree-like structure with time as the vertical axis and relationship distance as its horizontal axis and so is called the *tree of life*. There are different, topologically equivalent ways to depict the tree of life which vizualize the relationships between species in different ways. For example, the relationship distance can also be drawn on a circular axis and time would then lead from the inside of a circle to its rim. An example of such a diagram is shown in Figure 2.2. Currently, scientists are trying to obtain the detailed structural form of this tree of life by a statistical analysis of the genomes of different organisms [11].

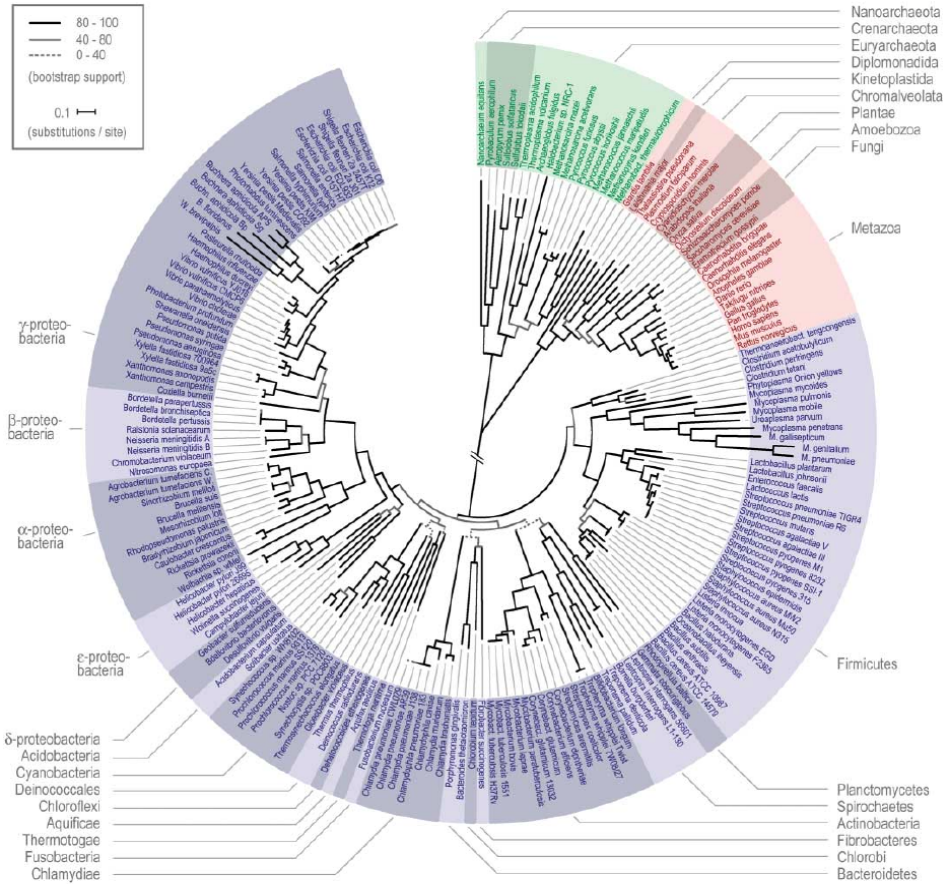

Figure 2.2.: A possible structure of the tree of life from [11]. Texts show the names of the sequenced organisms; blue denotes Bacteria, red Eukaryota and green Archaea. Lines show a possible structure of the relationship between different species. Time runs from the circle's center to its outer rim.

2.1.4. Genetic Drift

We already saw that stochasticity plays an important role for mutations and therefore evolution is not a deterministic process. Actually, stochasticity not only influences mutations, but it is also important in the reproduction process. How well adapted an individual is to a certain environment only determines its reproduction capabilities, but not the actual number of offspring it will receive. By pure chance a well adapted individual may die before receiving offspring and other less well adapted individuals may receive more offspring. Thus, the individuals of one generation are sampled in a random process to determine the individuals of the next generation. This random sampling process is called *genetic drift* [39].

This effect caused by the influence of stochasticity becomes negligible for large enough populations and disappears (theoretically) for infinitely large populations. Such populations evolving deterministically are actually studied in many models [24, 63, 79, 85, 93]. In small populations, genetic drift plays a major role that is well understood through simple mathematical models [20]. By pure chance a genotype being present in one generation may not be sampled in the reproduction process; this genotype would then be lost in the next generation. In this way the number of different genotypes present in a population can decline through the random sampling in the reproduction process. Therefore, genetic drift is a process reducing variability in a population [20]. However, different from natural selection which decreases variability in a deterministic way according to the fitness of the different genotypes, genetic drift reduces variability independent of individual's adaptation to environments and is only governed by the mathematics of chance. Thus, to reflect genetic drift, evolution models should be individual-based and use probabilistic reproduction events. Models based on deterministic differential equations are applicable in the approximation of large populations where genetic drift plays a minor role.

For a long time, it was thought that genetic drift plays a minor role in evolutionary dynamics until Motoo Kimura introduced the neutral theory of molecular evolution based on experimental estimates of amino acid substitutions in animal DNA [40]. Kimura proposes that many mutations do not affect fitness – the mutations are thus neutral –, but through genetic drift they can nontheless take hold in a population. Thus, in evolutionary dynamics, genetic drift may play an important role for mutations which do not or only slightly alter the fitness of a genotype.

2.1.5. Horizontal Gene Transfer

For a long time natural selection and mutations were thought to be the main ingredients of evolution. After Kimura's studies the importance of genetic drift was realized. However, there is another process shaping evolutionary dynamics: *Horizontal gene transfer (HGT)* [81, 82, 86], also referred to as lateral gene transfer or cross species gene transfer. Analysis of DNA-material indicates that many bacteria have aquired large portions of their

genomes via HGT [37, 46, 57]. In the picture of the tree of life, reproduction is seen as the vertical transfer of genetic material from one generation to the next. Following this picture, in general HGT refers to any transfer of genetic material between two organisms except the transfer from parent organism to its offspring in the process of reproduction. Thus, HGT may occur between totally unrelated organisms although the probability of a successful transfer is higher for more closely related organisms [86].

There are different mechansims for HGT, which we shortly discuss here. For more information see for example the review by Thomas and Nielsen [86]. First, bacteria may aquire new DNA segments through *natural transformation*, which means the uptake and integration of extracellular, free DNA. The ability of bacteria to take up this DNA is called *competence*. The individuals can stochastically switch between a noncompetent state and a compentent state depending on their environment [48]. Thus, the proportion of bacteria that are in the state of competence depends on their surrounding environment and can lie anywhere between 0 and 100 percent [12, 86]. The extracellular DNA is released in the surrounding environment either by decomposing or disrupted cells or through excretion from living cells. Another possibility for HGT to occur is *conjugative transfer*, where two cells link together for some time and build a junction from one cell to the other. Through this junction genetic material is transferred and then integrated into one or both cells' genome. A third process by which HGT occurs is *transduction*. Here a bacteriophage – a virus infecting bacteria – integrates into the DNA of a host-cell. Later it is expelled again taking with it some part of the bacteria's genetic material. Upon entering another bacterium this material is then integrated into the new host's genome.

Although HGT may play an important role in evolution, it is still heavily debated if the idea of a tree of life is also applicable in evolution considering the impact of HGT [11,14,17,52,69]: Due to frequently occuring HGT events in evolutionary history, the genes of one organism may have come from many different species. Thus, a statistical analysis cannot reveal the evolutionary history of the organism's genome in comparison to other species' genomes. This would make the construction of a unifying tree of life impossible since in this case for each gene there would exist a specific tree different from all other trees [14]. As the impact of HGT in evolutionary history is still under debate, it is not yet clear if the transfer of genetic material between different species introduces only some new connections in the tree [11] or completely destroys the notion of a tree [14, 17].

This demonstrates that there is still a lot unknown about the impact of HGT on evolutionary dynamics. There are nevertheless some theories how evolution might proceed under the influence of HGT [14, 46, 82]. It is for example proposed that HGT might help populations to adapt to rapidly changing environments [82], but there are yet only a few computational studies suggesting that in model systems HGT poses no evolutionary advantage in fixed environments [73, 92]. So, it is still unknown whether and how HGT increases a population's fitness in changing environments. Furthermore, with HGT there are different possible shapes for a "tree" of life (which look more like a bush or net) and up to now it is not yet clear which of these networks really represents the course of evolution [14, 18, 96].

Even more, Woese proposes that there is no universal first common ancestor at the root of the tree, but rather a point at which life began evolving with distinct species. Before that, he proposes, there were no distinct species but rather a soup of primordial forms rapidly exchanging genetic material via HGT [96,97]. However, how exactly evolution proceeds in such a setting and how the transition to distinct species could occur is yet unclear. This is one of the questions we address in this thesis.

There are already some theoretical studies on the impact of HGT (and also recombination [78] for sexually reproducing organisms) in the context of quasispecies theory [7,36,68] (see Section 2.2.4 for quasispecies theory). Boerlijst et al. showed [7] that in certain model settings the error threshold – a mutation rate above which no distinct species can exist – is shifted to lower mutation rates by introducing HGT to the system. Furthermore, HGT may introduce bistability between a selected state where the entire population is close to the fittest genotype and a distributed state where the population is distributed over all genotypes [36]. However, in all studies the distributed state vanishes for low mutation rates and it is not yet clear how exactly the bistability emerges. The theoretical results in all these studies [7, 36, 68] are based on population-level models where it is assumed that the populations are very large so that stochastic effects may be neglected. Yet, stochastic effects often play a major role in evolutionary dynamics which is well illustrated by the effect of genetic drift. We would like to gain a better insight into the evolutionary effects of HGT. As Black and McKane proposed recently, more general results than the one obtained from population-level models may come from individual-based models that include stochastic effects [6]. Therefore, here we introduce a new individual-based model to study HGT under the influence of stochastic dynamics. With this model we study the evolutionary effects of HGT with the aim of tackling the questions previously discussed. The models we use in this thesis are consequently explained in the next section.

2.2. Models

In this section we introduce some of the theoretical models used to describe evolutionary dynamics and also discuss results already obtained with them. We remark, that most models describe well-mixed populations in one small environment where all individuals interact with each other, i.e. the population has no spatial structure. This reduction yields a first fundamental understanding of the dynamics imposed on a population by selection, mutation, genetic drift or HGT. A spatially extended habitat adds additional complexity to the dynamics which is a research topic of its own. We only study models without spatial structure in this thesis. For evolutionary dynamics in spatially extended environments see e.g. [42, 43, 59, 63].

All the processes involved in reproduction, mutation and horizontal gene transfer are highly complex by themselves. To model these processes in detail may be the scope for simulation tools of detailed ecological setups, but most theoretical works focus on the essence of the

processes involved in evolution. Thus, the models presented here are strongly simplified, but are important for grasping the basic mechanisms driving evolution. Also, even these simple models are successfully used to analyze experimental data [58] which demonstrates that they are capable of capturing the essence of evolutionary dynamics.

2.2.1. The fitness landscape

Most models quantify the adaptation of an individual to its environment by a *fitness* measure of an individual, defined mathematically as the expected number of offspring (which themselves reach fertility before dying) produced by the individual. It is usually assumed that the genotype of an individual completely determines its fitness [20]. Thus, assigning a fitness value f_i to each genotype i yields a *fitness landscape*. Through selection and by following the paths of possible mutations populations evolve on such a fitness landscape usually by moving close to the highest peak, i.e. the fittest genotype. How exactly such a fitness landscape looks is a non-trivial problem: There are nonlinear interactions between the genes and each mutation can heavily alter the phenotype of an individual so that the mapping from genotype to fitness via the phenotype is highly complex [20].

Throughout this thesis the models we use are based on a standard model for fitness landscapes which is defined as follows [63]. Consider a population of individuals that all have a genome of fixed length l, i.e. their genotype is determined by l bases. Further consider that each base may assume two possible states, namely 0 and 1. In this model we only consider point-mutations, so that in one mutation event only one base is changed. The probability for one such mutation μ_{ij} from genotype i to j depends on the base which is changed. Then, the sequence space may be vizualized as an l-dimensional hypercube where the vertices are the genotypes and the edges are the possible mutations between the genotypes with the mutation probability μ_{ij} being the weight of an edge. In the standard model each genotype is assigned one fixed fitness value and we thus obtain a fitness landscape [20]. Figure 2.3 shows an example of such a fitness landscape for $l = 3$. The state space in this model is defined by the distribution of individuals $\underline{k}(t) = \{k_1(t), k_2(t), \ldots, k_{2^l}(t)\}$ on the different genotypes changing in time t. Here $k_i(t) \in \mathbb{N}$ is the number of individuals which are of genotype i at time t.

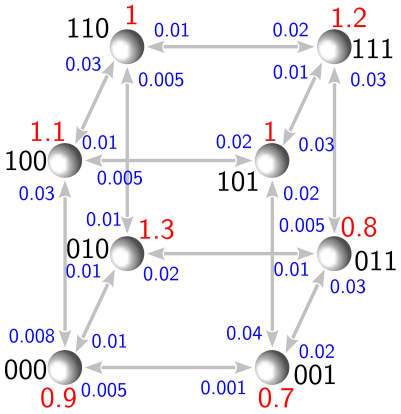

Figure 2.3.: An example fitness landscape for a genome of length $l = 3$. Nodes denote the eight different genotypes with their sequence (black, ranging from 000 to 111). Red numbers indicate the (fixed) fitness f_i of genotype i and blue numbers at the links the mutation probabilities μ_{ij} from genotype i to j. Mutations can only occur between genotypes with shown links, all other mutation probabilities are zero.

Fitness is often not a fixed property of the genotype alone as in the model above, but variable itself. This *dynamic fitness* may be caused by a changing environment [94], e.g. by the periodic fluctuations imposed on the environment by the seasons. Also, individuals interact and therefore the fitness of the individuals may depend on the possibly changing composition of the population [2,65,87,89]. This effect is called *frequency-dependent fitness* as the fitness of the individuals depends on the frequency of the different genotypes in the population. To study dynamic fitness, in this thesis we will not only use the model with fixed fitness values f_i defined above, but also extend this model to exhibit dynamic fitness. To this end we use the genotype space as defined above, but replace the fixed fitness values f_i for genotype i by state- and time-dependent fitness functions $f_i(\underline{k}, t)$. This function may depend explicitly on time t as the fitness of an individual can be explicitly time-dependent due to external influences. How the fitness depends on the distribution of individuals $\underline{k}$ is determined by their interactions which we discuss in Section 2.2.3.

To summarize, in this thesis we use models in which populations evolve on a fitness landscape that may change over time due to external environmental changes and interactions between the individuals. The basic topology of the landscape is defined by the possible mutations between the different genotypes. How exactly the population moves on this landscape depends further on the details of the reproduction and death processes. There are different models for these processes which we discuss in the following section.

2.2.2. Reproduction processes

In the model introduced in Section 2.2.1 we study the evolutionary mechanisms by analyzing the dynamics of the variable population sizes $k_i(t)$. To catch the features of the evolutionary dynamics elicited by the processes' underlying stochasticity, a fruitful approach is to apply individual-based models with finite population sizes $k_i(t) \in \mathbb{N}$ [6]. In such models simple individuals reproduce and die according to a simple stochastic reproduction process. Throughout this thesis we base our models on such stochastic reproduction processes which we introduce in the following.

The Moran process

Consider a population of overall fixed size N evolving in continuous time t on a fitness landscape as described in Section 2.2.1 with $k_i(t) \in \{0, 1, \ldots, N\}$ individuals on genotype i. The population is at all times described by the distribution of individuals $\underline{k}(t) = \{k_1(t), k_2(t), \ldots, k_{2^l}(t)\}$ with the additional condition that the overall population size is

$$\sum_{i=1}^{2^l} k_i(t) = N \tag{2.1}$$

at all times t. All individuals reproduce independently of each other and we consider the reproduction process to occur instantaneously because usually reproduction occurs on a very short time scale compared to the life span of an individual. Therefore, we call the instantaneous reproduction a reproduction event. As we consider a model with constant population size, whenever an individual produces one offspring, also one individual has to die and is removed from the population. We consider the death probability of all individuals to be equal, i.e. whenever one individual produces offspring, one individual from the population is chosen randomly with equal probability and removed from the population.

For one individual of genotype i with fitness f_i, reproduction shall be a Poisson process. Thus, the probability to reproduce in an infinitesimal time interval Δt is a time-independent constant $\Delta t \cdot f_i$ and the waiting time t_W to the next birth event of this individual is exponentially distributed with

$$p(t_W) = f_i \cdot e^{-f_i \cdot t_W} \tag{2.2}$$

so that the mean waiting time to the next birth event is given by the fitness f_i. Thus, for fitness values close to one, time is measured on the order of generations.

For the entire population, we describe the above introduced reproduction process in the following way (cf. Figure 2.4). As all individuals reproduce independently of each other as a Poisson process, the waiting time t_W to the next birth event occuring in the entire

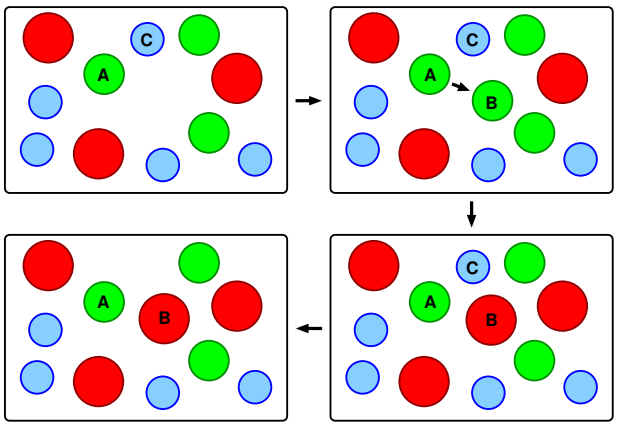

Figure 2.4.: The Moran process keeps the overall population size constant. In this example three different genotypes exhibiting different phenotypes with possibly different fitnesses are present. In the first step one of the individuals (A) of genotype i receives an identical offspring (B) with probability proportional to its fitness f_i. The newly created individual mutates to another genotype j with probability μ_{ij} and finally one of the individuals (C) is chosen with uniform probability to die. All of these steps are applied at each event of the Moran process and thus occur in zero time.

population of N individuals is given by

$$p(t_W) = N\overline{f} \cdot e^{-N\overline{f} \cdot t_W} \tag{2.3}$$

where

$$\overline{f} = \frac{1}{N} \sum_{i=1}^{2^l} f_i k_i \tag{2.4}$$

is the population's mean fitness defining the mean birth rate of the population. Which of the individuals gives birth at this event time is then determined by chosing one of the individuals with a probability proportional to its fitness f_i. This individual produces an identical offspring which may then mutate to genotype j with probability μ_{ij}. Finally, one of the individuals is chosen with equal probability to die and is thus removed from the population. Figure 2.4 illustrates these steps applied at each event time of the process.

This reproduction process is called Moran process is named after P. A. P. Moran [55] and is widely used because it captures essential features of evolution. Still, as it keeps the population size N constant, it also often allows for an analytical description of the system's dynamics [20, 88]. Originally, Moran designed this process to study a population evolving on a fitness landscape with only two genotypes and without any mutations occuring, but it was later generalized to landscapes with more genotypes and mutations in the way shown above [63, 89].

The Wright-Fisher process

Consider again a population with fixed population size N distributed on a fitness landscape as described in Section 2.2.1 with $k_i(t) \in \{0, 1, \ldots, N\}$ individuals on genotype i evolving in discrete time $t \in \mathbb{Z}$. This means that time is measured in generations. Then, the population is at all times described by the distribution of individuals $\underline{k}(t) = \{k_1(t), k_2(t), \ldots, k_{2^l}(t)\}$ and condition (2.1). Consider that each individual lives exactly for one generation and before dying may produce a number of offspring which will live in the next generation. As in the Moran process the reproduction and death of the individuals shall occur instantaneously, so that the population setup of one generation at time $t = n$ can be determined from the setup of the previous generation at time $t = n - 1$ in the following way. First, we draw one individual from the generation at $t = n - 1$ with a probability proportional to its fitness. The same type of individual is then created for the next generation at $t = n$. This process is repeated N times so that the new generation again is of size N. We remark that in this process some individuals may produce no offspring while some will produce multiple offspring. When the setup of the new generation is determined, each individual mutates with probability μ_{ij} to genotype j according to its actual genotype i. This process is illustrated in Figure 2.5 for one example time step.

This process was introduced by Wright and Fisher [26, 99] and mostly produces similar dynamics as the Moran process [20]. While the Moran process describes populations with

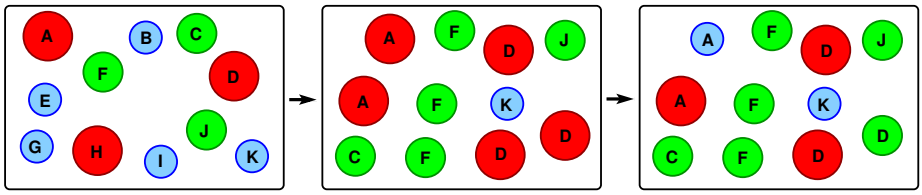

Figure 2.5.: The population setup can change strongly in one step of the Wright-Fisher process. In this example the same population as in Figure 2.4 goes through one step of the Wright-Fisher process. At each such time step, N individuals are drawn from the parent generation resulting in the population shown in the second panel. Letters in the second frame indicate the origin from the parent generation. Then, in this population mutations may occur (e.g. here the individuals in the upper left and lower right corner) leading to the third panel showing the resulting new generation.

overlapping generations, i.e. individuals of different generations interact and the individuals reproduce at different times, the Wright-Fisher process describes populations with non-overlapping generations, i.e. they reproduce and die and only after their death does the new generation arise. Both types of reproduction processes may roughly describe certain features of real evolutionary systems. For example many bacteria live and interact in overlapping generations which may rather be described by the Moran process while many insects live in non-overlapping generations which is rather capture by Wright-Fisher reproduction.

Independent birth and death process

Although the assumption of constant population size N in the above processes makes a mathematical treatment of the resulting dynamics more feasible, it may miss important effects in evolutionary dynamics arising from fluctuations in the population size. Therefore, we developed a reproduction process where the overall population size N is itself a stochastic time-dependent variable. We describe this process in the following.

We adapt the Moran process to model a process with independent birth and death events (IBD process). In the Moran process at some time one individual reproduces and at the same time another individual dies so that the population size remains constant. In the IBD process we consider a population where individuals of genotype i have a birth rate f_i given by their fitness and a constant death rate $\kappa_i = 1$, so that an individual with fitness f_i produces on average f_i offspring. Let both birth and death events occur independently with exponentially distributed waiting times t_W^B for birth and t_W^D for death events. Thus, similar to equation (2.3) the waiting times are distributed according to

$$p(t_W^B) = N\overline{f} \cdot e^{-N\overline{f} \cdot t_W^B} \tag{2.5}$$

for the birth events and

$$p(t_W^D) = N \cdot e^{-N \cdot t_W^D} \tag{2.6}$$

for the death events. As in the Moran process, when a birth event occurs one individual of genotype i is chosen with probability proportional to its fitness f_i and produces one offspring. This newly created individual may mutate with probability μ_{ij} from genotype i to genotype j . When a death event occurs, one individual is chosen randomly with equal probability and the chosen individual is removed from the system.

Here, the population size is increased by one with each birth event and decreased by one with each death event. Note however, that for a general fixed fitness $f_i \neq 1$ the population size is intrinsically unstable, as the population will on average either grow infinitely ($f_i > 1$) or go extinct ($f_i < 1$). Even more, the population can even go extinct by random fluctuations when the birth rate equals the death rate ($f_i = 1$). We will study this in more detail in Section 4 and show how the IBD process may be stabilized by a dynamic fitness that depends on the population size.

2.2.3. Individuals' interactions and game theory

As we already discussed shortly in Section 2.2.1, individuals may interact in various ways. For example, they may compete for a food source, individuals may prey on other individuals, or some individuals may benefit from mutual cooperation. One approach to study the effects of such interactions is game theory [63, 79, 85]. This theory models the interactions of two individuals through simple, well defined games. In such a game each player has the option to choose between different strategies how to interact with the other player, for example strategies A and B. However, the players get to know the strategy chosen by their opponent only after they have chosen their own strategy. Then, each player will receive a payoff P from the game, which for both players does not only depend on their own strategy, but also on the opponent's strategy. Therefore, the payoff of both individuals is determined by the payoff matrix

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array} \tag{2.7}$$

so that the payoff for playing strategy A versus A is a, while playing A versus B yields a payoff b. In the same way playing B yields c when playing versus A and d versus B. In game theory, it is analyzed how individuals can maximize their payoffs if they play such a game repeatedly against the same opponent [63].

The concept of game theory is applied to evolutionary dynamics in evolutionary game theory [10, 35, 63, 79, 85]. To this end, a model is introduced where the genotype of an individual determines its strategy in the game. Further consider, that all individuals play the game defined by the payoff matrix against all individuals at all times. Then, an

individual's overall payoff obtained from one round of these games is related to its actual fitness by adding the overall payoff to its basic fitness $f = 1$. For example, let there be j individuals of genotype A and $N - j$ individuals of genotype B. Then, in the game determined by the payoff matrix (2.7) their fitness is given by

$$f_A(j) = 1 + \frac{a(j-1) + b(N-j)}{N} \tag{2.8}$$

for individuals of genotype A as they play versus $j-1$ players of type A and $N-j$ players of type B; individuals of genotype B similarly obtain a fitness

$$f_B(j) = 1 + \frac{cj + d(N-j-1)}{N}. \tag{2.9}$$

If we let the population reproduce with these (time-dependent) fitnesses according to one of the reproduction processes defined in Section 2.2.2, we thus obtain a model for the evolution of interacting individuals. In this model the individuals' fitness is frequency-dependent as the fitnesses $f_A(j)$ and $f_B(j)$ depend linearly on the genotype frequencies j/N and $(N-j)/N$ of genotype A and B.

2.2.4. The replicator equation

In very large populations the stochastic effects of replication and mutation even out and the deviations from the expected values tend to zero. Models studying such large populations consequently study the dynamics of the frequency $x_i(t) \in [0, 1]$ with which genotype i is present in an infinitely large population. The frequency $x_i(t)$ equals one when the entire population is of genotype i at time t and it equals zero at times when genotype i is extinct. In this setup a set of deterministic differential equations for the genotype frequencies x_i describe the evolutionary dynamics [20, 63]. We will shortly discuss this approach here so that we may later compare our results with the results obtained by previous studies on evolutionary dynamics in such model systems.

To study the impact of frequency-dependent fitness on evolutionary dynamics the so-called replicator equation is often used [35, 63, 79, 85]. Consider a population evolving on a fitness landscape with M genotypes where x_i denotes the frequency of genotype i in the population as described above. We assume that mutations occur so rarely that they can be neglected on the time scale modelled by the replicator equation. Let the individuals interact according to a game (cf. Section 2.2.3) with entries a_{ij} of the payoff matrix (2.7) for genotype i individuals interacting with genotype j individuals. Then the fitness for individuals of genotype i becomes

$$f_i(\underline{x}) = \sum_{j=1}^{M} a_{ij} x_j \tag{2.10}$$

depending on the actual composition $\underline{x}$ of the population as described in Section 2.2.3.

We further assume that the population with genotype i will grow proportionally to its actual fitness $f_i(\underline{x})$ compared to the average fitness

$$\overline{f}(\underline{x}) = \sum_{k=1}^{M} x_k f_k(\underline{x}) = \sum_{j=1}^{M} \sum_{k=1}^{M} a_{jk} x_i x_j \tag{2.11}$$

in the population, which yields the replicator equation

$$\dot{x}_i = x_i \left(f_i(\underline{x}) - \bar{f}(\underline{x}) \right) \tag{2.12}$$

where both the fitness $f_i(\underline{x})$ of genotype i and the average fitness $\overline{f}(\underline{x})$ may depend on the actual composition of the population. By subtracting the mean fitness $\overline{f}(\underline{x})$ it is ensured that the population size remains constant.

As the average fitness $\overline{f}(\underline{x})$ in equation (2.11) may depend quadratically on the genotype frequencies x_i, the fitness term in the replicator equation (2.12) may thus be quadratic. Similarly, the fitness functions (2.8) and (2.9) for the individual-based model in Section 2.2.3 are linear – and will maximally become quadratic through a normalization – in the frequencies j/N and $(N-j)/N$ of both genotypes A and B. Nonlinearities of higher order in the frequencies are not possible in this model framework, i.e. game theoretic considerations of individuals' interactions always imply very specific dependencies of fitness on genotype frequency [2]. Yet, experiments suggest that fitness may also depend on genotype frequencies in a more general nonlinear way [51]. Therefore, in this thesis we apply a more general approach to frequency-dependent fitness in Chapter 3 to clarify how interactions that cause nonlinear fitness functionality influence evolutionary dynamics.

The resulting dynamical system defined by the replicator equation (2.12) and the interaction matrix a_{ij} was thoroughly analyzed [63] showing that the dynamics converge to stable fixed points where either one of the genotypes outcompetes all other genotypes or a certain mixture of genotypes is present in the population. The system may have more than one stable state so that the convergence to a state depends on the initial setup of the population. However, starting from any initial condition, the dynamics will always converge to one fixed point and then stay there for all times. Only in models considering stochastic effects can the dynamics escape from such stable states and move to other stable states due to the stochastic fluctuations. For more informations on the topic of game theory and the replicator equation see for example the book on evolutionary dynamics by M. Nowak [63].

2.2.5. The quasispecies equation

The replicator equation is commonly used to study the influence of selection in evolutionary dynamics. However, it neglects mutations which can also play a major role in evolution. Indeed, when mutations occur numerously they have a strong impact on the composition of a population. This was first described by Eigen and Schuster in the quasispecies model [23, 24]. They introduced the quasispecies equation for a set of M genotypes

$$\dot{x}_i = \sum_{j=1}^{M} x_j f_j \mu_{ji} - \overline{f} x_i \tag{2.13}$$

where similarly to the replicator equation x_i is the frequency of genotype i, f_j is the (fixed) fitness of genotype j and $\overline{f}$ is the mean fitness of the population. Newly introduced is the mutation matrix μ_{ij} determining the mutation rate from genotype i to genotype j.

Due to the mutations the population does not necessarily converge to a composition that maximizes the fitness of the population. The difference comes from the fact that the model does not only consider the speed with which different genotypes reproduce, but also the flow from one genotype to others. Thus, less fit genotypes can maintain a substantial population frequency because they receive a constant mutational input from more fit genotypes. The population will evolve to a distribution around the fittest genotype, where the mutation rate μ_{ij} and the fitness values f_i together determine the stationary distribution $\underline{x}^*$ the dynamics converge to. This distribution is called a quasispecies and is determined by the eigenvalue problem

$$\underline{x}^* W = \overline{f} \underline{x}^* \tag{2.14}$$

where $W_{ij} = f_j \mu_{ji}$ is a matrix containing the effects of both selection and mutation [63].

We conclude that the mutations tend to decrease the mean fitness of the population. The population will not converge to the genotype of maximal fitness. Not only does the stationary distribution in this setting fail to maximize the population's fitness, but the population can even completely fail to adapt to the underlying fitness landscape: If the overall mutation rate exceeds a certain value, the population is dispersed over the entire available genotype space. This value defines a critical mutation rate μ_c, where the dynamical system defined by the quasispecies equation (2.13) undergoes a bifurcation. The critical mutation rate μ_c is called *error threshold* because above it so many errors are made in the reproduction process that the population cannot adapt to the given fitness landscape. In Chapter 6 we find a similar effect for the evolution with HGT. Above a critical HGT rate a new state emerges similarly to the dispersed state caused by too many mutations. However, differently to the effects of mutation, we find that the dynamical system with HGT becomes bistable above the critical HGT rate. How we model HGT in detail is explained in the next section.

2.2.6. Stochastic modelling of Horizontal Gene Transfer

Here, we introduce a new, stochastic model for horizontal gene transfer (HGT). HGT has already been studied in the deterministic quasispecies model where it is represented by an additional term in the quasispecies equation [7, 36, 68]. However, the stochastic nature of the process may play an important role in the evolutionary dynamics so that a stochastic model is needed. Similarly to models for reproduction or mutation, we do not describe the processes involved in HGT in detail, but rather capture the essence of HGT as described in Section 2.1.5.

Basically, HGT may be described by the following process: One individual of genotype A meets an individual of type B and inserts some of the genetic material from B with a rate defined by a basic rate c and the probability k_j/N to meet B. By taking up the material, the individual of type A mutates to genotype C. This genotype C is determined by the sequence transferred to A and the position in A's genome where the new sequence is inserted. Thus, HGT introduces a 3-genotype interaction to the evolutionary dynamics. The basic rate c at which the process occurs is determined by the frequency with which the individuals meet and their competence to exchange genetic material [86] (cf. Section 2.1.5). Actually, it was observed that this competence may be a time-dependent property of the individuals [48, 86]. However, as we first want to grasp the overall influence of HGT we restrict our model to time independent base rates c for each HGT-link.

How should we model this basic process? Consider a fitness landscape as defined in Section 2.2.1 for a genome length l on which a population of N individuals evolves according to the Moran process introduced in Section 2.2.2. Thus, at all times t there is a distribution $\underline{k}(t) = \{k_1(t), k_2(t), \ldots, k_{2^l}(t)\}$ of individuals on the genotype space. To model the above described HGT events we introduce HGT-links into the genotype space (see Figure 2.6). One HGT-link is defined in the following way: We randomly choose two different genotypes A and B. Then we randomly choose a subsequence of length between 2 and $l-2$ bases from the sequence of genotype B. This sequence is then inserted at a random position into the sequence of genotype A. To keep the sequence length of A constant, the remaining bases at the end of sequence A are then cut off. The resulting sequence defines the genotype C to which the individual of genotype A will mutate through this HGT-link. If the genotype C is identical with genotype A we do not keep this HGT-link, but rather repeat the above procedure because such a HGT-event would leave the population unchanged and thus be irrelevant for the evolutionary dynamics. We repeat this procedure until a predefined number of m new HGT-links has been introduced to the system.

To clarify the introduction of HGT-links, we present a simple example in the following which is illustrated in Figure 2.6. Consider a sequence space of genomes of length $l = 4$. To create a new HGT-link in this space we randomly choose two genotypes, e.g. $A = 0100$ and $B = 1101$. We randomly choose a subsequence of length 2 from genotype B, e.g. the first two bases 11 in **11**01. This sequence is inserted at a random position into the sequence of A, e.g. position three so that the sequence of genotype A becomes $A = 01\mathbf{11}00$. Finally,

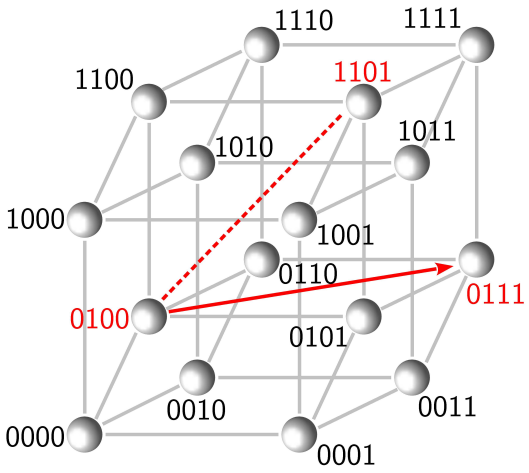

Figure 2.6.: HGT-links introduce a three-genotype interaction to the evolutionary dynamics. Here we show an example for the introduction of one HGT-link to the sequence space of genomes with the length $l = 4$ (see text). The HGT-link is marked with red: Individuals of genotype $A = 0100$ take up the first two bases 11 from genotype $B = 1101$ which are inserted at position three into $A = 011100$. The remaining bases of A are cut off so that the sequence has again length $l = 4$ and the individuals' new sequence is thus $C = 0111$.

the remaining bases of genotype A are cut off to obtain again a sequence of length l and we obtain the sequence $C = 01\mathbf{11}$. Thus, through this HGT-link individuals of genotype $A = 0100$ meeting genotype $B = 1101$ may be transformed to individuals of type $C = 0111$.

Each of these HGT-links defines one type of HGT event in which genotype A mutates to genotype C by interacting with B. We consider these events to occur independently of each other at a rate

$$r_{HGT}^{A\to C} = c \cdot k_A \frac{k_B}{N} \tag{2.15}$$

where c is the base rate of the process as described above and k_A is the number of individuals of type A which may meet an individual of genotype B with probability k_B/N according to the number of type B individuals k_B. As in the Moran process the events should occur with equal probability $\Delta t \cdot c \cdot k_A \cdot k_B/N$ in each infinitesimal time interval Δt, so that the waiting times t_W^{HGT} between HGT events of one HGT-link are exponentially distributed with

$$p(t_W^{\mathrm{HGT}}) = c \cdot k_A \frac{k_B}{N} \cdot \exp\left[-c \cdot k_A \frac{k_B}{N} \cdot t_W^{\mathrm{HGT}}\right]. \tag{2.16}$$

2.3. Mathematical fundamentals

This chapter provides the basic mathematical concepts needed for the analyses presented in this thesis. We first introduce Markov processes and then show how to describe their probabilistic dynamics using master equations. Thereafter, we describe absorbing states of Markov processes. For birth and death processes (cf. Section 2.2.2) the extinction of a population is such an absorbing state of the underlying Markov process. We present an analytic formula for the probability that a birth and death process reaches the absorbing state and the average time it needs to reach such a state. We then describe the basics of Kramers' method to obtain approximations for these absorption times. Finally, we present the general form of Fokker-Planck equations and explain their connection to master equations.

2.3.1. Markov processes

Let S be a countable set of states. Consider a time-discrete stochastic process $X(t)$, $t \in \mathbb{N}$ assuming one of the states of S at each time t, i.e. the dynamic process is described by random variables $X(t)$ indexed by the time t. The process is completely determined by the joint probability distribution

$$P\left(X(t_1) = s_1; X(t_2) = s_2; X(t_3) = s_3; \ldots\right) \tag{2.17}$$

for all ordered times $t_1 < t_2 < t_3 < \ldots$ and states $s_1, s_2, s_3 \ldots \in S$. Furthermore, we define the conditional probability distribution

$$\begin{aligned} &P\left(X(t_1) = s_1; \ldots \,|\, X(T_1) = s_2;\, X(T_2) = s_3; \ldots\right) \\ &= \frac{P\left(X(t_1) = s_1; \ldots;\, X(T_1) = s_2;\, X(T_2) = s_3; \ldots\right)}{P\left(X(T_1) = s_2;\, X(T_2) = s_3; \ldots\right)} \end{aligned} \tag{2.18}$$

if $P\left(X(T_1) = s_2;\, X(T_2) = s_3; \ldots\right) \neq 0$.

Consider the times being ordered according to $T_1 > T_2 > \ldots$. Then the *Markov assumption* states that the conditional probability distribution of the process is entirely determined by the most recent condition

$$\begin{aligned} &P\left(X(t_1) = s_1; \ldots \,|\, X(T_1) = s_2;\, X(T_2) = s_3; \ldots\right) \\ &\quad = P\left(X(t_1) = s_1; \ldots \,|\, X(T_1) = s_2\right). \end{aligned} \tag{2.19}$$

This means that the actual state of the process $X(t)$ at time t completely determines the probability for the process to be in state s after the next time step; the history of the process does not influence the processes further evolution. We say that the process is memoryless and define this process as a *time-discrete Markov process*.

For time-continuous stochastic processes the definitions are the same. Assume, that the process takes a certain state $X(t) = s_1$ at time $t \in \mathbb{R}$. Then, after a time interval Δt the probability to find the process in state s_2 is given by

$$P\left(X(t+\Delta t) = s_2 \middle| X(t) = s_1\right). \tag{2.20}$$

The memoryless property of the process now implies that the state at time $t + \Delta t$ only depends on $X(t)$. Especially, this means that that the state at time $t + \Delta t$ is independent of how long the process stayed in the state s_1 before moving on to state s_2. The only distribution of waiting times τ with this property is the exponential distribution, so that after an exponentially distributed waiting time τ the process moves on to the state s_2 depending only on the previous state s_1.

For a time-discrete Markov process being in state s_k at time t the probability to arrive in state s_l at time $t + 1$ only depends on the state s_k. We may thus assign *transition probabilities*

$$p_{k,l} = P\left(X(t+1) = s_l \middle| X(t) = s_k\right) \tag{2.21}$$

to move from state s_k to state s_l. As the process always has to move somewhere, the transition probabilities fulfill

$$\sum_l p_{k,l} = 1 \tag{2.22}$$

for all states s_k. Note, that the transition probability $p_{k,k}$ need not necessarily be zero, meaning that the process can stay in state s_k for more than one time step. A memoryless dynamical process $X(t)$ with the transition probabilities (2.21) between a countable set of states S and the initial distribution $P(X(0))$ is also called a *Markov chain*.

For continuous-time Markov processes, similarly to the time-discrete Markov process we define *transition rates* $q_{k,l}$, which determine how fast the process moves from state s_k to state s_l. Unlike the transition probabilities the transition rates fulfill the condition

$$q_{k,k} = -\sum_{l \neq k} q_{k,l} \tag{2.23}$$

so that $|q_{k,k}|$ defines how long (on average) the process will stay in state s_k. For more informations on Markov processes see e.g. the book by J. Norris [62].

2.3.2. The master equation

Assume now, that we know all transition rates $q_{k,l}$ (or probabilities $p_{k,l}$) of a continuous-time (or discrete-time) Markov process on a set S of n states. How do these transition rates determine the dynamics of the Markov process? To answer this question we need to consider how the probability of finding the process in a state s_k changes with time. For

the sake of brevity, we write

$$p_k(t) := P\left(X(t) = s_k\right) \tag{2.24}$$

for the probability to find the process in state s_k at time t. How this probability changes with time is determined by the *master equation*

$$\frac{\partial p_k(t)}{\partial t} = \sum_{l=1,\, l\neq k}^{n} \left[p_l(t) q_{l,k} - p_k(t) q_{k,l}\right] \tag{2.25}$$

for all $k \in \{1, 2, \ldots, n\}$. Here, the first term describes the probability flux into the state s_k out of all other states s_l which is given by the probability $p_l(t)$ that the process is in state s_l multiplied with the transition rate $q_{l,k}$ from state s_l to s_k. Similarly, the second term gives the probability flux out of state s_k into all other states s_l.

If only nearest neighbour transitions $q_{k,(k+1)}$ and $q_{k,(k-1)}$ are possible (i.e. $q_{kl} = 0$ for $l \notin \{k-1, k+1\}$) in an ordered set of states S, the master equation simplifies to

$$\frac{\partial p_k(t)}{\partial t} = p_{k-1}(t) q_{(k-1)k} + p_{k+1}(t) q_{(k+1)k} - p_k(t) q_{k(k+1)} - p_k(t) q_{k(k-1)}. \tag{2.26}$$

Such a system is called a linear Markov chain [38]. In this thesis we repeatedly study systems which are effectively described by such linear Markov chains.

2.3.3. Absorbing states in birth-death processes

If for a certain state s_k of a continuous-time Markov process the transition rates fulfill $q_{k,l} = 0$ for all $l \neq k$ – or the transition probabilities of a time-discrete Markov process fulfill $p_{k,l} = 0$ for all $l \neq k$ – the process can never leave this state. Such a state is called an *absorbing state* of the Markov process.

An important example of such a state is the state $N = 0$ in birth-death processes where a population of N individuals gets offspring at rate λ_N and death events occur at rate μ_N [29, 38, 62]. As the population becomes extinct on reaching the state $N = 0$, new offspring cannot emerge so that the dynamics will stay in this absorbing state. The master equation of birth-death processes takes the form of equation (2.26) with $q_{k(k+1)} = \lambda_k$ and $q_{k(k-1)} = \mu_k$. We study such systems in Chapter 4.

What is the probability p_k^H that the process will hit the absorbing state if it initially starts in a state k? Considering the transition rates λ_k and μ_k one obtains through a simple recursion formula [38]

$$p_k^H = \begin{cases} 1 & \text{if } \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} = \infty \\ \frac{\sum_{i=k}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}}{1+\sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}} & \text{if } \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} < \infty \end{cases} \tag{2.27}$$

which we derive in detail in Appendix A. Here, the sum $\sum_{i=1}^{\infty}\prod_{j=1}^{i}\mu_j/\lambda_j$ quantifies how often every state would be visited by the dynamics if there was no absorbing state. If the sum evaluates to ∞ as in the first case in (2.27), the dynamics would visit every state infinitely often if there was no absorbing state. Thus, the birth-death process is *recurrent* and it will eventually end in the absorbing state so that $p_k^H = 1$ [62]. If the sum converges, the process can go to arbitrarily large population sizes without hitting the absorbing state. It is *transient* and the dynamics will end in the absorbing state with a probability $p_k^H < 1$.

Through a similar recursion formula (cf. Appendix A) we obtain the mean time to absorption [38]

$$\overline{T}_k = \begin{cases} \infty & \text{if } \sum_{i=1}^{\infty}\chi_i = \infty \\ \sum_{i=1}^{\infty}\chi_i + \sum_{i=1}^{k-1}\left[\prod_{j=1}^{i}\frac{\mu_j}{\lambda_j}\cdot\sum_{j=i+1}^{\infty}\chi_j\right] & \text{if } \sum_{i=1}^{\infty}\chi_i < \infty \end{cases} \tag{2.28}$$

from state k where we defined $\chi_i = \frac{1}{\mu_i}\prod_{j=1}^{i-1}\lambda_j/\mu_j$. Here, the sum $\sum_{i=1}^{\infty}\chi_i$ quantifies the average time to move from state 1 to state 0. The sum ranging from $i = 1$ to $k-1$ thus quantifies how long the dynamics need on average to move from state k to state 1.

We remark that according to the conditions in equations (2.27) and (2.28) parameter settings are possible where the dynamics hit the absorbing state with probability $p_k^H = 1$, but the mean time to absorption is $\overline{T}_k = \infty$. For example, if $\mu_j = \lambda_j = 1$ for all $j \in \mathbb{N}$, the dynamics are recurrent as

$$\sum_{i=1}^{\infty}\prod_{j=1}^{i}\frac{\mu_j}{\lambda_j} = \sum_{i=1}^{\infty}1 = \infty, \tag{2.29}$$

but the mean time to absorption diverges because

$$\sum_{i=1}^{\infty}\frac{1}{\mu_i}\prod_{j=1}^{i-1}\frac{\lambda_j}{\mu_j} = \sum_{i=1}^{\infty}1 = \infty. \tag{2.30}$$

In such a setting the dynamics will hit the absorbing state eventually, but for every $t > 0$ there exists a positive fraction of trajectories which reach the absorbing state only after a waiting time $t_W > t$. Thus, the mean time to absorption diverges.

2.3.4. Kramers' method

The example of a diverging mean time to absorption in settings where the probability of absorption is $p_k^H = 1$ illustrates that it is often not sufficient to study the mean time to absorption for gaining a thorough understanding of the absorption process. Rather, it is useful to analyze the extinction time distribution $p_E(t)$, i.e. the probability distribution of the time it takes to hit the absorbing state (e.g. $N = 0$). Naturally, the shape of this distribution depends on the birth and death rates λ_N and μ_N. In Section 4 we analyze a problem where the rates are such that the dynamics of reaching the absorbing state may be seen as the escape over a potential barrier (see Figure 2.7). In this situation, we apply

Kramers' method [33, 44], which Kramers originally devised for molecular transformations between two stable states divided by a potential barrier. In [29, p. 384-386] Gardiner describes how to modify it such that we obtain the extinction time distribution for reaching an absorbing state via a potential barrier in Markovian birth-death processes.

Consider a Markovian birth-death process with an absorbing state at $k = 0$. In this system, the transition rates fulfill $\lambda_k > \mu_k$ for all $k < N_*$ and $\lambda_k < \mu_k$ for all $k > N_*$ with $N_* \in \mathbb{N}$. That means, that birth events occur at a higher rate than death events for population sizes smaller than N_* and vice versa for population sizes larger than N_*. This results in a potential well at $k = N_*$ for the dynamics as depicted in Figure 2.7. Kramers' approximation is that the potential barrier between $k = N_*$ and $k = 1$ is high enough compared to the stochastic diffusion process, so that the dynamics will first settle into a quasistationary distribution p_k^* before escaping into the absorbing state. For initial conditions $k(t = 0) \gg 1$, here high enough means, that the stationary distribution fulfills $p_{N_*}^* \gg p_1^*$. This can be rewritten to

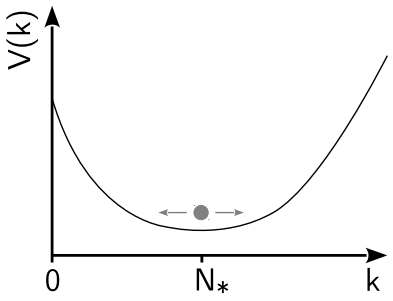

Figure 2.7.: A sketch of a potential $V(K)$ where the stochastic dynamics of a linear Markov chain may be conceived as the escape from the metastable state $k = N_*$ to the absorbing state $k = 0$ via a potential barrier.

$$\prod_{j=1}^{k-1} \frac{\lambda_{j-1}}{\mu_j} \gg 1 \tag{2.31}$$

(cf. equation (3.6)). If this condition is fulfilled, we may use a separation ansatz for the probability function $p_k(t)$. We assume that the distribution $p_k(t)$ approximately takes the form of the quasistationary distribution p_k^* multiplied with the probability $p_W(t)$ to be in the potential well, i.e. the probability that the dynamics have not reached the absorbing state at time t. Thus, the probability $p_W(t)$ to be in the well yields the survival time distribution $p_S(t)$ of the process. Formally, with this approximation the quasistationary distribution takes the form

$$p_k(t) = p_k^* \cdot p_W(t). \tag{2.32}$$

The birth-death master equation is given by

$$\frac{\partial p_k(t)}{\partial t} = \mu_{k+1} p_{k+1}(t) - \lambda_k p_k(t) - \mu_k p_k(t) + \lambda_{k-1} p_{k-1}(t) \tag{2.33}$$

and its quasistationary distribution p_k^* fulfills the detailed balance equation

$$\lambda_{k-1} p_{k-1}^* = \mu_k p_k^*, \tag{2.34}$$

i.e. the probability flow out of state k into $k-1$ equals the probability flow out of $k-1$ into

k in the quasistationary solution. In the following we use these two equations to derive a formula for the change of the probability $p_W(t)$ in time, following the arguments in [29]. For the absorbing state we define $p_0(t) = 0$, i.e. dynamics reaching the absorbing state are removed from the system. We obtain the probability flow to j from $j+1$ by summing equation (2.33) from $j+1$ to ∞ which yields

$$\frac{\partial}{\partial t} \sum_{i=j+1}^{\infty} p_i(t) = -\mu_{j+1} p_{j+1}(t) + \lambda_j p_j(t) \tag{2.35}$$

where we used the fact that $\lim_{i\to\infty} p_i(t) = 0$ as the probability is normalized to $\sum_{i=1}^{\infty} p_i(t) = 1$. Using the balance equation (2.34) we rewrite this to

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{i=j+1}^{\infty} p_i(t) &= -\mu_{j+1} p_{j+1}(t) + \lambda_j p_j(t) \\ &= -\mu_{j+1} p_{j+1}(t) + \mu_{j+1} \frac{p^*_{j+1}}{p^*_j} p_j(t) \\ &= -\mu_{j+1} p^*_{j+1} \left[\frac{p_{j+1}(t)}{p^*_{j+1}} - \frac{p_j(t)}{p^*_j} \right]. \end{aligned}$$

We divide by the factor $\mu_{j+1} p^*_{j+1}$ and sum this up from 0 to $N_* - 1$ to obtain the probability from N_* into the absorbing state. Using $p_0(t) = 0$ this yields

$$\frac{\partial}{\partial t} \sum_{j=0}^{N_*-1} \frac{\sum_{i=j+1}^{\infty} p_i(t)}{\mu_{j+1} p^*_{j+1}} = -\frac{p_{N_*}(t)}{p^*_{N_*}} \tag{2.36}$$

and using the separation ansatz (2.32) we obtain

$$\frac{\partial}{\partial t} p_W(t) \cdot \sum_{j=0}^{N_*-1} \frac{\sum_{i=j+1}^{\infty} p^*_i}{\mu_{j+1} p^*_{j+1}} = -p_W(t) \tag{2.37}$$

describing the exponential decay of the probability $p_W(t)$ to be in the potential well [29]. As mentioned above, we identify the probability $p_W(t)$ to be in the well with the survival time distribution $p_S(t)$ of the process. Thus, under the condition (2.31) the survival time distribution is well approximated by an exponential distribution

$$p_S(t) = \exp(-t/\tau) \tag{2.38}$$

which is a solution to equation (2.37). The time scale τ is given by the factor in equation (2.37) which becomes

$$\tau = \left[\sum_{j=1}^{N_*} \frac{1 - \sum_{i=1}^{j-1} p^*_i}{\mu_j p^*_j} \right]^{-1} \tag{2.39}$$

using an index shift and the normalization $\sum_{i=1}^{\infty} p^*_i = 1$.

2.3.5. The Fokker-Planck equation

Consider a system with states $x \in \mathbb{R}$ in which a deterministic force field with additional diffusion determine the dynamics. In such a system probability density functions $\rho(x,t)$ determining the probability to find the process in state x at time t are appropriate to describe the system's dynamics. The time evolution of such a probability density function $\rho(x,t)$ is described by a Fokker-Planck equation [29, 74]. Such equations may be used to describe the dynamics of high-dimensional systems. However, as we will only use Fokker-Planck equations describing one-dimensional systems in this thesis, we will only introduce the one-dimensional form of the equation here. For the general form of the equation see for example [74].

The one-dimensional *Fokker-Planck equation* (FPE)

$$\frac{\partial \rho(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[D^{(1)}(x)\rho(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[D^{(2)}(x)\rho(x,t)\right] \tag{2.40}$$

describes the time evolution of the probability density $\rho(x,t)$, where the *drift coefficient* $D^{(1)}(x)$ specifies the deterministic force field and the *diffusion coefficient* $D^{(2)}(x)$ contains the stochastic effects of the process. We remark that the FPE considers processes where both the force force field as well as the diffusion process may be state dependent, as $D^{(1)}(x)$ and $D^{(2)}(x)$ may both depend on the state x of the system.

Consider a Markov chain with N states where only nearest neighbour transitions are possible, so that the master equation takes the form of equation (2.26). For such a system with many states N, the relative step sizes k/N with $k \in \{0, 1, \ldots, N\}$ between the neighbouring states become small compared to the overall system size. Thus, the system may be approximately described by a continuous system with a variable $x = k/N \in [0,1]$; the master equation for the probability distribution $p_k(t)$ is transformed to a FPE for the probability density $\rho(x,s)$ with a rescaled time s. To achieve this transformation all system parameters as well as the system time have to be rescaled with N such that in the limit $N \to \infty$ all terms stay finite. A parameter a is hence rescaled to a parameter $\tilde{a} = a \cdot F(N)$ with a functional dependence $F(N)$ on the system size N. This functional dependence is determined by the transformation $k \to x$, so that a FPE of the form of equation (2.40) is obtained. For more details see e.g. [74] or the calculations in Appendix B.

3. Frequency-dependent fitness in evolutionary dynamics

Before we set out to analyze the evolutionary dynamics under the influence of HGT, we first need to thoroughly understand the effects imposed on the evolutionary dynamics by selection, mutation and genetic drift. In this chapter we study the effects of selection, genetic drift and mutations in a simple system of two genotypes evolving under the Moran process (cf. Section 2.2.2). Such systems of only two genotypes allow for a thorough mathematical analysis, but still yield much insight into the basic effects of selection, mutation and genetic drift in evolutionary dynamics [20, 65, 87, 89]. That a mathematical analysis of these systems is feasable stems from the fact that they are effectively one-dimensional; the frequency x of individuals of genotype A in the population determines the frequency of individuals of the other genotype B [87, 88]. The long term dynamics of the population may be analyzed using the stationary probability distribution $\rho^*(x)$ that genotype A has a frequency x in the population. In this way it was analytically shown how genetic drift and mutations impose opposing forces on the dynamics [88]. Mutations drive the population to higher diversity, genetic drift reduces such diversity. Furthermore, it was shown that, depending on the individuals' interactions, frequency-dependent selection can give rise to two different metastable states for the population dynamics [65, 88]. Due to frequency-dependent selection, the dynamics can be drawn towards a mixed state, where both genotypes are present in the population, or towards uniformity so that only one genotype remains present in the population. However, all these results were only obtained for symmetric mutation rates and special forms of frequency-dependent selection.

Here we use a more general approach where we concentrate specifically on asymmetric mutation rates and frequency-dependent fitness in a more general setting. Previous studies on the influence of frequency-dependent fitness in evolutionary dynamics were inspired by evolutionary game theory (cf. Section 2.2.3 and 2.2.4) which resulted in fitness functions $f_i(\underline{x})$ depending linearly on the frequencies x_i of the different genotypes i (cf. [65, 88] and section 2.10). However, interactions may give rise to nonlinear dependencies which has already been reported in experiment [51] and which is illustrated by the following example: Let us assume that individuals of a given genotype A cooperate so that they receive a better fitness when meeting other individuals of genotype A. Then the fitness $f_A(x_A) = 1 + a \cdot x_A$ increases linearly with the frequency x_A of genotype A [65, 88]. Yet, within the habitat in which the individuals are living there is only a limited amount of resources which all individuals of genotype A are living off. If too many of them compete

for these resources the competition will be stronger than the cooperative effects and thus the fitness of the individuals should decline. Therefore, the resulting fitness function has to contain a nonlinear factor to reflect both of these effects, e.g.

$$f_A(x_A) = 1 + a \cdot x_A - b \cdot x_A^2 \tag{3.1}$$

with $a > 0$ and $b > 0$ constants reflecting cooperative and competitive effects respectively. Such nonlinear fitness functions were not yet considered and we will analyze their impact on the dynamics in this chapter.

Furthermore, different genotypes may exhibit diverse mutation probabilities [22,77] due to each genome having its own stability properties [78]. For example, the mutation probability for genotype A to mutate to B may be very different from the probability for B to mutate to A (see also Figure 2.3). However, most studies focus on symmetric mutation rates where all mutation probabilities are identical [1,88,100]. Therefore, in this chapter we will also study asymmetric mutation probabilities.

3.1. Model setup

The model we use throughout this chapter is defined in the following way. Consider a population of N individuals with genome length $l = 1$. The individuals are hence distributed on a fitness landscape of only two genotypes A and B. The population evolves under the Moran process defined in Section 2.2.2, i.e. at exponentially distributed event times one individual produces offspring and one individual dies. Thus, the Moran process keeps the overall population size N constant. We define k_A and k_B as the population sizes on the genotypes A and B. Because the overall population size N is constant, we have the identity $k_A + k_B = N$ and thus describe the actual state of the effectively one-dimensional system with the variable $k \equiv k_A = N - k_B$. We denote the mutation probability for an individual of genotype A to mutate to B by μ_{AB} and the probability for B to mutate to A by μ_{BA}. Furthermore, the fitness functions $f_A(k) = 1 + g_A(k)$ and $f_B(k) = 1 + g_B(k)$ may take any functional form with the only restriction that $g_A(k) > -1$ and $g_B(k) > -1$ for all $k \in \{0, 1, \ldots, N\}$, because a negative fitness would mean a negative birth rate and is thus not defined. Here, we introduced the functions $g_A(k)$ and $g_B(k)$ which represent the effect of the individuals' interactions on the fitness. This system exhibits highly diverse dynamics for different parameter sets. Figure 3.1 shows three example trajectories of the variable k for three different parameter sets and also illustrates the corresponding stationary probability distributions to find the system in state k.

The dynamics of this system were already studied for special parameter sets. The studies considered identical mutation probabilities $\mu_{AB} = \mu_{BA}$ [1,88,100] or no mutations occuring at all $\mu_{AB} = \mu_{BA} = 0$ [3,84]. Furthermore, the fitness functions were derived in the context of evolutionary game theory [1,3,65,84,87,88,100], which yields fitness functions $f_A(k)$ and

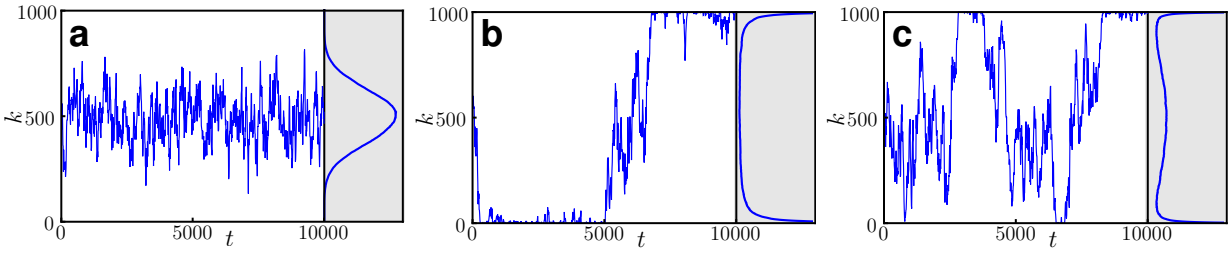

Figure 3.1.: Example trajectories of the number of individuals of genotype A for three parameter sets in the two genotype system with a population evolving under the Moran process. The gray insets show the stationary probability distributions to find the system in state k which were obtained from equation (3.19). In (a) mutations are the dominating effect ($\mu_{AB} = \mu_{BA} = 10^{-2}$) driving the dynamics towards $k = 500$. In (b) genetic drift dominates ($\mu_{AB} = \mu_{BA} = 10^{-4}$) so that the dynamics stochastically switch between the states $k = 0$ and $k = 1000$. In (c) genetic drift is stronger than the effects of mutations ($\mu_{AB} = \mu_{BA} = 10^{-4}$) as in (b), but additionally frequency-dependent fitness causes a metastable state at $k = 500$. The population size was $N = 1000$ in all simulations, in (a) and (b) there was no frequency-dependent fitness ($g_A = g_B = 0$) and in (c) the fitness functions were $g_A(k) = -10^{-5} \cdot k$ and $g_B(k) = -10^{-5} \cdot (N - k)$. In all three simulations the initial condition was $k = 500$.

$f_B(k)$ being linear or quadratic in k (cf. Section 2.2.3). However, mutation probabilities are often diverse [22, 77] and interactions may cause more complex fitness functions [51]. Therefore, here we study this system in a more general setting with different μ_{AB} and μ_{BA} as well as fitness functions that may take any functional form.

3.2. Statistical analysis

How will a population in this model system evolve under the Moran process? First we note that at each time an event of the Moran process occurs, k changes at most by ± 1. The transition probabilities at these events depend on the actual state k of the system alone, so that the system is described by a linear Markov chain of $N + 1$ states ranging from $k = 0$ to $k = N$. The fitness functions and the mutation probabilities determine the rate r_k^+ at which the transition $k \to k + 1$ occurs. At each event time of the Moran process the reproduction, mutation and death events are applied in sequence in zero time determining r_k^+ in the following way:

- **Reproduction:** The population creates an offspring of genotype A with rate $k \cdot f_A(k)$ and of genotype B with rate $(N - k) \cdot f_B(k)$ as there are k individuals of genotype A reproducing at rate $f_A(k)$ and $(N - k)$ individuals of genotype B reproducing at rate $f_B(k)$.
- **Mutation:** The offspring increases k if the newly created individual is of genotype A after applying the mutation process in the Moran step. Thus, with probability

$(1-\mu_{AB})$ an offspring of genotype A increases k, if the offspring does not mutate, and with probability μ_{BA} an offspring of genotype B increases k, if the offspring mutates to A.

- **Death:** One individual in the population is chosen with uniform probability and dies. The population of genotype A will only increase if one individual of genotype B is chosen to die. The probability that an individual of genotype B dies is $(N-k)/(N+1)$ as the individuals are chosen with equal probability.

The transition $k \to k+1$ will thus occur at a rate

$$r_k^+ = [k f_A(k) \cdot (1-\mu_{AB}) + (N-k) f_B(k) \cdot \mu_{BA}] \frac{N-k}{N+1}. \tag{3.2}$$

Analogously to the above derivation we obtain the rate

$$r_k^- = [(N-k) f_B(k) \cdot (1-\mu_{BA}) + k f_A(k) \cdot \mu_{AB}] \frac{k}{N+1} \tag{3.3}$$

for the transition $k \to k-1$.

These rates (3.2) and (3.3) yield the master equation

$$\frac{\partial p_k(t)}{\partial t} = p_{k-1}(t) r_{k-1}^+ + p_{k+1}(t) r_{k+1}^- - p_k(t) r_k^+ - p_k(t) r_k^- \tag{3.4}$$

for a linear Markov chain (2.26) describing the time evolution of the probability $p_k(t)$ of finding the population in the state with $k_A = k$ and $k_B = N-k$ at time t. The stationary solution p_k^* of the master equation yields the probability of finding the system in state k after fadeaway of initial conditions (cf. also Figure 3.1). For a linear Markov chain as described by the master equation (3.4), in the stationary solution all probability fluxes are balanced [29], i.e. in all states k the probability flux out of state k into state $k-1$ equals the probability flux out of state $k-1$ into state k. This may be understood with the following considerations. In the stationary solution $\frac{\partial p_k^*}{\partial t} = 0$, so that according to the master equation (3.4) the probability flux out of state 0 into state 1 has to equal the flux out of state 1 into 0. Then this is also true for the flux out of state 1 into state 2 and iteratively for all further states. This is quantified by the detailed balance equation

$$r_{k-1}^+ p_{k-1}^* = r_k^- p_k^*, \tag{3.5}$$

cf. also [29]. Using the balance equation (3.5) iteratively we arrive at

$$p_k^* = p_0^* \prod_{j=0}^{k-1} \frac{r_j^+}{r_{j+1}^-} \tag{3.6}$$

where we eliminate the prefactor p_0^* using the normalization condition

$$\sum_{k=0}^{N} p_k^* = 1. \tag{3.7}$$

In this way we obtain the stationary solution

$$p_k^* = \frac{\prod_{j=0}^{k-1} \frac{r_j^+}{r_{j+1}^-}}{\sum_{l=0}^{N} \prod_{j=0}^{l-1} \frac{r_j^+}{r_{j+1}^-}} \tag{3.8}$$

which may be evaluated numerically, but does not yield much insight analytically as the functional form of the stationary solution is hard to grasp in the above form (3.8).

Instead of analyzing the stationary solution of the master equation, we rather advance to the Fokker-Planck equation (see Section 2.3.5) [74] which describes the system approximately in the limit of large population sizes. The Fokker-Planck equation has a stationary solution which is more explicit than the solution of the master equation and thus yields more insight on how the evolutionary dynamics depend on the system parameters. Previous works have shown that already for population sizes $N \gtrsim 100$ the approximation works well [88].

The master equation is transformed to a Fokker-Planck equation by rescaling the system according to the population size N (cf. Section 2.3.5). Thus, all parameters have to be rescaled such that the limit $N \to \infty$ is non-degenerate which is sometimes referred to as the weak-selection limit [65, 87, 88]. We use the transformation

$$x = \frac{k}{N} \in [0,1], \qquad s = \frac{t}{N}, \qquad \tilde{\mu}_{ij} = \mu_{ij} \cdot N \tag{3.9}$$

$$\rho(x,s) = p_{xN}(sN)N, \qquad \tilde{g}_j(x) = g_j(xN)N \tag{3.10}$$

which we derive in Appendix B together with the Fokker-Planck equation corresponding to the above master equation (3.4). Here, x is now the frequency of genotype A and $(1-x)$ the frequency of genotype B. s is the rescaled time and $\tilde{\mu}_{ij}$ are the mutation rates. The probability of finding the population with a certain frequency x of genotype A at time s is then given by the probability density $\rho(x,s)$. We rescaled the interaction functions $\tilde{g}_A(x)$ and $\tilde{g}_B(x)$ but not the fitness functions as only fitness differences $f_A(x) - f_B(x) = g_A(x) - g_B(x)$ enter the Fokker-Planck equation. We obtain the Fokker-Planck equation

$$\frac{\partial \rho(x,s)}{\partial s} = -\frac{\partial}{\partial x}\Big[\{(\tilde{g}_A(x) - \tilde{g}_B(x))x(1-x) + \tilde{\mu}(1-2x) - \Delta\tilde{\mu}\}\rho(x,s)\Big] + \frac{\partial^2}{\partial x^2}\Big[x(1-x)\rho(x,s)\Big] \tag{3.11}$$

where we introduced the mean mutation rate $\tilde{\mu} := (\tilde{\mu}_{AB} + \tilde{\mu}_{BA})/2$ and the mutation rate difference $\Delta\tilde{\mu} := (\tilde{\mu}_{AB} - \tilde{\mu}_{BA})/2$. Additionally to the Fokker-Planck equation, we have the

normalization condition

$$\int_0^1 \rho(x,s)dx = 1 \tag{3.12}$$

for all $s \geq 0$.

What does the Fokker-Planck equation (2.40) tell us about the evolutionary dynamics of the system? The drift coefficient (cf. Section 2.3.5)

$$D^{(1)}(x) = (\tilde{g}_A(x) - \tilde{g}_B(x))x(1-x) + \tilde{\mu}(1-2x) - \Delta\tilde{\mu} \tag{3.13}$$

contains three different effects. The fitness difference $\tilde{g}_A(x) - \tilde{g}_B(x)$ causes a drift towards the fitter genotype which may depend on the genotype frequency as $\tilde{g}_A(x) - \tilde{g}_B(x)$ may depend on the frequency x. The mean mutation rate $\tilde{\mu}$ causes a drift towards $x = 1/2$ where both genotypes occur with equal frequency and the mutation rate difference $\Delta\tilde{\mu}$ causes a drift towards the genotype which receives more mutational input. The diffusion coefficient

$$D^{(2)}(x) = x(1-x) \tag{3.14}$$

reflects the genetic drift directed towards the edges of the system, meaning that either genotype A or genotype B take over the population.

The stationary solution of the Fokker-Planck equation is [74]

$$\rho^*(x) = Ce^{-\Phi(x)} \tag{3.15}$$

with the potential

$$\Phi(x) = \ln\left[D^{(2)}(x)\right] - \int \frac{D^{(1)}(x)}{D^{(2)}(x)}dx \tag{3.16}$$

and the normalization constant

$$C = \frac{1}{\int_0^1 e^{-\Phi(x)}dx}. \tag{3.17}$$

With the drift coefficient (3.13) and diffusion coefficient (3.14) the potential (3.16) becomes

$$\begin{aligned}\Phi(x) &= \ln\left[x(1-x)\right] - \int \left[\tilde{g}_A(x) - \tilde{g}_B(x) + \frac{\tilde{\mu}(1-2x)}{x(1-x)} - \frac{\Delta\tilde{\mu}}{x(1-x)}\right]dx \\ &= \ln\left[x(1-x)\right] - \int \left[\tilde{g}_A(x) - \tilde{g}_B(x)\right]dx - \tilde{\mu}\ln\left(x(1-x)\right) + \Delta\tilde{\mu}\ln\left[\frac{x}{1-x}\right]\end{aligned} \tag{3.18}$$

where we used $(1-2x) = \frac{\partial}{\partial x}\left[x(1-x)\right]$. We obtain the stationary solution

$$\rho^*(x) = C \cdot e^{\int[\tilde{g}_A(x) - \tilde{g}_B(x)]dx} \cdot \left[x(1-x)\right]^{\tilde{\mu}-1} \cdot \left[\frac{x}{1-x}\right]^{-\Delta\tilde{\mu}} \tag{3.19}$$

where the constant C has to be computed numerically from equation (3.17) for given parameter values.

Four different effects enter the stationary distribution, factorized in three different terms:

- The selection effects enter the solution exponentially in the term $e^{\int [\tilde{g}_A(x)-\tilde{g}_B(x)]dx}$, where the frequency-dependent interaction functions $\tilde{g}_A(x)$ and $\tilde{g}_B(x)$ can take any form and thus the selection influence depends strongly on their form. We further note that as the interaction functions $\tilde{g}_i(x) = Ng_i(k/N)$ depend linearly on the population size, already small fitness differences imply a strong selectional force in large populations.
- The term $[x(1-x)]^{\tilde{\mu}-1}$ reflects the opposing forces of mutations ($\tilde{\mu}$ in the exponent) and genetic drift (-1 in the exponent). For mean mutation rates $\tilde{\mu} < 1$ the dynamics are driven towards the edges of the system, i.e. the population has mostly individuals of one genotype; for mean mutation rates $\tilde{\mu} > 1$ a mixture of both genotypes is more probable. If $\mu_{AB} = \mu_{BA} = 1/N$ and no interactions are involved ($\tilde{g}_A(x) = \tilde{g}_B(x) = 0$) any composition k of the population has the same probability to be observed, as mutational force and genetic drift cancel each other (cf. also [88]).
- The asymmetry in the mutation rates $\Delta\tilde{\mu}$ adds an additional term $[x/(1-x)]^{-\Delta\tilde{\mu}}$ reflecting a force driving the dynamics towards a higher frequency of the genotype which receives more mutational input than the other genotype.

3.3. Analysis of the stationary solution

How do the different evolutionary forces shape the dynamics of the two-genotype system? With the stationary solution (3.19) of the Fokker-Planck equation we have now a means to answer this question. In the following we analyze the possible shapes of the stationary distribution in dependence of the system parameters, namely the interaction functions $\tilde{g}_A(x)$, $\tilde{g}_B(x)$, the mean mutation rate $\tilde{\mu}$ and the mutation rate difference $\Delta\tilde{\mu}$.

We start by studying the influence of the interaction functions on the stationary solution (3.19). First, we note that the selection term $\int \tilde{g}_A(x) - \tilde{g}_B(x)dx$ in the stationary solution has an extremum at every point x_E, where $\tilde{g}_A(x_E) - \tilde{g}_B(x_E)$ switches its sign. Thus, $\tilde{g}_A(x_E) = \tilde{g}_B(x_E)$ determines the fixed points of the (deterministic) selection dynamics. These points x_E can be either stable or unstable. They are stable, if $\tilde{g}_A(x) > \tilde{g}_B(x)$ for $x < x_E$ and $\tilde{g}_A(x) < \tilde{g}_B(x)$ for $x > x_E$ resulting in a maximum of the stationary distribution. Under these conditions, for a frequency $x < x_E$ genotype A outcompetes B and thus the frequency of A increases on average; for a frequency $x > x_E$ genotype B outcompetes A and the frequency decreases which makes x_E a stable fixed point of the selection driven dynamics. Following the same argument, a fixed point is unstable, if $\tilde{g}_A(x) < \tilde{g}_B(x)$ for $x < x_E$ and $\tilde{g}_A(x) > \tilde{g}_B(x)$ for $x > x_E$ because then the dynamics are driven away from this point. We conclude that for continuous interaction functions $\tilde{g}_A(x)$ and $\tilde{g}_B(x)$ there can be half as many metastable states of the dynamics as there are intersections of the interaction functions, because every second intersection yields a maximum of the selection

term in the stationary solution. We identify each maximum of the stationary solution with a metastable state of the dynamics because the dynamics normally stay there for a certain time before stochastically switching to another metastable state which is for example illustrated in Figure 3.3. Finally, we note that also at the edges of the system $x = 0$ or $x = 1$ there can be metastable states: If $\tilde{g}_A(0) < \tilde{g}_B(0)$ in a uniform population of genotype B individuals any newly emerging individuals of genotype A are outcompeted by genotype B individuals making $x = 0$ a metastable state, and similarly, if $\tilde{g}_A(1) > \tilde{g}_B(1)$ a metastable state at $x = 1$ emerges.

We illustrate the above considerations using example interaction functions similar to the ones provided in equation (3.1) where both genotype A and B interact only with individuals of their own genotype. Here, cooperation and resource competition yield the interaction functions

$$\tilde{g}_A(x) = N(a_A x - b_A x^2) \quad \text{and} \quad \tilde{g}_B(x) = N(a_B(1-x) - b_B(1-x)^2) \tag{3.20}$$

where the constants a_A and a_B reflect the fitness increase through cooperation of genotypes A and B and b_A and b_B the fitness decrease due to resource competition. Figure 3.2b shows an example for such interaction functions. We remark that in this example the difference $\tilde{g}_A(x) - \tilde{g}_B(x)$ of the interaction functions is of the order x^2, exhibiting maximally two zero-crossings. Hence, the selection term may in general introduce two maxima (or less) to the stationary solution, where one maximum may be given for a mixed population $0 < x_{\max} < 1$ and one for a uniform population $x_{\max} = 0$ or $x_{\max} = 1$. This is illustrated in Figure 3.2 for a parameter set where one maximum at $x \approx 0.2$ emerges as genotype A is fitter than B for $x < 0.2$ and A is less fit than B for $x > 0.2$. Another maximum is at $x = 1$ because for $x > 0.7$ genotype B outcompetes A so that B's frequency increases on average until the whole population is made up of genotype B individuals. Additionally, in this example the dynamics exhibits a maximum at $x = 0$ due to genetic drift (because $\tilde{\mu} < 1$) which also amplifies the maximum at $x = 1$. Hence, this figure illustrates that different mechanisms such as selection and genetic drift may lead to different metastable states which are determined by the maxima of the stationary distribution.

For nonlinear interaction functions there may be more than two metastable states of the dynamics, so that the population exhibits many stable genotype frequencies between which the dynamics switch back and forth stochastically. An example in Figure 3.3 illustrates that nonlinear interactions can give rise to multiple stable states which had not been realized before. Here, we considered interaction functions

$$\tilde{g}_A(x) = \alpha\,[1 + \sin(\beta x)] \quad \text{and} \quad \tilde{g}_B(x) = \alpha\,[1 + \cos(\beta x)] \tag{3.21}$$

which are periodical in the frequency x. In most applications this may not be a realistic interaction function, but it nonetheless demonstrates what is possible if a population exhibits nonlinear interaction functions. We conclude that nonlinear interaction functions may theoretically lead to an arbitrarily large number of stable states.

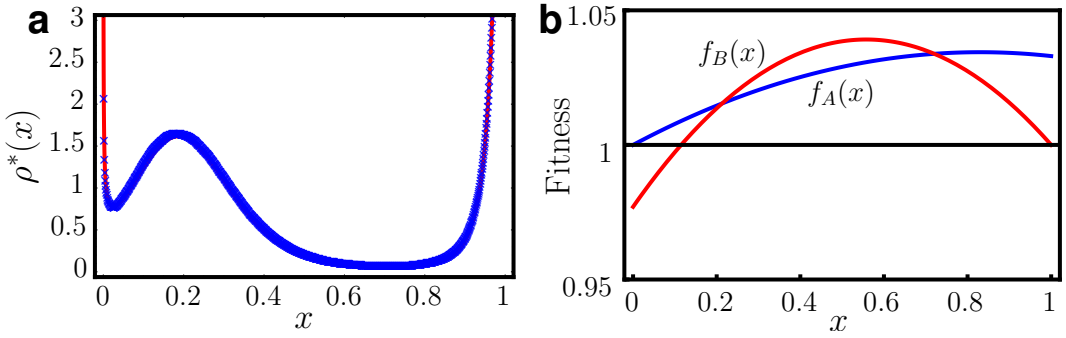

Figure 3.2.: The population dynamics exhibit metastable states due to selective forces and genetic drift. (a) shows the stationary solution (red, solid) from equation (3.19) together with data from simulations (blue, $\times$) for the fitness functions shown in (b) defined in equation (3.20). At the point $x \approx 0.2$ where the interaction function $\tilde{g}_B(x)$ intersects $\tilde{g}_A(x)$ from below the dynamics exhibit a metastable state. Furthermore, a metastable state at $x = 1$ is caused by genotype B outcompeting A for frequencies larger than $x \approx 0.7$. The metastable state at $x = 0$ is caused by genetic drift, which also enhaces the stability of the state at $x = 1$. The population size for the simulation was $N = 1000$ showing that the Fokker-Planck equation approximates the dynamics well already for relatively small population sizes. Further parameters were $a_A = 0.083$, $b_A = 0.05$, $a_B = 0.177$ and $b_B = 0.2$ and the mutation rate was set to $\tilde{\mu} = 0.5$ with no asymmetry ($\Delta\tilde{\mu} = 0$).

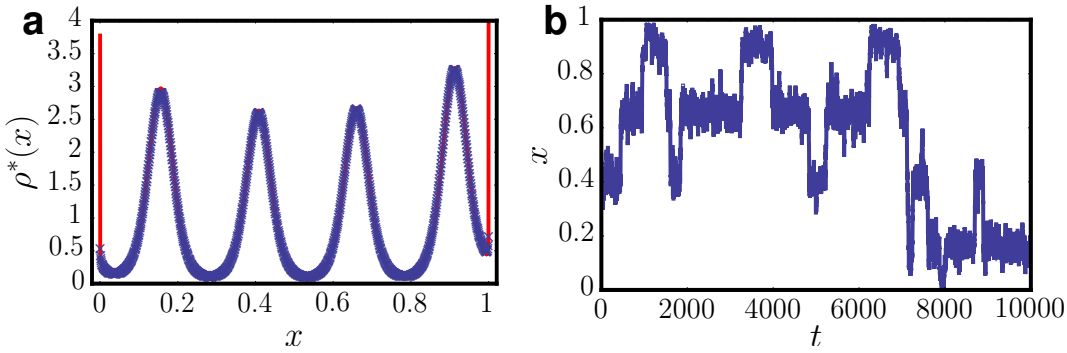

Figure 3.3.: The stationary solution (3.19) exhibits multiple metastable states for periodic interaction functions. (a) shows the stationary solution (red, solid) from equation (3.19) for the periodic fitness functions defined in equation (3.21) together with data from simulations (blue, $\times$). Theoretically genetic drift also causes metastable states at $x = 0$ and $x = 1$, but due to the finite number of individuals ($N = 1000$) in the simulations the maxima of the stationary solution remain small at these points. (b) shows a sample path demonstrating repeated stochastic switching between the different metastable states. The paramters for this system were $\alpha = 30$, $\beta = 25\tilde{\mu} = 0.5$ and $\Delta\tilde{\mu} = 0$.

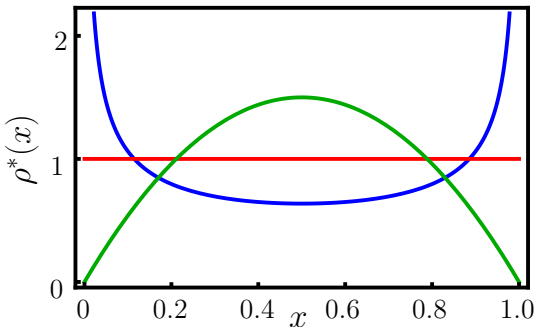

Figure 3.4.: At a critical mean mutation rate $\tilde{\mu} = 1$ the effects of genetic drift and mutations cancel each other. Shown are three stationary solutions (3.19) of the Fokker-Planck equation for a population without selective pressure ($f_A = f_B = 1$) with three different mean mutation rates $\tilde{\mu}$ without asymmetry $\Delta\tilde{\mu} = 0$. For small mutation rates genetic drift pushes the dynamics to the boundaries which is illustrated for $\tilde{\mu} = 0.5$ (blue). For the critical mutation rate $\tilde{\mu} = 1$ (red) the forces of genetic drift and mutation are equally strong so that a uniform distribution is observed and for higher mutation rates the dynamics are pushed by the numerously occuring mutations towards a mixed population $x = 0.5$ which is here illustrated for $\tilde{\mu} = 2$ (green).

Let us advance to study the influence of the mutations on the population dynamics. It is already known [88], that in a population without selective pressure mean mutation rate and genetic drift determine the distribution of individuals. The forces of genetic drift and mutations are opposed to each other, which is reflected by the term $(x(1-x))^{\tilde{\mu}-1}$ in the stationary solution (3.19). For $\tilde{\mu} = 1$ these forces cancel out and the population may be found in any state with equal probability. For smaller mutation rates $\tilde{\mu} < 1$ genetic drift drives the dynamics towards the system's edges so that either genotype A or genotype B dominates the dynamics which stochastically switch between these two states (cf. Figure 3.1b). For larger mutation rates $\tilde{\mu} > 1$ a mixture of both genotypes is maintained as the often occuring mutations push the dynamics towards a frequency of $x = 0.5$. The different shapes of the stationary distribution for the different mutation rates are illustrated in Figure 3.4 (cf. also Figure 1 in [88]).

While the influence of the mean mutation rate $\tilde{\mu}$ on the population dynamics was already well understood, the influence of asymmetric mutation rates has not yet been studied in this model system. We note, that the term $(x/(1-x))^{-\Delta\tilde{\mu}}$ may have a strong impact on the stationary solution (3.19) near $x = 0$ for $\Delta\tilde{\mu} > 0$ and near $x = 1$ for $\Delta\tilde{\mu} < 1$. Furthermore, per definition $\Delta\tilde{\mu} \in [-\tilde{\mu}, \tilde{\mu}]$ so that, if the mean mutation rate $\tilde{\mu}$ is small, the influence of asymmetric mutation rates on the stationary solution is small, too. However, for large mean mutation rates $\tilde{\mu}$ the shape of the stationary solution may strongly depend on the asymmetry of the mutation rates if $\Delta\tilde{\mu}$ is large. Under these conditions – due to the asymmetric mutation rates – the population dynamics will be drawn towards an increased frequency of the genotype that receives more mutational input. Figure 3.5a illustrates that the asymmetry in the mutation rates can elicit the emergence of a new metastable state at the edge of the system. Actually, as Figure 3.5b shows, this newly emerging metastable

state even minimizes the fitness of the population and is thus clearly not induced by selection. On the other hand, as Figure 3.6 illustrates, the asymmetry can also influence the dynamics in such a way that metastable states are shifted or even vanish for high enough mutational asymmetries.

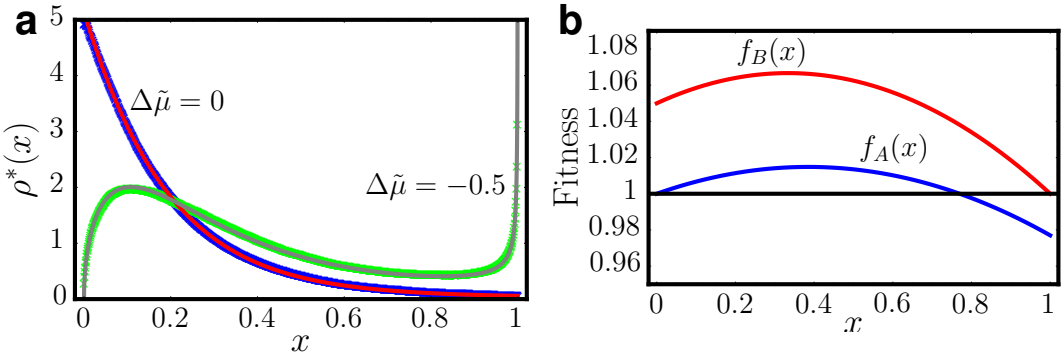

Figure 3.5.: New metastable states emerge due to asymmetric mutation rates. (a) shows the stationary distriubtions of a system where genotype B outcompetes genotype A for symmetric ($\Delta\tilde{\mu} = 0$) and asymmetric mutation rates ($\Delta\tilde{\mu} = -0.5$). The red and gray solid curves show the theoretically obtained solution (3.19) of the Fokker-Planck equation and the blue and green crosses data from simulations with $N = 1000$. (b) shows the fitness functions of the genotypes demonstrating that the new metastable state emerging for $\Delta\tilde{\mu} = -0.5$ minimizes the fitness of the population. The mean mutation rate was $\tilde{\mu} = 1$ and the interaction functions were given according to equation (3.20) with $a_A = 0.077$, $b_A = 0.1$, $a_B = 0.2$ and $b_B = 0.15$.

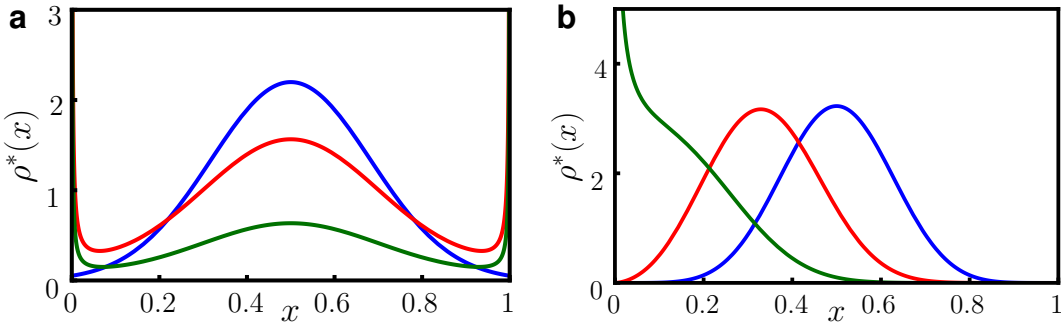

Figure 3.6.: At high mutation rates $\tilde{\mu}$ the asymmetry $\Delta\tilde{\mu}$ of the mutation rates has a strong impact on the population dynamics. (a) shows that for low mutation rates $\tilde{\mu} = 0.01$ (green) and $\tilde{\mu} = 0.1$ (red) genetic drift has a strong impact at the edges of the system creating metastable states at $x = 0$ and $x = 1$ which vanish at the critical mutation rate $\tilde{\mu} = 1$ (blue). (b) illustrates the effect of asymmetries for high mutation rates (here $\tilde{\mu} = 5$) where a metastable state at $x = 0.5$ for symmetric mutation rates $\Delta\tilde{\mu} = 0$ (blue) is first shifted for intermediate asymmetries $\Delta\tilde{\mu} = 2.5$ (red) and finally vanishes for strong asymmetries $\Delta\tilde{\mu} = 4.5$ (green). Here, the parameters $a_A = a_B = -0.01$ and $b_A = b_B = 0.005$ in the interaction functions (3.20) are chosen such that selection drives the dynamics towards $x = 0.5$.

3.4. The quality of the Fokker-Planck approximation

Here, we quantify the quality of the approximation we made on switching from a Markovian to a Fokker-Planck description of the system's dynamics. In Figure 3.2, Figure 3.3 and 3.5 the theoretical solutions well fit the data from simulations with popualtion sizes $N = 1000$; in a similar study Traulsen et al. [88] find good fits already for $N = 100$. In the following we quantitatively analyze how the quality of the Fokker-Planck approximation depends on system parameters such as the population size N. We define the empirical distribution

$$\pi_k := \frac{1}{T_{\text{meas}}} \sum_{t=0}^{T_{\text{meas}}} \delta(X_{t+T_{\text{mix}}}, k) \tag{3.22}$$

to compare our theoretically obtained solution (3.19) of the Fokker-Planck equation with data from simulations. Here, $(X_t : t \geq 0)$ is the evolutionary process defined by the master equation (3.4), T_{mix} is a time large enough for the process to reach stationarity and T_{meas} is the measurement time of the simulation. We quantify the fit's quality using the mean distance measure

$$\overline{d} := \frac{1}{N} \sum_{k=1}^{N-1} \left| \pi_k - \int_{\frac{k}{N}-\frac{1}{2N}}^{\frac{k}{N}+\frac{1}{2N}} \rho^*(x)dx \right| \tag{3.23}$$

comparing the mean distance of the empirical distribution π_k from the theoretical distribution, and the maximum distance measure

$$d_{\max} := \max_{k\in[1,N-1]} \left\{ \left| \pi_k - \int_{\frac{k}{N}-\frac{1}{2N}}^{\frac{k}{N}+\frac{1}{2N}} \rho^*(x)dx \right| \right\} \tag{3.24}$$

which returns the maximum distance of the empirical distribution π_k from the theoretical distribution. We obtain the theoretical distribution by integrating the theoretical density $\rho^*(x)$ over the bin size $1/N$ around the points k/N. Furthermore, we leave out the points $k = 0$ and $k = N$ in both measures because there the theoretical density $\rho^*(x)$ can diverge.

We find that both distances $\overline{d}$ and $d_{\max}$ decay with increasing N which is illustrated in Figure 3.7 for the example from Figure 3.2. Due to the divergence of $\rho^*(x)$ at the boundaries the maximum distance $d_{\max}$ decays slower than the mean distance $\overline{d}$ showing that the slow convergence at the domain boundaries dominates the quality of the fit. We conclude that the fit for $N \gtrsim 1000$ is already good, but special care has to be taken at the domain boundaries where the divergence of the stationary solution $\rho^*(x)$ may lead to larger deviations.

It is often assumed that the Fokker-Planck approximation holds only for weak selection [65, 88]. Thus in the following we use the two above defined distance measures $\overline{d}$ and $d_{\max}$ to analyze the quality of the stationary solution in dependence of the selection strength. We start by introducing a scaling factor ξ so that the interaction functions from

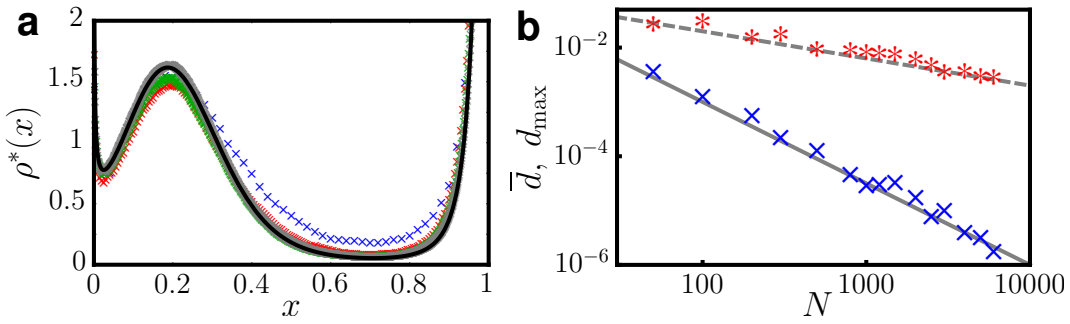

Figure 3.7.: For large population sizes N the theoretically obtained stationary distribution well fits data from simulations. (a) shows the stationary distribution (black, solid) from Figure 3.2 together with empirical densities (crosses) as defined in equation (3.22) from simulations with $N = 50$ (blue), $N = 200$ (red), $N = 500$ (green) and $N = 1000$ (gray). (b) shows that the mean distance $\overline{d}$ and maximum distance $d_{\max}$ between simulation data and the stationary solution decrease with increasing population size N. The measured mean distance $\overline{d}$ (blue, $\times$) as defined in equation (3.23) decreases faster with N than the maximum distance $d_{\max}$ (red, $*$) defined in equation (3.24) because of a slow convergence at the domain boundaries $x = 0$ and $x = 1$. We added the gray lines $N^{-1.5}$ (solid) and $2N^{-0.5}$ as a guide to the eye to show that the distances decrease with N approximately as a power law. For both (a) and (b) we obtained the empirical distributions by simulating the dynamics from an initial state drawn from $\rho^*(x)$ for a mixing time $T_{\text{mix}} = 100N$ and then recording the density for a measurement time $T_{\text{meas}} = 10N^2$.

equation (3.20) become

$$\tilde{g}_A(x) = \xi N(a_A x - b_A x^2) \quad \text{and} \quad \tilde{g}_B(x) = \xi N(a_B(1-x) - b_B(1-x)^2). \tag{3.25}$$

To exclude errors from the finite measurement times T_{meas} we also compare the stationary solution (3.19) of the Fokker-Planck equation (2.40) with the stationary solution (3.8) of the master equation (3.4). The distance between these to distributions is obtained by simply replacing the empirical distribution π_k in the definition of the mean distance measure (3.23) with the stationary solution p_k^* of the master equation. We find that the stationary solution of the Fokker-Planck equation well approximates the solution of the master equation for selection strengths up to ξ of the order 1 which is shown in Figure 3.8. As Figure 3.8d illustrates, only for $\xi < 1$ does the measurement error due to finite simulation times T_{meas} play a role. Figure 3.8c demonstrates that even if the Fokker-Planck approximation causes a larger error for strong selection strengths (here $\xi = 50$), it still catches the overall trend of the dynamics. Thus, for large selection strengths it predicts the trend of the dynamics qualitatively, but makes a quantitative error.

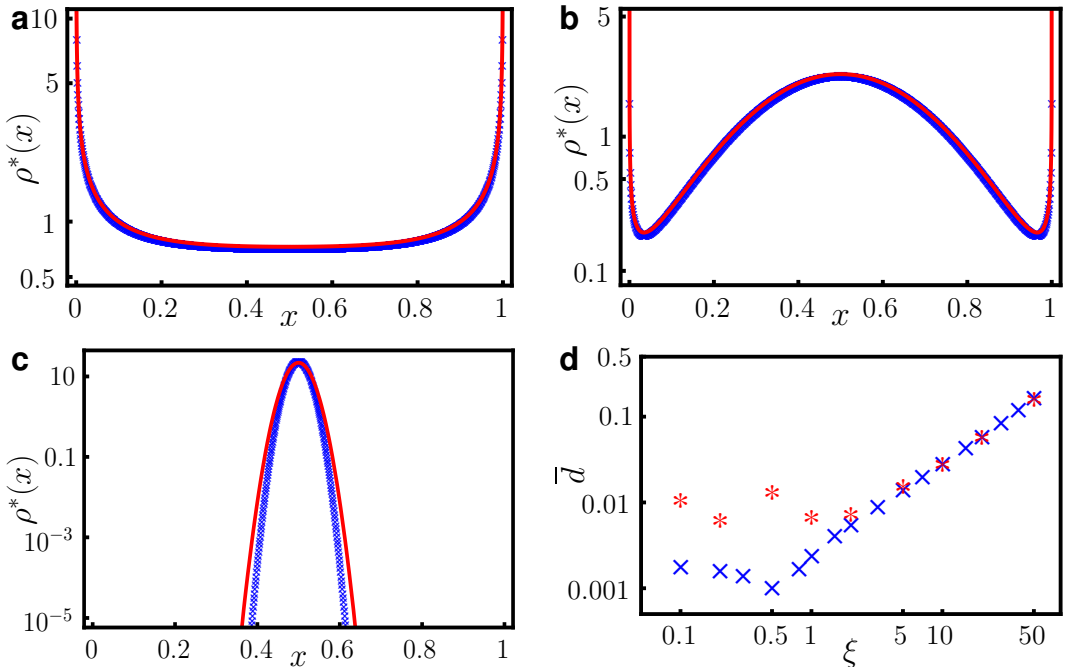

Figure 3.8.: For weak selection the stationary solution (3.19) from the Fokker-Planck equation well approximates the exact solution (3.8) from the master equation, but only qualitatively catches the overall trend of the dynamics for strong selection. (a)-(c) show the stationary solution from the Fokker-Planck equation (red, solid) together with the solution from the master equation (blue, $\times$) for the interaction functions from (3.25) for weak selection $\xi = 0.1$ (a), intermediate selection $\xi = 1$ and strong selection $\xi = 50$. In (c) a large deviation of the Fokker-Planck equation's solution from the master equation's exact solution is observable. (d) quantifies the dependence of the mean distance $\overline{d}$ on the selection strength ξ for measured data (red, $*$) and the solution of the master equation (blue, $\times$). For $\xi > 1$ we observe an increasing distance between the approximate and the exact solution. The system parameters were $N = 1000$, $\tilde{\mu} = 0.5$ and $a_A = a_B = -0.01$ and $b_A = b_B = 0.005$ in equation (3.25). We obtained the empirical distributions by simulating the dynamics from an initial state drawn from $\rho^*(x)$ for a mixing time $T_{\text{mix}} = 10^5$ and then recording the density for a measurement time $T_{\text{meas}} = 10^7$.

3.5. Conclusion

Here, we have analyzed in a simple system of two genotypes how genetic drift, mutations and frequency-dependent fitness shape evolutionary dynamics. In particular, we considered a general class of functions for the frequency-dependent fitness where previous studies had only considered special instances of this class of functions [65, 88]. Also, our general approach included arbitrary mutation rates for both genotypes where previous studies had only considered identical mutation rates [1,88,100]. Our analysis revealed how the interplay of dynamic fitness, mutations and genetic drift induce metastable states on the population dynamics between which the dynamics switch stochastically. At which genotype frequencies the dynamics have a metastable state thereby depends on the respective strength of the involved processes. At low mutation rates genetic drift induces metastable states at the edges of the system where only one genotype is present in the population. These metastable states vanish for higher mutation rates. Furthermore, we found that nonlinear fitness functions may cause many metastable states so that there is no theoretical limit to the number of such states in the system. Additionally, asymmetric mutation rates may shift these states and also cause the emergence of new metastable states.

We found that already at moderate population sizes N a Fokker-Planck equation well approximates the dynamics. However, our analysis based on the distance measures $\overline{d}$ (3.23) and $d_{\max}$ (3.24) revealed that special care has to be taken at the domain boundaries where one genotype dominates in the population. At these boundaries the stationary solution of the Fokker-Planck equation may diverge and thus result in larger deviations. Also, we found that in the strong selection regime where fitness differences are large the Fokker-Planck approach leads to an increasing quantitative error. Thus, for evolutionary systems exhibiting strong selection differences a Wentzel-Kramers-Brillouin (WKB) method may be more appropriate to quantitatively study the evolutionary dynamics. As discussed for example in [3] by Assaf and Mobilia a WKB approach leads to analytical results which well fit simulation data even in the strong selection regime. Still, we have found that the Fokker-Planck approximation qualitatively predicts the trend of the dynamics, so that the application of a WKB approach here would not yield additional insight into how the stochastic evolutionary dynamics are driven by selection and mutation.

We conclude that individuals' interactions that create a nonlinear dependence of the fitness on genotype frequency together with asymmetric mutation rates induce complex evolutionary dynamics [2]. Thus, if in a stable environment repeated shifts between different frequencies are observed, this may indicate individuals' interactions which cause a nonlinear dependence of fitness on the frequency.

4. Dynamic fitness stabilizes populations with variable population size

How will variations in the overall size of a population affect the course of evolutionary dynamics? Experimental studies indicate that dynamically changing population sizes may influence the genotype distribution of a population and thus be important for the evolutionary dynamics [71, 83] which is also suggested by our study of the two-genotype system in Chapter 3 where selectional and mutational effects scaled with the population size. However, the standard replication processes from theoretical models such as the Moran or Wright-Fisher processes keep the population size constant. Therefore, in Section 2.2.2 we introduced a new reproduction process, the IBD process, based on independent birth and death events which yields dynamically changing population sizes. While such birth and death processes are well known in the mathematical theory of Markov chains [38, 62], they are not common in the context of populations evolving on fitness landscapes, although the IBD process allows for studying the impact of dynamically changing population sizes on evolutionary dynamics.

In this chapter we first demonstrate that a population evolving under the IBD process tends to go extinct with high probability after just relatively short times because of random fluctuations in the population size. We then show how dynamic fitness may stabilize the dynamics so that the population will persist for very long times with high probability. Finally, we present a model system where the population size fluctuations of the IBD process together with frequency-dependent fitness induce rich evolutionary dynamics. These dynamics exhibit evolutionary features such as quasi-cycles – cyclic population size dynamics modulated by stochastic background dynamics [6, 53, 70] – and punctuated equilibrium dynamics – self-organized critical dynamics that produce extinction avalanches of power-law distributed sizes [4, 31, 60, 67]. This model is thus a promising approach to gain a better understanding of how individuals' interactions and stochastic reproduction processes cause the emergence of such complex dynamical features.

4.1. The unstable IBD process

Here, we show that the IBD process for populations exhibiting only static fitness is intrinsically unstable, so that a population evolving under this process will either grow infinitely or go extinct. For simplicity of the argument, let us assume that all genotypes have the same fixed fitness f, so that all individuals in the population have equal fitness. As each individual has a birth rate given by the fitness f, the overall birth rate of the population with N individuals is given by $\lambda_N = fN$. On the other hand the death rate of the population is always $\mu_N = N$. Thus, if the individuals' fitness is larger than one ($f > 1$), the birth rate exceeds the death rate for all population sizes and the population will grow on average, if it does not go extinct through stochastic fluctuations from the initial population size N_0. Similarly, for a fitness smaller than one ($f < 1$) the death rate exceeds the birth rate for all population sizes so that the population will almost always go extinct in finite time from any initial population size N_0. We conclude that any population where the individuals have a fitness other than $f = 1$ is unstable under the IBD process.

Even if the fitness is $f = 1$ for all individuals, the population remains unstable. If we consider such a population where all individuals have a fitness $f = 1$, the birth rate $\lambda_N = N$ and the death rate $\mu_N = N$ are equal at all times. Although the rates are equal the actual population size $N(t)$ still fluctuates because birth and death events occur stochastically and independently of each other. The dynamics of $N(t)$ are given by a Markov chain similar to the one described by the master equation (3.4) for the two-genotype system in Chapter 3. We label the states with the actual number of individuals N and thus obtain the transition rates

$$r_N^+ = \lambda_N = N \quad \text{and} \quad r_N^- = \mu_N = N \qquad (4.1)$$

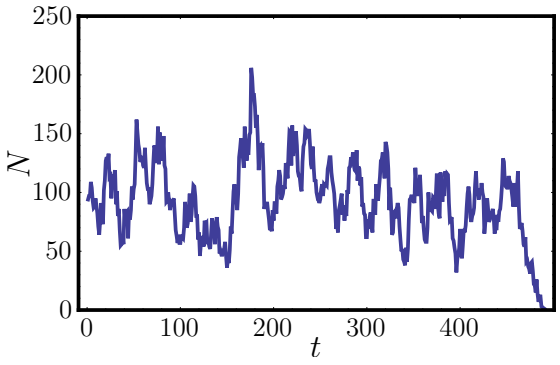

Figure 4.1.: An example dynamics of the population size $N(t)$ for a population with fitness $f = 1$ evolving under the IDB-process. The initial condition was $N_0 = 100$ and the population in this example goes extinct at $t_E = 491$.

to go from state N to the states $N + 1$ and $N - 1$ respectively. Random fluctuations of the population size in this setting can still lead to the extinction of the population, i.e. the dynamics of the Markov chain hits the absorbing state $N = 0$. An example in Figure 4.1 illustrates this, where a population with initially $N_0 = 100$ individuals goes extinct after approximately $t_E \approx 500$ generations.

The dynamics of the underlying Markov chain are recurrent, i.e. the process would visit every state infinitely often if there was no absorbing state. As discussed in Section 2.3.3

this is quantified by the sum

$$\sum_{k=1}^{\infty}\prod_{j=1}^{k}\frac{\mu_j}{\lambda_j}=\sum_{k=1}^{\infty}\prod_{j=1}^{k}\frac{j}{j}=\sum_{k=1}^{\infty}1 \tag{4.2}$$

which here diverges. We conclude that the population will go extinct with probability $p_{\text{Ext}}=1$ as it will hit the absorbing state eventually (cf. Section 2.3.3 and [38]).

According to equation (2.28) in Section 2.3.3, the mean time to absorption from the initial state $N_0=1$ diverges, if the sum

$$\sum_{k=1}^{\infty}\frac{1}{\mu_k}\prod_{j=1}^{k-1}\frac{\lambda_j}{\mu_j} \tag{4.3}$$

diverges [38]. With the transition rates (4.1) this yields the sum

$$\sum_{k=1}^{\infty}\frac{1}{\mu_k}\prod_{j=1}^{k-1}\frac{\lambda_j}{\mu_j}=\sum_{k=1}^{\infty}\frac{1}{k}\prod_{j=1}^{k-1}\frac{j}{j}=\sum_{k=1}^{\infty}\frac{1}{k} \tag{4.4}$$

which diverges. We conclude that the mean time to absorption from the state $N_0=1$ is $\bar{t}_E^{(1)}=\infty$.

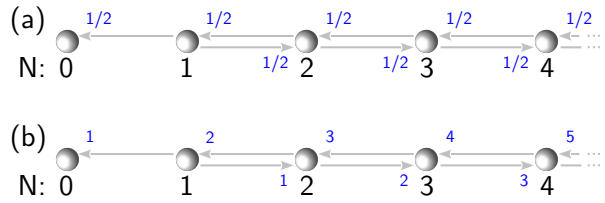

Figure 4.2.: The Markov chain representations of the standard random walk (a) and the IBD process (b) only differ in the speed of the transitions. Black numbers indicate the index of the state indentified with the population size. Blue numbers determine the transition rates between the states. In both the standard random walk in (a) and the IBD process in (b) the probability to move to a higher state equals the probability to go to a lower state.

What is the probability that the population is still alive after a certain time? To answer this question we need to determine the extinction time distribution $p_E(t)$ which determines the probability that the population will go extinct at time t. First, we remark that the process we study is very similar to another well known Markov process: The one-dimensional random walk [29, 62]. In this process similarly to the IBD process only nearest neighbour transitions are possible, i.e. the dynamics are described by a master equation of the form of equation (3.4). The transition rates in this process are $r_k^+=r_k^-=0.5$ for all states k. This Markov process and the IBD process are illustrated in Figure 4.2. The only difference between both processes is that in the one-dimensional random walk transition events always occur at rate one while in our evolutionary system the rate at which a transition occurs is $2N$ in state N. The factor 2 originates from both birth and death events occuring at rate N, yielding an accumulated rate of $2N$.

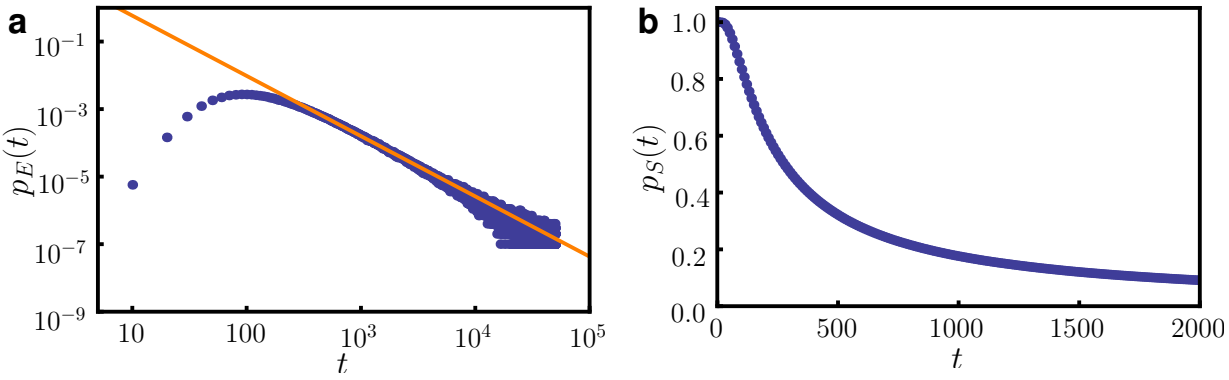

Figure 4.3.: The extinction time distribution $p_E(t)$ of the IDB-process well fits a power law. (a) shows an extinction time distribution $p_E(t)$ (blue, dots) obtained from simulations of the IBD process with initial condition $N_0 = 100$. The process was started at N_0 in 10^6 trials and the extinction time was recorded. The orange line shows a power law fit to the data according to equation (4.5) yielding an exponent $\alpha = 1.784 \pm 0.001$. The deviation from the power law for small times t is determined by the initial condition N_0. (b) shows the survival probability distribution $p_S(t)$ for the measured data from (a) obtained using equation (4.6).

The hitting time distribution of random walks is a research field still under investigation [29, 90, 91]. For the one-dimensional random walk it is well known that the extinction time distribution has a power law shape

$$p_E(t) \propto t^{-\alpha} \tag{4.5}$$

for large times t. At small times t the shape of the distribution $p_E(t)$ is determined by the initial condition N_0 of the process. The tail of the distribution follows the power law (4.5) with exponent $\alpha = 3/2$ independently of the initial condition [90, 91]. As the evolutionary IBD process is very similar to the one-dimensional random walk we expect that it will also exhibit a power law extinction time distribution. Only the exponent α should be different than in the one-dimensional random walk because the process is faster for higher population sizes. We could not compute the exponent α for this process analytically, but determine it by measuring the extinction time distribution in simulations and fitting a power law distribution to it. To achieve this we first cut of the part of the measured distribution that is dominated by the initial condition. We then applied a nonlinear least squares fit to the remaining data and in this way find that the exponent is approximately $\alpha \approx 1.8$. Figure 4.3a shows a measured extinction time distribution together with the fit of the power law (4.5) to the data.

From the extinction time distribution $p_E(t)$ we obtain the survival time distribution

$$p_S(t) = 1 - \int_0^t p_E(t')dt' \tag{4.6}$$

determining the probability that the population is still alive at time t. The second term

here gives the probability that the population has gone extinct until time t, while the initial condition is $p_S(0) = 1$. Figure 4.3b illustrates the survival time distribution for the unstable IBD process. Note that as $\alpha < 2$ for the unstable IBD process the mean time $\bar{t}_E$ to extinction does not exist as $\bar{t}_E = \int_0^\infty t \cdot p_E(t)dt = \infty$ which confirms the result obtained in equation (4.4). Yet, the probability that the population is alive drops quickly with time, so that the population is extinct with probability 0.5 after only approximately 280 generations in the example illustrated in Figure 4.3 starting from a population size $N_0 = 100$. However, in reality populations are usually much more persistent, so that we conclude that there must be a mechanism stabilizing the population dynamics. In the next part we show that dynamic fitness can be such a mechanism as it is able to stabilize the population dynamics of the IBD process.

4.2. The stabilized IBD process

As we discussed in Section 2.2.1, the fitness of individuals is often dynamic. This does not only include frequency-dependent selection which we studied in Chapter 3 but also other effects, such as changing environments. Here, we show how a fitness changing with the overall population size stabilizes the dynamics of the IBD process. We consider a population living in an environment providing only enough resources to sustain a population of N_* individuals. All individuals of the population compete for these resources so that the fitness of each individual is decreased by every other individual through competition for the resources. Such considerations are common in evolutionary models since Verhulst first introduced a carrying capacity in the mid 1800s [56]. If the population is much smaller than N_* then there are resources in abundance and the fitness of the individuals is high. We thus propose a fitness dependence

$$f(N) = \frac{N_*}{N} \tag{4.7}$$

so that the population on average grows if it is smaller than N_* and shrinks whenever it is larger than N_*. As Figure 4.4 illustrates for $N_* = 100$, the population dynamics stay close to N_* and the population is thus less prone to go extinct. Consequently, we call the process defined here the stabilized IBD process.

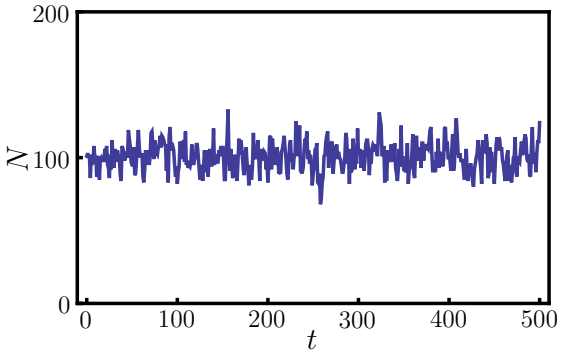

Figure 4.4.: The population size $N(t)$ of a population evolving under the stabilized IDB-process with fitness function (4.7) mainly fluctuates around the sustained population size N_* which in this example is $N_* = 100$.

In the following we analyze the properties of the stabilized IBD process in comparison to

the unstable IBD process. With the fitness of the process (4.7) we obtain the birth and death rates

$$\lambda_N = f(N)N = N_* \text{ and } \mu_N = N. \tag{4.8}$$

With these rates we find that the probability of the stabilized IBD process going extinct is the same $p_E = 1$ as for the unstable IBD process, because the sum

$$\sum_{k=1}^{\infty} \prod_{j=1}^{k} \frac{\mu_j}{\lambda_j} = \sum_{k=1}^{\infty} \prod_{j=1}^{k} \frac{j}{N_*} = \sum_{k=1}^{\infty} \frac{k!}{N_*^k} \tag{4.9}$$

does not converge (cf. Section 2.3.3). However, for this process the mean time to absorption from $N_0 = N_*$ is (cf. equation (2.28))

$$\begin{aligned}
\overline{T} &= \sum_{k=1}^{\infty} \frac{1}{\mu_k} \prod_{j=1}^{k-1} \frac{\lambda_j}{\mu_j} + \sum_{k=1}^{N_*-1} \left(\prod_{j=1}^{k} \frac{\mu_j}{\lambda_j} \right) \sum_{m=k+1}^{\infty} \frac{1}{\mu_m} \prod_{n=1}^{m-1} \frac{\lambda_n}{\mu_n} \\
&= \sum_{k=1}^{\infty} \frac{N_*^{k-1}}{k!} + \sum_{k=1}^{N_*-1} \frac{k!}{N_*^k} \sum_{m=k+1}^{\infty} \frac{N_*^{m-1}}{m!} \\
&= \frac{e^{N_*} - 1}{N_*} + \sum_{k=1}^{N_*-1} \frac{k!}{N_*^k} \frac{e^{N_*}}{N_*} \left(1 - \frac{\Gamma(k, N_*)}{k!} \right) \\
&= \frac{e^{N_*} - 1}{N_*} + (e^{N_*} - 1)\mathcal{O}\left(N_*^{-2}\right)
\end{aligned} \tag{4.10}$$

where $\Gamma(y, k) = \int_k^{\infty} x^{y-1} e^{-x} dx$ is the incomplete Gamma function. Here, in the third step we evaluated the sum term by term using MATHEMATICA. The sum's first term $k = 1$ yields $(e^{N_*} - 1)/N_*^2$ as $\Gamma(1, N_*) = e^{-N_*}$. The sum's higher order terms are of the order $\mathcal{O}(N_*^{-3})$; they may be neglected under the assumption $N_* \gg 1$. This mean time to absorption is finite while for the unstable IBD process the mean time to absorption diverges (cf. equation (4.4)). This seems surprising at first glance as the dynamic fitness pushes the dynamics of the stabilized IBD process away from the absorbing state while there is no such mechanism in the unstable IBD process. The explanation for this finding is the fact, that the unstable IBD process can reach infinitely large population sizes, while the stabilized IBD process cannot because for large enough population sizes N the birth rate is much smaller than the death rate. Thus, the dynamics will always stay close to the stable state N_* until reaching the absorbing state through a random fluctuation.

What is the survival time distribution for the stabilized IBD process? The shape of the distribution is determined by the fact that the birth rates are larger than the death rates for population sizes $N < N_*$ and vice versa for $N > N_*$. Thus, the dynamics stay near the metastable point $k = N_*$ for a long time and only in rare events reach the absorbing state $k = 0$. In such a scenario Kramers' method as discussed in Section 2.3.4 applies [33, 44].

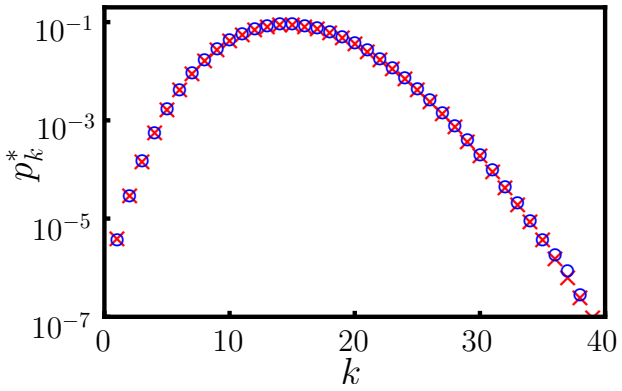

Figure 4.5.: The quasistationary distribution of the stabilized IBD process is very small close to the absorbing state. Shown are the theoretical distribution (red, ×) and a distribution (blue, ○) obtained from simulating the system dynamics for a time $t_{\mathrm{Meas}} = 10^8$ from the initial condition $N_0 = N_*$. The sustainable population size was set to $N_* = 15$.

As we saw there, the survival time distribution $p_S(t)$ fulfills the law

$$\dot{p}_S(t) = -p_S(t) \cdot \left[\sum_{j=1}^{N_*} \frac{1 - \sum_{i=1}^{j-1} p_i^*}{\mu_j p_j^*} \right]^{-1} . \tag{4.11}$$

Here, p_k^* is the quasistationary probability distribution of the metastable state fulfilling the detailed balance equation

$$\lambda_k p_k^* = \mu_{k+1} p_{k+1}^* . \tag{4.12}$$

With $\mu_k = k$ and $\lambda_k = N_*$ the quasistationary distribution p_k^* is thus given by (see [29, p. 266] and the derivation of equation (3.8) in Chapter 3)

$$p_k^* = \frac{\prod_{j=1}^{k-1} \frac{\lambda_j}{\mu_{j+1}}}{\sum_{l=1}^{\infty} \prod_{j=1}^{l-1} \frac{\lambda_j}{\mu_{j+1}}} = \frac{N_*^{k-1}}{k!} \cdot \frac{N_*}{e^{N_*} - 1} = \frac{N_*^k}{k!} \cdot \frac{1}{e^{N_*} - 1} . \tag{4.13}$$

for $k \geq 1$. As the dynamics are absorbed in the state $k = 0$ the quasistationary distribution is not defined there. Therefore, in equation (4.13) all sums and products start with the index 1 instead of 0 which was the case in equation (3.8). The distribution is defined on $k \in \mathbb{N}$, so that the normalization sum ranges from $l = 1$ to infinity. This quasistationary distribution has a high maximum close to $k = N_*$ and already for moderate N_* the probability to come close to the absorbing state is very small, so that the conditions for Kramers' method are clearly fulfilled for $N_* \gg 1$. Figure 4.5 illustrates this for $N_* = 15$.

Using the condition $N_* \gg 1$ we evaluate the time scale in equation (4.11) and obtain

$$\tau = \sum_{j=1}^{N_*} \frac{1 - \sum_{i=1}^{j-1} p_i^*}{\mu_j p_j^*} \approx \left(e^{N_*} - 1 \right) \cdot \left(N_*^{-1} + \mathcal{O}\left(N_*^{-2} \right) \right) \tag{4.14}$$

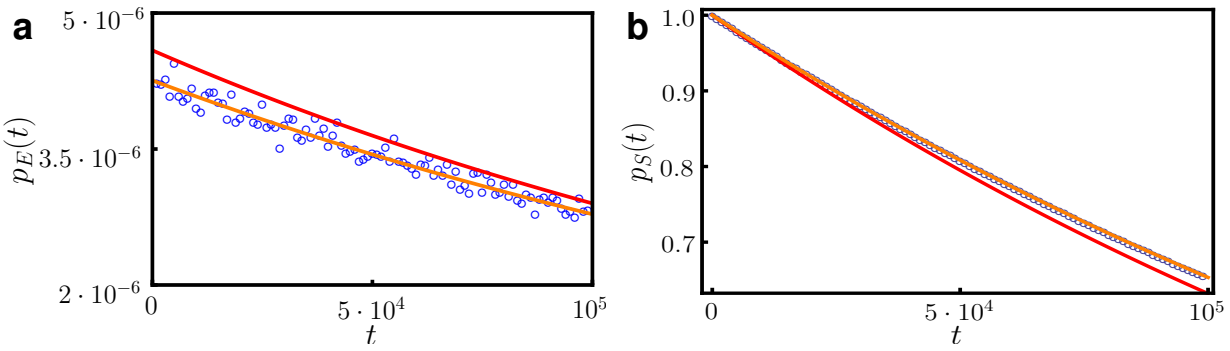

Figure 4.6.: The approximations made for the derivation of the survival time distribution (4.15) and extinction time distribution (4.16) of the stabilized IBD process work well already for relatively small N_*. Both (a) and (b) compare the theoretical predictions (red and orange, solid lines) with data (blue, circles) obtained from recording the extinction times from $5 \cdot 10^5$ trials of simulating the dynamics for $N_* = 15$. The red lines are given by using equation (4.14), while the orange lines were obtained by numerically evaluating equation (4.11). Thus, the main error in our calculations are not determined by Kramers' approximation, but by the following neglecting of terms of the order $\mathcal{O}(N_*^{-2})$. (a) shows the resulting extinction time distribution $p_E(t)$ and (b) the survival time distribution $p_S(t) = 1 - \int_0^t p_E(t')dt'$.

which we derive in detail in Appendix C. We thus find that the survival time distribution in the limit of large sustained population sizes ($N_* \gg 1$) is

$$p_S(t) = \exp\left(-\frac{N_*}{e^{N_*} - 1}t\right). \tag{4.15}$$

The extinction time distribution is easily calculated to

$$p_E(t) = \frac{N_*}{e^{N_*} - 1} \exp\left(-\frac{N_*}{e^{N_*} - 1}t\right) \tag{4.16}$$

using $p_E(t) = -\dot{p}_S(t)$. This extinction time distribution yields the same mean time to extinction $\overline{T} = \int_0^\infty t \cdot p_E(t)dt = \left(e^{N_*} - 1\right)/N_*$ as the result in equation (4.10).

The approximations made to derive the survival time and extinction time distributions work well even for relatively small N_*. This is demonstrated in Figure 4.6 where the theory only slightly underestimates the survival probability of the stabilized IBD process for $N_* = 15$. Actually, the approximations made to derive Kramers' formula (4.11) work very well which we checked by evaluating the factor in equation (4.11) numerically for given parameters (see orange lines in Figure 4.6). We conclude that the main error in our calculation (4.14) is determined by the fact that we neglect terms of the order $\mathcal{O}(N_*^{-2})$.

4.3. A scalable model

In the above section we used a dynamic fitness function (4.7) which implies large differences in fitness. The fitness of an individual can become of the order of the sustained population size N_*. Usually, fitness differences are assumed to be small [88] as large fitness differences often result in the fast extinction of all genotypes but the fittest (cf. Chapter 3), i.e. all fitnesses are usually assumed to be of the order of 1. Kimura's neutral theory of molecular evolution even states that many genotypes have equal fitness, so that many mutations do not affect fitness at all [41]. We conclude that the above fitness function (4.7) is probably not applicable to real systems. Therefore, in this section we introduce a fitness function which allows for keeping the fitness differences small and still stabilizes the IBD process. To this end, we introduce a scaling factor a in the fitness function

$$f(N) = 1 + a\left[\frac{N_*}{N} - 1\right] \tag{4.17}$$

where now with $a = 0$ we obtain the unstable IBD process and with $a = 1$ the stabilized IBD process. Thus, a population evolving under this fitness function (4.17) exhibits more stable dynamics than the unstable IBD process, but also more variable dynamics than the stabilized IBD process, depending on a. Figure 4.7 illustrates this for two different values of a. We call this new process the scaled IBD process.

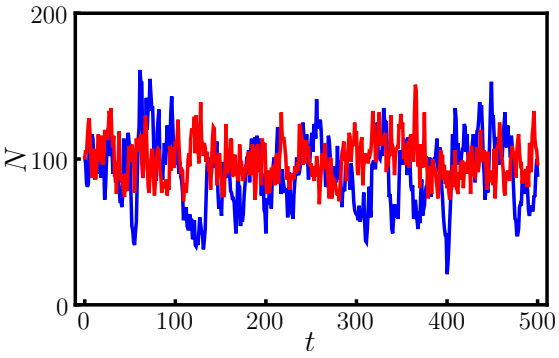

Figure 4.7.: A population evolving under the scaled IDB-process with fitness function (4.17) exhibits fluctuations for which the size depends on the scaling parameter a: For larger a the dynamics stay close to N_* while the fluctuation size increases with decreasing a. Shown are two example trajectories for $N_* = 100$ and $a = 0.1$ (blue) and $a = 0.5$ (red).

If the force driving the system towards N_* is strong enough to stabilize the system, we have a similar situation as in Section 4.2 and Kramers' method applies here, too. This means, that the survival probability of the scaled IBD process should decline as an exponential function with a timescale τ according to equation (2.39). However, the fitness function (4.17) yields the birth rates $\lambda_j = j(1-a) + aN_*$ and death rates $\mu_j = j$ which make a calculation of this time scale τ (2.39) too complicated to obtain a closed expression for τ. Thus, the exact time scale may only be obtained numerically for given parameters a and N_*. Yet, we require an analytical estimate of the time scale to enable a prediction of the necessary parameter value a to let the scaled IBD process survive for a certain time with high probability.

To obtain such an estimate of the time scale we note that if Kramers' method applies the

survival time distribution will be given by

$$p_S(t) = \exp\left(-\frac{t}{\overline{T}}\right) \tag{4.18}$$

where $\overline{T}$ is the mean time survival time which is equal to the mean time to extinction defined by equation (2.28) [38]. With this formula we derive an estimate value for $\overline{T}$ in Appendix D. This estimate

$$\overline{T} = \frac{a^{-\frac{aN_*}{1-a}} - 1}{aN_*} \tag{4.19}$$

is derived under the assumption $aN_* \gg 1$ where we remark that all correction terms for equation (4.19) are positive (cf. equation (D.12) in Appendix D). There are two important facts about this estimate to note. First, in the limit $a \to 1$ the above mean time to extinction (4.19) is equal to the mean time (4.10) derived for the stabilized IBD process in Section 4.2. Secondly, because all correction terms for equation (4.19) are positive, increasing the value of $\overline{T}$, this result always underestimates the mean time to extinction. Thus, equation (4.19) is a strict lower bound for the mean time to extinction. This is illustrated in Figure 4.8 where the theoretically predicted extinction time distribution

$$p_E(t) = \frac{aN_*}{a^{-\frac{aN_*}{1-a}} - 1} \exp\left(\frac{aN_*}{a^{-\frac{aN_*}{1-a}} - 1} \cdot t\right) \tag{4.20}$$

and the survival time distribution

$$p_S(t) = \exp\left(\frac{aN_*}{a^{-\frac{aN_*}{1-a}} - 1} \cdot t\right) \tag{4.21}$$

decline faster than the measured distributions. Figure 4.8 also illustrates that the prediction well approximates the data for increasing aN_*.

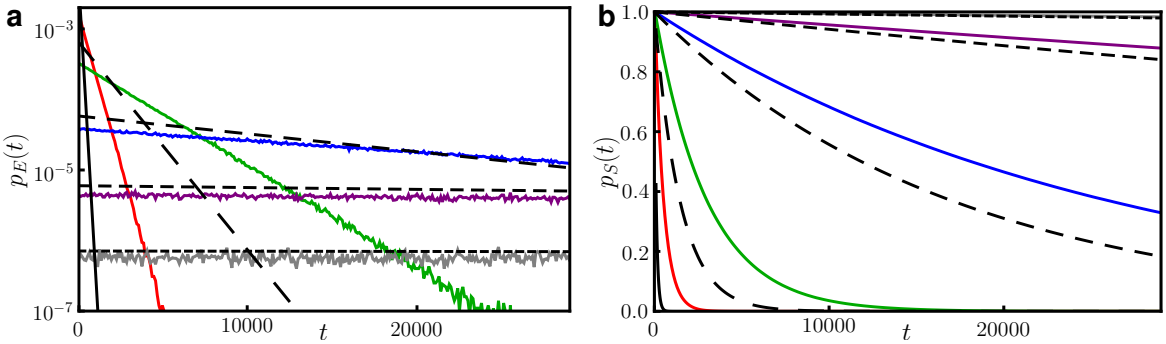

Figure 4.8.: The theoretically computed survival time distribution (4.21) underestimates the actual survival time distribution of the scaled IBD process. Shown are extinction time distribution (a) and survival time distribution (b) of the scaled IBD process obtained from simulations (color) and the corresponding theoretical lower bounds (black). We obtained the data by recording the absorption times for 10^6 trials starting at the initial condition $N_0 = N_* = 100$. The scaling factor was set to $a = 0.01$ (red data, solid line), $a = 0.02$ (green data, long dashes), $a = 0.03$ (blue data, normal dashes), $a = 0.04$ (violet data, short dashes) and $a = 0.05$ (gray data, very short dashes). Already for $aN_* = 5$ the approximations made in the derivation of equation (4.19) work well with relatively small deviations between simulation data and theory. (b) also illustrates that the theory always underestimates but never overestimates the survival probability.

4.4. Comparison of the IBD processes

Comparing the survival time distributions of the stabilized (4.15) and the scaled IBD process (4.21) with the survival time distribution of the unstable IBD process (4.6) we find that the unstable process goes extinct much faster than the other processes on short time scales, which confirms that both the stabilized and the scaled IBD process exhibit more stable dynamics on relevant time scales. However, at extremely long time scales the unstable IBD process has a higher survival probability than the other processes due to the power law distributed extinction times. This is illustrated in Figure 4.9 also demonstrating that the survival time distribution of the unstable IBD process has a fundamentally different shape than the survival time distributions of the other processes. The unstable IBD process can make excursions to arbitrarily large population sizes and these highly improbable excursions can result in a very long survival time of the process. This also causes the divergence of the mean extinction time of the unstable IBD process. On the other hand the other processes will always stay close to N_* with only few stochastic escapes and so they always stay relatively close to the absorbing state. However, as Figure 4.9 demonstrates, the survival probability of the unstable IBD process only surpasses the survival probability of the other processes in a regime where the survival probability of the process has already reached marginal values. We conclude that the population size dependend fitness introduced in equations (4.7) and (4.17) stabilize the population dynamics so that these fitness

functions almost always increase a population's survival probability.

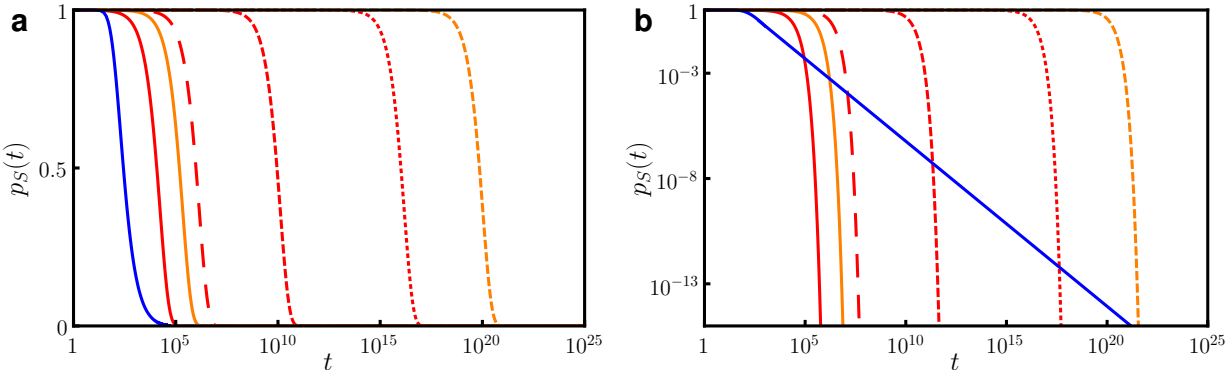

Figure 4.9.: The unstable IBD process has a low survival probability at normal time scales but a higher survival probability than the other processes for extremely long time scales. Shown are the survival time distributions of the unstable IBD process (blue), the stabilized IBD process (orange) and the scaled IBD process (red) in a log-linear plot (a) and a log-log-plot (b). The survival time distribution of the unstable IBD process was obtained by using the data in Figure 4.3 and extending it using the fit in Figure 4.3a. The theoretical survival time distributions (4.16) for the stabilized IBD process are shown in orange for $N_* = 15$ (solid) as in Figure 4.6 and $N_* = 50$ (dashed). The theoretical survival time distributions (4.21) for the scaled IBD process are shown in red for $N_* = 100$ and $a = 0.03$ (solid), $a = 0.05$ (long dashes), $a = 0.1$ (normal dashes) and $a = 0.2$ (short dashes).

The stabilized IBD process survives for extremely long times even at moderate sustained population sizes N_* (cf. Figure 4.9). For example, for $N_* = 100$ the survival probability only goes below $p_S(t) < 0.5$ for times larger than $t = 10^{42}$ (not shown in Figure 4.9 as it is indistinguishable from 1 in the entire plot). Thus, if we require a model of an independent birth death process which will not die out for very long times, the stabilized IBD process is a good choice. However, the fluctuations of the population size in this model remain very small, as the population almost always remains close to N_*. Therefore, if larger fluctuations of the population size are important for the system under study, one should rather turn to the scaled IBD process.

We only analyzed the stability properties of the IBD process under the condition that all genotypes of a population exhibit identical fitness. If we need to study populations with different fitnesses such as in Chapter 3, we therefore have to take special care whether the basic conditions for the theory presented here are fulfilled. The frequency- and population-size dependent fitness function of genotype i becomes

$$f_i(\underline{x}, N) = f_i(\underline{x}) \cdot f(N) \tag{4.22}$$

where $f_i(\underline{x})$ represents frequency-dependent fitness of genotype i (cf. equation (3.1)) and $f(N)$ represents fitness depending on population size (cf. equation (4.17)). Thus, for

the population to be stable there has to be a population size below which $f(N) > (\min_i \{f_i(\underline{x})\})^{-1}$ for all possible configurations $\underline{x}$ so that $f_i(\underline{x}, N) > 1$ at small population sizes even if the population is in a genotype configuration $\underline{x}$ of low fitness $f_i(\underline{x})$. Otherwise the population will go extinct with high probability upon reaching such a genotype configuration of low fitness. Further, for the population to not grow infinitely there must be a population size above which $f(N) < (\max_i \{f_i(\underline{x})\})^{-1}$ for all configurations $\underline{x}$ so that $f_i(\underline{x}, N) < 1$ even if the entire population is of the fittest genotype.

For the stabilized IBD process the fitness depending on population size is of the form $f(N) = N_*/N$. Thus, this fitness is on the order of N_* for small population sizes and can become arbitrarily small for large population sizes. If the genotype-dependent fitness $f_i(\underline{x})$ takes values on the order of one, a population evolving under the stabilized IBD process is therefore stable with the fitness function defined in equation (4.22). For the scaled IBD process the population-size-dependent fitness is given by $f(N) = 1 + a\,[N_*/N - 1]$, so that the maximal possible fitness is given by $1 + aN_* - a$ at population size $N = 1$. For large population sizes the fitness converges to the value $1 - a$. Thus, for a population to evolve under the scaled IBD process in a stable way, the smallest genotype-dependent fitness value has to fulfill $f_i^{\min}(\underline{x}) > 1/(1 + aN_* - a)$ and the largest value has to fulfill $f_i^{\max}(\underline{x}) < 1/(1 - a)$. If these conditions are fulfilled, a population evolving under the scaled IBD process is stable, i.e. the selectional force drives the population towards a stable population size that depends on the population's genotype configuration $\underline{x}$. Still, we remark that the population can go extinct through stochastic fluctuations as described previously.

4.5. Application: A predator-prey model

Populations of predators and their prey exhibit cyclic population dynamics where at some times prey is abundant while predators are rare and vice versa at other times. This striking phenomenon has been studied for many years [5] and models of such systems already exist since the 1920s when e.g. Volterra designed a first model to conceive how the predator-prey interactions shape the population dynamics [93]. He described the population dynamics using coupled, deterministic differential equations for the predator and prey frequencies in the population. However, only under the assumption of special kinds of interactions do these models predict cyclic behaviour; only recently was it realized that the deterministic dynamics may not be oscillatory at all and the cyclic dynamics may rather be caused by a resonant amplification of the stochasticity underlying the dynamics [53]. Thus, stochastic models are required to wholly grasp all effects occuring in predator-prey dynamics [6]. An individual-based model with dynamic fitness may thus be a promising tool to better understand how individuals' interactions cause the emergence of cyclic dynamics in predator-prey system.

Another important feature of evolutionary dynamics is that evolution does not proceed

gradually but rather in sudden steps [31, 72]. This effect has been called punctuated equilibrium dynamics because the dynamics stay in an equilibrium for long times until the equilibrium is "punctuated" through stochastic fluctuations and the dynamics switch to another equilibrium on a much shorter time scale. It has been proposed that this effect is due to the evolutionary system being in a self-organized critical state [4]. To study this effect models have been developed showing that species may go extinct in avalanches of power-law distributed size [4, 60, 67]. However, these models are more abstract than the individual-based models we use in this thesis and it is therefore not clear how the interactions of individuals influence the distribution of the extinction events. Here, we present a simple model based on a reproduction process with independent birth and death events. The interactions of predator and prey individuals are reflected by dynamic fitness functions. We show that this model both exhibits cyclic behaviour as well as punctuated equilibrium dynamics, which indicates that both effects may be induced by the co-action of dynamic fitness and stochastic reproduction.

In the following we describe an individual-based model which combines the stabilized dynamics of IBD processes with the general interactions between individuals of different genotypes (cf. Chapter 3). In this model we consider M different genotypes each representing another species in a food web [21,95]. An individual of a certain genotype may thus be the prey of individuals of another genotype positioned higher in the food web, and at the same time prey on individuals of genotypes further down in the food web. In our model this is reflected by the fitness functions of the different genotypes. We consider interactions between the different individuals in such a way that the fitness of an individual is reduced if there are many predators hunting it, and its fitness is increased if prey is abundand in the population. A food web has a certain depth d, which is the number of levels from the basic prey species up to the species on the top level of the food web that is not hunted by any other species. In the following, for clarity of the argument we describe how to model a food web of depth $d = 3$. However, the model is easily generalizable to any depth d.

For a food web of depth $d = 3$ we label the different species by the three levels A, B and C in the food web. Individuals of type A are feeding on some limited resources, e.g. plants, and are the prey of type B individuals which are in turn the prey of type C individuals. The species of one type are positioned on one level of the underlying food web, but differ in the species they hunt and in the species they are hunted by (cf. Figure 4.10).

We model the interactions between the different individuals using dynamic fitnesses as in Chapter 3: First, we assume that populations of type A individuals will rapidly grow as long as enough recources are there, and we assume that each species is specialized on one food source, but also using some of the other type A individuals' food sources. Taken together, this results in a fitness function

$$f_i^A(\underline{k}) = 1 + a - \alpha k_i - \sum_{j \neq i, j \in A} \beta k_j \tag{4.23}$$

for population i of type A. Here, $\underline{k} = (k_1, k_2, \ldots, k_M)$ is the vector of all population

sizes, $a > 0$ models an increased fitness due to abundant resources,$\alpha > 0$ models the strong competition for resources of the individuals of species i and $\beta > 0$ models the weak competition with the other type A individuals. If a and α are large enough, the population is stabilized according to the considerations in Section 4.3. Each of the type B species are specialized on certain types of prey, i.e. they hunt only some of the type A species while they ignore others. We assume that they have some species which they mainly prey on and some which are also hunted with a less strong impact. The predation decreases the type A populations' fitness so that the overall fitness function becomes

$$f_i^A(\underline{k}) = 1 + a - \alpha k_i - \sum_{j \neq i, j \in A} \beta k_j - \sum_{j \in B} \gamma_{ij} k_j \tag{4.24}$$

where $\gamma_{ij} = \gamma_S$ if there is strong predation of species i from j, $\gamma_{ij} = \gamma_W$ if predation is weak, and $\gamma_{ij} = 0$ if species j does not prey on species i. While the fitness of the type A individuals decreases with predation, the fitness of type B individuals increases so that we obtain

$$f_i^B(\underline{k}) = 1 - b + \sum_{j \in A} \delta_{ij} k_j \tag{4.25}$$

with $\delta_{ij} = \delta_S$, $\delta_{ij} = \delta_W$ or $\delta_{ij} = 0$ as above. Here, $b > 0$ models the species' reduced fitness if no prey is to be found. Similarly, individuals of type C strongly or weakly prey on type B species, so that the type B fitness function becomes

$$f_i^B(\underline{k}) = 1 - b + \sum_{j \in A} \delta_{ij} k_j - \sum_{j \in C} \varepsilon_{ij} k_j \tag{4.26}$$

and type C fitness

$$f_i^C(\underline{k}) = 1 - c + \sum_{j \in A} \zeta_{ij} k_j \tag{4.27}$$

with $\epsilon_{ij} \in \{\epsilon_S, \epsilon_W, 0\}$ and $\zeta_{ij} \in \{\zeta_S, \zeta_W, 0\}$ as above. Figure 4.10 illustrates the resulting food web structure for an example system.

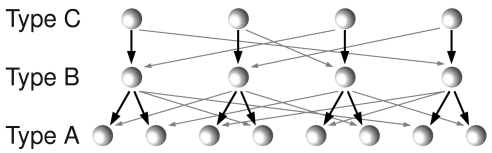

Figure 4.10.: The interaction structure of the predator-prey model. Strong black arrows indicate strong predation and thin gray arrows weak predation.

We consider mutations between the different species in the following way. New mutants may turn to new food sources without changing their position in the level of the food web. We represent this with the mutation probabilities $\mu_{ij} = \mu_{AA}$ for an individual of a type A species i to mutate to another type A species j. The same applies for type B and type C with mutation probabilities $\mu_{ij} = \mu_{BB}$ and $\mu_{ij} = \mu_{CC}$. With a lower probability new mutants may also turn to a new role in the food web, i.e. they may be positioned on a new level. We represent this with the lower mutation probabilities $\mu_{ij} = \mu_{AB}$ for individuals of type A species i to mutate to a type B species j and similarly with $\mu_{ij} = \mu_{BC}$ for individuals of type B species

i to mutate to a type C species j.

In the following, we study an example system with eight different species of type A, and four of type B and C respectively. The overall food web of this model is illustrated in Figure 4.10. The dynamics of this system exhibits two features that are proposed to occur in real evolutionary systems [5, 31]. An example dynamics is shown in Figure 4.11, illustrating the points discussed in the following. The features of the dynamics are:

1. Quasi-cycles. As Figure 4.11c illustrates the population sizes of predators and prey fluctuate cyclically, however not perfectly periodically. Rather, the fluctuations are influenced by the stochastic fluctuations from the random events of the IBD process. Periodic fluctuations are well known for predator-prey systems [5, 93], however in these systems the fluctuations are perfectly periodic and deterministic. Already in the 1970s it was suggested that the stochastic nature of the underlying dynamics in predator-prey systems may be the cause of periodic dynamics [61], but only recently was this analyzed in detail [53, 70]. Thus, it was revealed that normally stable states in predator-prey systems become unstable under stochastic dynamics, so that fluctuations arise, which are not perfectly periodic, but still cyclic. That means, that the microscopic stochastic dynamics excite macroscopic oscillations which are phase-forgetting. Therefore, this dynamical feature is called a quasi-cycle [6, 53, 70]. As the dynamics in Figure 4.11c and its power spectrum in Figure 4.11d demonstrate, our system clearly exhibits these quasi-cycles (compare also with power spectrum in Figure 2 in [53]). Figure 4.11d also shows, that the amplitude of the fluctuations is highest for the species which are on top of the food chain. Even more, the influence of the stochastic dynamics on the fluctuations seems to give rise to a resonance-like effect (see dynamics between $t = 500$ and $t = 1000$ in Figure 4.11c) where the fluctuations become very large. This effect might be caused by extinction events which make the dynamics more unstable than in the system studied in [53] by McKane and Newman. Thus, the introduced system is an example system for the important influence of the stochastic nature of the dynamics on predator-prey dynamics.

2. Punctuated equilibrium dynamics. Panels a and b of Figure 4.11 show that the dynamics often stay close to a metastable state for long times before sudden switches to new metastable states occur. We had observed similar behaviour already in Chapter 3. Here, this behaviour often includes the extinction of species which is visible in Figure 4.11e, where the number n of alive species is plotted versus time. Actually, if one species goes extinct through stochastic fluctuations, the metastable state the dynamics were in can become unstable so that the dynamics will move to a new metastable state. Thus, the extinction of one species often has a catastrophic impact on the evolution of other species. For example, if a species of type A vanishes, the food for a type B species becomes scarce. Or if a type C species vanishes, a type B species can grow strongly which in turn decreases the fitness of type A species. Thus, one extinction event is often followed by more extinction events a few generations later in an avalanche-like fashion. We measured the size of such avalanches in the

following way. We define that an avalanche was initialized at a time t if the number of alive species $n(t)$ has decreased in the last generation: $n(t) < n(t-1)$. After such an initialization of an avalanche we checked if further species had vanished from the system after a time $t + 50$. This waiting time is large enough to allow the effect of the vanishing species spread through the system so that other species can go extinct, but still small enough so that the probability for another unrelated extinction event to occur remains small. If $n(t + 50) < n(t)$, we then checked whether this lead to further extinctions after a further time $t + 100$. We repeated this procedure iteratively until we found a time $t + i \cdot 50$, $i \in \mathbb{N}$ at which no further species had vanished from the system, i.e. $n(t + (i-1) \cdot 50) = n(t + i \cdot 50)$ we obtained the avalache size as $s = n(t) - n(t + i \cdot 50)$. Figure 4.11f shows the thus obtained distribution of avalanche sizes $m(s)$ in a log-log-plot. Also shown is a line, suggesting that the avalanche sizes decrease like a power law. This is an indicator that the dynamics are in a self-organized critical state, i.e. extinction events affecting many species can be caused by the stochastic reproduction dynamics alone. Therefore, our results suggest that our system exhibits punctuated equilibrium dynamics, which is an important feature of evolutionary dynamics as it determines that evolution does not proceed gradually but rather in sudden steps [31]. There are already evolutionary models which exhibit punctuated equilibrium dynamics [4, 60, 67], however these models are abstract and study the emergence of the phenomenon on the population level. Thus, these models cannot determine how individuals' reproduction processes and the interactions between the individuals influence the phenomenon. With the results presented here, our model appears promising to study the influence of reproduction processes and individuals' interactions on punctuated equilibrium dynamics.

The interesting dynamics of this example system demonstrate that the IBD process with dynamic fitness seems to be a promising model to study the impact of stochastic reproduction processes and interactions on evolutionary dynamics. To our knowledge quasi-cycles and punctuated equilibrium dynamics were until now only observed in different models, but not in one unifying model. Thus, our model shows that both effects can be induced by the co-action of dynamic fitness and stochasticity alone. Additionally, our model may clarify how interactions and stochastic reproduction cause these effects. Also, it could show whether they mutually influence each other; e.g. a higher amplitude of the quasi-cycles could cause populations to become more unstable as they have a small population size in the valley of the cycle, so that this could induce more avalanches of large size. Finally, we remark that the stochastic switching of the dynamics in Panels a and b of Figure 4.11 together with the distribution of extinction events in Figure 4.11f suggests that our model system exhibits punctuated equilibrium. However, further studies in larger systems are necessary to confirm this. In our system the avalanche sizes s could only be observed on a scale of the order $\mathcal{O}(10)$ which, in a model of 16 species, is already a large avalanche. Still, to gain better statistics with larger avalanche sizes it is thus necessary to study systems with more species.

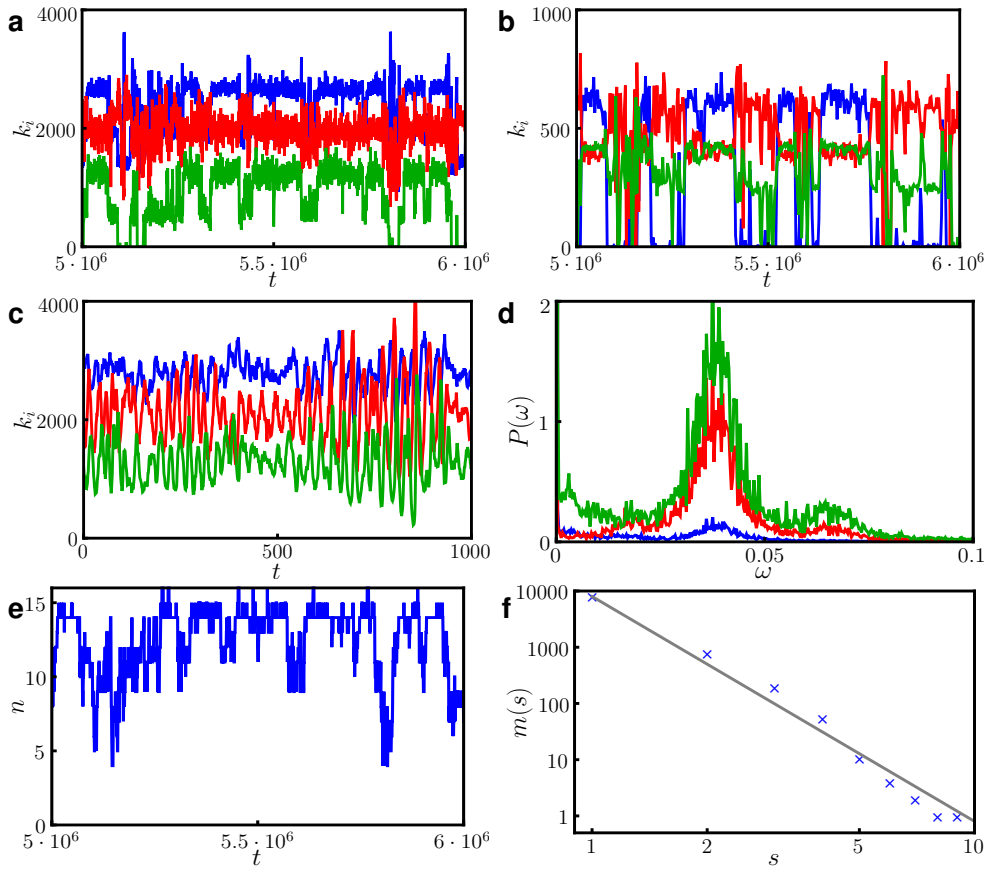

Figure 4.11.: The predator-prey model system exhibits rich dynamics. (a) shows the long term dynamics of the population sizes k_A (blue), k_B (red) and k_C (green) of type A, B and C individuals respectively. Each data point is averaged over a time $t = 500$ so that the cyclic fluctuations shown in (c) are averaged out. Shown is a time interval of length 10^6 of a simulation of the dynamics lasting 10^7 time units. Long periods of quasistationarity are followed by short bursts of rapid changes which is also illustrated by (b) showing the population sizes of three single species of type A (blue), type B (red) and type C (green), also averaged over a time $t = 500$. (c) shows the same population size dynamics k_A, k_B and k_C as (a) on a smaller time scale. The dynamics exhibit cyclic fluctuations influenced by stochastic noise which can also give rise to resonance effects as seen between $t = 500$ and $t = 1000$. (d) shows a plot of the power spectrum $P(\omega)$ of the three time series $k_A(t)$, $k_B(t)$ and $k_C(t)$ which confirms the impression given by (c) that they exhibit cyclic fluctuations. The intensity of the fluctuations is highest for the species which are on the highest level in the food web. (e) shows the number n of alive species for the dynamics shown in (a). n often decreases in avalanches of size s. (f) shows the number $m(s)$ of avalanches of size s that occured in the example dynamics of (a) in a time $t = 10^7$ after fadeaway of initial conditions, i.e. the avalanche sizes were only recorded after a waiting time $t = 10^6$. The gray curve $m(s) = 8000 \cdot s^{-4}$ is added as a guide to the eye. The interaction parameters for this system were $a = 0.2$, $b = 0.15$, $c = 0.3$, $\alpha = 2.5 \cdot 10^{-4}$, $\beta = 10^{-5}$, $\gamma_S = 1.5 \cdot 10^{-4}$, $\gamma_W = 5 \cdot 10^{-5}$, $\delta_S = \varepsilon_S = 3 \cdot 10^{-4}$, $\delta_W = \varepsilon_W = 1.5 \cdot 10^{-4}$, $\zeta_S = 5 \cdot 10^{-4}$ and $\zeta_W = 2 \cdot 10^{-4}$. The mutation probabilities were $\mu_{AA} = \mu_{BB} = \mu_{CC} = 2 \cdot 10^{-6}$ and $\mu_{AB} = \mu_{BC} = 2 \cdot 10^{-7}$.

4.6. Conclusion

In this chapter we have studied populations evolving under a reproduction process with independent birth and death events. In this setting, the overall population size fluctuates stochastically so that a population can go extinct. We showed that the population dynamics are unstable under this reproduction process if the individuals exhibit fixed fitnesses: If the population's fitness is larger than one the population will on average grow without limit, if it is smaller than one it will go extinct. If the fitness is exactly $f = 1$ we found that the population exhibits an extinction time distribution $p_E(t)$ – the probability that the population will go extinct at time t – following a power law. Thus, the the population will go extinct with high probability after relatively short times and only persist for long times with low probability. However, assuming a carrying capacity N_* for the population we obtained a dynamic population-size-dependent fitness yielding a qualitatively different extinction time distribution; in such a setting the distribution follows an exponential decay. Our results show that a population with dynamic fitness modelling a carrying capacity will persist with high probability for much longer times than the population with fixed fitness. We conclude that in models studying populations with fluctuating population sizes dynamic fitness may be an important mechanism stabilizing the population.

In the following we presented an example system modelling predator-prey interactions based on the previously studied independent birth and death process. We thus illustrated how the here presented framework may be used to model ecological systems. The resulting system exhibits rich dynamical features such as quasi-cycles [6, 53] or punctuated equilibrium dynamics [31] which was to our knowledge not yet studied in individual-based models, but only in more abstract models. Also, we do not know of any previously defined model system where both of these evolutionary features were observable. As the underlying model is based only on dynamic fitness and a simple reproduction process, an analytic study of how the individuals' interactions cause these effects seems feasible. We therefore suggest that this model may be a good approach to gain a better understanding under which conditions quasi-cycles or punctuated equilibrium dynamics emerge.

5. Horizontal gene transfer in changing fitness landscapes

Recent studies suggest that horizontal gene transfer (HGT) – the exchange of genetic material between individuals of different species – may have played an important role in early evolution and still contributes to todays ongoing evolution [14, 45, 46]. However, the influence of HGT in evolutionary dynamics is still not well understood. There are rather many assumptions about the role of HGT in evolution, but only few theoretical studies on how HGT affects evolutionary dynamics. One of these assumptions is that HGT might help populations to adapt to changing environments [82]. Recently, Raz and Tannenbaum showed in a simple model that in any static environment HGT has a deleterious effect on a population at mutation-selection balance [73]. This result was confirmed by Vogan and Higgs [92] in simulations using an agent-based model. As HGT still plays a role in the bacterial evolution today [37, 46, 57], they proposed that HGT may confer a fitness advantage in changing fitness landscapes. However, to our knowledge it has still not been shown explicitly that HGT yields a fitness advantage for populations in changing fitness landscapes.

Therefore, in this chapter we study HGT in changing environments using an individual-based model. Here, using the individual-based model introduced in Sections 2.2.2 and 2.2.6 with a slowly fluctuating fitness landscape we explicitly show, that HGT can give a population a fitness advantage in changing environments where the optimal rate at which HGT is most advantageous for the population depends on the speed at which the fitness landscape changes. Thus, depending on the frequency of environmental change there is an optimal competence for HGT exhibited by the individuals of the population which maximizes their fitness.

5.1. Model setup

Consider a population evolving on a changing fitness landscape defined by a genome of length l as described in Section 2.2.1, so that genotype space assumes a hypercube structure with mutation probabilities μ_{ij}. Let us for now assume that all mutation probabilities are equal $\mu_{ij} = \mu$. Later we will also consider distributed mutation probabilities. As we have seen in Chapter 3, individuals interactions can imply rapid changes in a population's fitness

which we exclude here to focus on the effect of HGT on an adapting population. Therefore, in the following we will not include interactions in our model, but rather apply a fitness landscape driven by external forces. This means that the fitness $f_i(t)$ of genotype i depends explicitly on time, but not on the composition of the population. We consider a Fujiyama landscape [20] – i.e. the fitness increases gradually towards one single fitness peak – with periodically shifting peak defined by the periodically fluctuating fitness function

$$f_i(t) = 1 + A\sin(\omega t)\cos\left(\pi\frac{\mathrm{HD}(i,0)}{l}\right) \tag{5.1}$$

for genotype i. Here A is the amplitude and ω the frequency of the fluctation and $\mathrm{HD}(i,0)$ indicates the Hamming distance between genotype i and genotype 0. Thus, depending on time only, half of the time genotype 0 exhibits the highest fitness and the other half of the time genotype $2^l - 1$ has the highest fitness. Such a periodic fluctuation may for example model seasonal effects.

We further assume that the individuals are open to exchange genetic material which we represent in the model by introducing a number m of HGT-links. These HGT-links are inserted randomly in the way described in Section 2.2.6. We assume that each HGT link has the same HGT base rate c.

5.2. Adaptation to changing landscapes

Studying the population dynamics in this system we find (as expected) that the population tries to move to the peak of the Fujiyama landscape. As the peak is shifted periodically the population has to adapt to the changing landscape repeatedly. Thus, the mean fitness $\langle f\rangle$ of the population increases when adapting to the new peak, but decreases when the peak is shifted to a new position. Depending on the frequency with which the landscape changes, sometimes the population is capable of adapting in time to the new peak through repeated mutations and sometimes the fitness decreases more strongly as the population does not reach the newly emerging fitness peak in time. The population has more problems to adapt to the landscape the faster it changes. This is illustrated in Figure 5.1 which shows example trajectories of the population's mean fitness for different frequencies ω with which the fitness fluctuates (cf. equation (5.1)).

Does HGT increase or decrease the average fitness of a population evolving in such a fitness landscape? To answer this question, we now randomly introduce to the system a number m of HGT-links as defined in Section 2.2.6. Assigning different values for the base rate c for different HGT-links leads to qualitatively similar results as using the same value c for all HGT-links. Hence, for simplicity we set the base rate for all links to the same value c. In this way, we study the overall influence of HGT on the fitness of the population by varying the HGT base rate c. In simulations for each value c we let the population evolve for a long

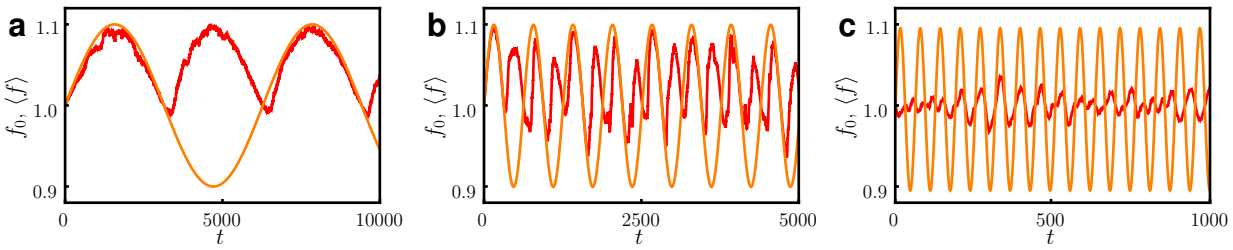

Figure 5.1.: A population adapts best to slowly changing fitness landscapes. The Figures show example trajectories of the mean fitness $\langle f \rangle$ (red) of a population evolving in a changing fitness landscape with the fitness $f_0(t)$ of genotype 0 shown in orange. The frequency of change was $\omega = 0.001$ in (a), $\omega = 0.01$ in (b) and $\omega = 0.1$ in (c). Further parameters were $l = 7$, $N = 1000$, $\mu = 0.001$ and $A = 0.1$ and in this example no HGT occured ($c = 0$).

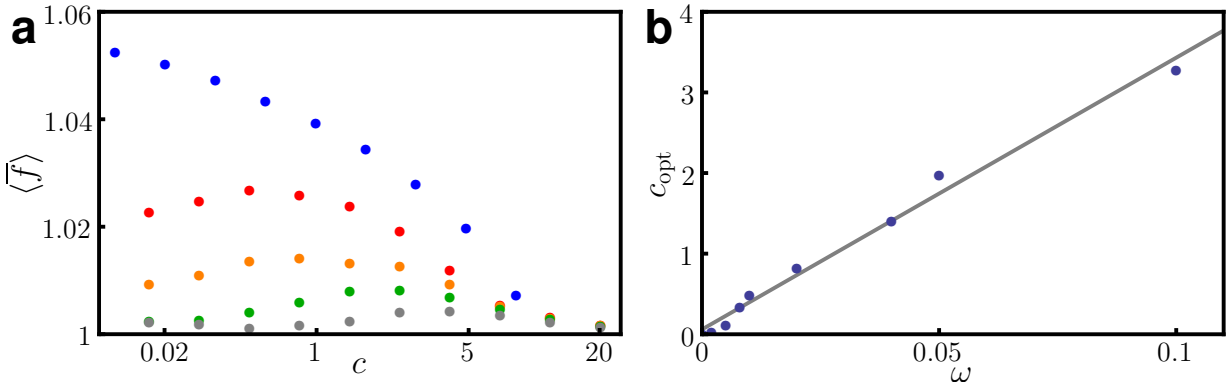

Figure 5.2.: The optimal HGT base rate c_{opt} where HGT is most beneficial for a population depends on the frequency ω with which the fitness landscape changes. (a) shows the population's overall mean fitness $\left\langle \overline{f} \right\rangle$ in dependence of the HGT base rate c for different frequencies $\omega = 0.0001$ (blue), $\omega = 0.01$ (red) $\omega = 0.02$ (orange), $\omega = 0.05$ (green) and $\omega = 0.1$ (gray). HGT improves the population's potential for adaptation in changing environments as for each frequency $\omega > 0$ there is an optimal base rate $c_{\text{opt}} > 0$ which maximizes $\left\langle \overline{f} \right\rangle$. We measured this optimal base rate c_{opt} in dependence of ω which is shown in (b). The data suggests that the optimal base rate increases linearly with the frequency which is illustrated by the least squares fit to the data $c_{\text{opt}}(\omega) = 0.06 \pm 0.049 + (33 \pm 1.3) \cdot x$ (gray line). System parameters were $l = 7$, $m = 1000$,$N = 1000$, $\mu = 0.001$ and $A = 0.1$. The simulation time for each data point in (a) was $T = 10^5$. For all data points in (a) the variance is on the order of 10^{-7} or smaller, so that we did not add error bars to the data points.

time and measure the population's overall mean fitness $\langle \overline{f} \rangle$, which is the time average of the population's mean fitness. We repeat this procedure for different frequencies ω. Over a wide range of frequencies we found in our simulations that for each frequency $\omega > 0$ an optimal HGT base rate $c_{\text{opt}} > 0$ maximizes the population's overall mean fitness. This finding is illustrated in Figure 5.2a for different frequencies ω. Clearly c_{opt} decreases with decreasing ω which is illustrated in Figure 5.2b suggesting a linear relationship between frequency ω and optimal HGT base rate c_{opt}. We observe that in the limit $\omega \to 0$ of slowly changing landscapes also the optimal HGT base rate becomes neglectable ($c_{\text{opt}} \to 0$). This confirms the findings in [73, 92] that HGT does not increase a population's fitness in fixed environments. We remark, that for the fitness function (5.1) the limit $\omega = 0$ does not exist as the (fixed) fitness value f_i depends on the initial condition. Thus, the special case $\omega = 0$ is qualitatively different from the fluctuating fitness landscape we studied here and we therefore cannot determine c_{opt} for $\omega = 0$ directly. Figure 5.2a also illustrates that with increasing frequency ω the maximal overall mean fitness $\langle \overline{f} \rangle_{\text{opt}}$ decreases, because it becomes more and more difficult for the population to adapt to the fast changes in the landscape. Yet, the fitness gain due to HGT is clearly visible. All in all, we have now shown explicitly that HGT can confer a fitness advantage in changing environments.

5.3. Conditions for the beneficial effect of HGT

Is there also an optimal mutation probability μ_{opt} (at a given HGT base rate c) for the adaptation to the changing fitness landscape and how does the optimal HGT base rate c_{opt} depend on this mutation probability? In the previous simulations we kept the mutation probability μ fixed, so that we found an optimal HGT base c_{opt} for this special value μ only. Now, we need to clarify how mutations and HGT act together to help a population move in a changing landscape, i.e. we search for the parameter setting $(\mu_{\text{opt}}, c_{\text{opt}})$ maximizing the population's overall mean fitness $\langle \overline{f} \rangle$ at a given frequency ω. To this end, we varied both the mutation probability μ and the HGT base rate c for a given frequency ω and measured the overall mean fitness $\langle \overline{f} \rangle$ as above. The simulation results indicate that there is an optimal mutation probability μ_{opt} maximizing the fitness of the population which is illustrated in Figure 5.3 for two different frequencies ω. Furthermore, the results shown in Figure 5.3 suggest that for this mutation probability μ_{opt} HGT does not increase the fitness of the population; for a given frequency ω the mutation probability μ_{opt} seems to determine the best possible speed of adaptation for the population which cannot be increased by HGT, i.e. $c_{\text{opt}} = 0$. For all parameter sets with $\mu > \mu_{\text{opt}}$ HGT was deleterious. However, close to the optimal mutation probability HGT can still be beneficial, as for example Figure 5.3b shows where the overall mean fitness for the parameter set $(\mu = 0.001, c = 1)$ comes very close to the optimal overall mean fitness at the parameter set $(\mu = 0.005, c = 0)$. In this setting, also for the optimal mutation probability μ_{opt} the deleterious effect of HGT seems to set in only for $c > 1$.

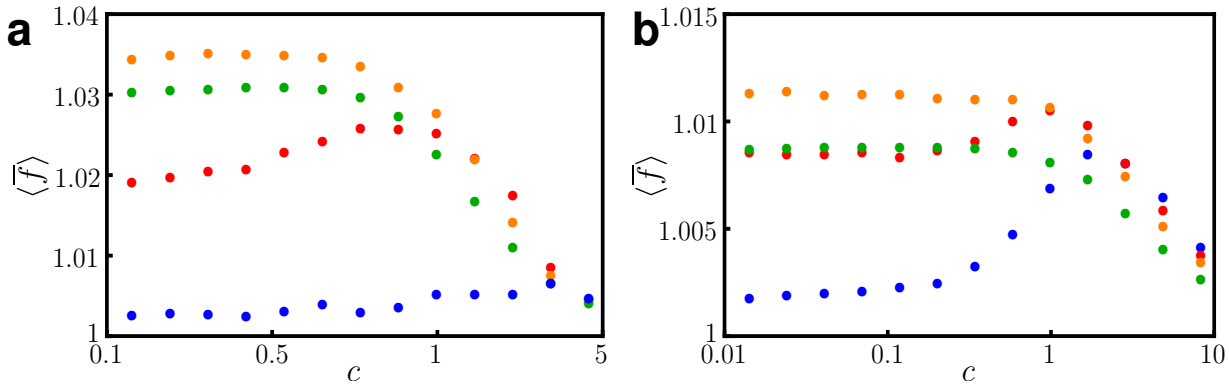

Figure 5.3.: In fitness landscapes changing with one frequency ω there is an optimal mutation probability μ_{opt} maximizing the overall fitness of the population. Both panels (a) and (b) show the measured overall mean fitness $\left\langle \overline{f} \right\rangle$ in dependence of the HGT base rate c for different mutation probabilities μ and frequencies $\omega = 0.01$ (a) and $\omega = 0.05$ (b). Mutation probabilities were $\mu = 0.0001$ (blue), $\mu = 0.001$ (red) and $\mu = 0.005$ (orange) and $\mu = 0.01$ (green) in (a) and $\mu = 0.001$ (blue), $\mu = 0.01$ (red), $\mu = 0.02$ (orange) and $\mu = 0.05$ (green) in (b). Further system parameters were $l = 7$, $m = 1000$, $N = 1000$, and $A = 0.1$. Each data point was obtained by simulating the dynamics for a time $T = 10^5$. The variance for all data points is on the order of 10^{-7} or smaller.

The dynamic change of the fitness landscape should – to be more realistic – include more complex features than just a periodic switching with only one frequency ω. Is there a dynamically changing landscape with optimal mutation probability μ_{opt} for which HGT gives an additional advantage ($c_{\text{opt}} > 0$)? To answer this question we studied fitness landscapes changing in different ways. For example, we used landscapes changing with multiple frequencies, but also landscapes changing with sudden randomly occuring jumps mimicking punctuated equilibrium dynamics [31], or landscapes fluctuating randomly. Yet, the results of the simulations were qualitatively the same as above. In all of these landscapes we found that a parameter set ($\mu_{\text{opt}}, c_{\text{opt}} = 0$) maximizes the population's overall mean fitness, i.e. HGT does not yield a fitness advantage if the population exhibits the optimal mutation probability. Only, for $\mu < \mu_{\text{opt}}$ does HGT increase the population's overall mean fitness. As we could not check all possible dynamic fitness landscapes we cannot exclude, that a landscape exists, where HGT also confers an advantage at μ_{opt}, but the collection of all our simulation results indicate that this is not the case.

In real biological systems the mutation probabilities μ_{ij} between different genotypes are usually diverse, i.e. there is not one uniform value $\mu_{ij} = \mu$ [22, 77]. How does such mutational diversity influence the positive impact of mutations and HGT in changing environments? We considered a fitness landscape as defined above with the only difference, that the mutation probabilities were chosen randomly from a uniform distribution $\mu_{ij} = 2\mu\xi_{ij}$ with the random number ξ_{ij} uniformly drawn from the interval $[0, 1]$ and the mean mutation probability μ. In this fitness landscape we varied the parameters for the mean

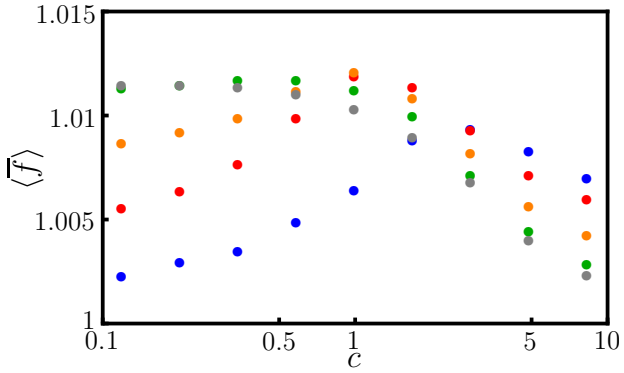

Figure 5.4.: In fitness landscapes with non-uniform mutation probabilities μ_{ij} there is an optimal mean mutation probability $\mu_{\text{opt}} > 0$ and optimal HGT base rate $c_{\text{opt}} > 0$ maximizing the overall fitness of the population, if the landscape is changing with one frequency ω. The Panel shows the measured overall mean fitness $\left\langle \overline{f} \right\rangle$ in dependence of the HGT base rate c for different mean mutation probabilities μ where the mutation probabilities were distributed uniformly with $\mu_{ij} = 2\mu\xi_{ij}$ for $\xi_{ij} \in [0, 1]$. The fitness fluctuated with a frequency $\omega = 0.05$. The mean mutation probabilities here were $\mu = 0.001$ (blue), $\mu = 0.005$ (red), $\mu = 0.01$ (orange), $\mu = 0.02$ (green) and $\mu = 0.03$ (gray). The highest overall fitness value was measured for $\mu_{\text{opt}} = 0.01$ and $c_{\text{opt}} = 1$. Each data point was obtained by simulating the dynamics for a time $T = 10^5$. Further system parameters were $l = 7$, $m = 1000$, $N = 1000$, and $A = 0.1$.

mutation probability μ and the HGT base rate c. For each of these values we again measured the overall mean fitness $\left\langle \overline{f} \right\rangle$ of the population. In this setting, we found that HGT actually gives a slight advantage to the population even for the optimal mean mutation probability μ_{opt}. An example is shown in Figure 5.3. The results suggest that for low mean mutation probabilities μ HGT is highly beneficial for the population, while close to the optimal mutation probability $\mu_{\text{opt}} = 0.01$, the increase in fitness due to HGT is small but noticeable. Thus, our results suggest that if mutation rates between different genotypes are diverse, a population evolving in changing environments may obtain optimal fitness if it exhibits HGT.

5.4. Conclusion

All in all, we explicitly showed that the assumption in [73, 92] that HGT can improve the fitness of a population in changing environments is true. However, we found that close to an optimal mean mutation probability μ_{opt}, HGT gives only a slight advantage for non-uniform mutation probabilities. For uniform mutation probabilities $\mu_{ij} = \mu$ HGT is only beneficial if the mutation probabilities are too small. Our results thus suggest that HGT may play a beneficial role for populations adapting to changing environments by providing a mechanism to cross regions in genotype space where mutation probabilities are small. The beneficial impact of HGT on the population observed in our simulations was relatively

small.

We speculate that dynamically changing competences may make HGT more advantageous for adapting populations. It has been shown that competence can depend on the well-being of an individual so that in our model the parameter c would depend on the fitness of an individual [47, 48]. In such a setting the deleterious effects of HGT for a well-adapted population would be reduced while still yielding the advantage of fast adaptation when the landscape changes. To check if such a mechanism may increase the positive impact of HGT will be a goal of future work on evolutionary dynamics in changing fitness landscapes.

6. Evolutionary dynamics with frequent horizontal gene transfer

The emergence of the first species is still a puzzle. In the last decade it has been proposed e.g. by Woese [96,97] that before the emergence of distinct species there was a reactive soup where HGT dominated the evolutionary dynamics. In this reactive soup no distinct species exist, but rather every life form has its own set of genes which are frequently exchanged with other individuals. Many studies agree on HGT probably having played a prominent role in early evolution [14,18,45,96,97]. Yet, the properties of such HGT-dominated dynamics are not known and it is completely unclear how distinct species could emerge from the reactive soup [14,18,97]. In this chapter we analyze how frequently occurring HGT affects evolutionary dynamics and how it may cause a reactive soup state.

The few existing previous theoretical studies considering HGT mainly focussed on how HGT influences the fitness of a population [73,92] (cf. Chapter 5) or the impact of HGT in the quasispecies model [7,36,68]. With the quasispecies model the influence of mutations in a population may be analyzed and thus the inclusion of HGT in this model revealed how HGT may change the impact of mutations in population dynamics. The studies [7,36,68] revealed that HGT influences the mutational error threshold – a critical mutation rate above which the population completely spreads out in genotype space (cf. Section 2.2.5) – in the quasispecies model and can give rise to a bistability in the system. However, these results do not reveal the general impact of HGT on evolutionary dynamics which remains unknown.

Here, we use the HGT model introduced in Section 2.2.6 and study the impact of HGT on the population dynamics in a static fitness landscape. We concentrate on the effect of changing the HGT base rate c on the evolutionary dynamics. The individuals' inclination to exchange genetic material scales with c so that a large value c models the frequently occuring HGT at early evolution. By varying the HGT base rate c we find a critical transition similar to the error threshold in quasispecies theory [23,24], where a new stable state emerges which is dominated by HGT instead of selection dynamics. The system becomes bistable, so that a selection-dominated state is stable as well and the dynamics can stochastically switch between these two states. Our analysis reveals that HGT alone suffices to cause the emergence of this bistability. This stochastic switching between the HGT-dominated and the selection-dominated state and the vanishing of the HGT-dominated state by lowering the competence may help explain the transition from the reactive soup to distinct species.

6.1. Model setup and the introduction of an entropy variable

Consider a population evolving on a static Fujuyama fitness landscape [20] where fitness increases gradually in genotype space towards one single fitness peak. For the genotype space we assume a hypercube defined by a genome of length l as described in Section 2.2.1. The fitness function of genotype i is defined by

$$f_i(t) = 1 + A\cos\left(\pi\frac{\mathrm{HD}(i,0)}{l}\right) \tag{6.1}$$

throughout this chapter, where A defines the height of the fitness peak in the Fujiyama landscape and $\mathrm{HD}(i,0)$ is the Hamming distance between genotype i and 0 as in Chapter 5. Thus, genotype 0 is the fittest and genotype $2^l - 1$ is the least fit in this fitness landscape at all times. We assume a uniform mutation probability $\mu_{ij} = \mu$ for all mutation links in the hypercube. We have checked that moderately distributed mutation probabilities do not change the results of this chapter qualitatively.

We further assume that the individuals are open to exchange genetic material which we represent in the model by introducing a number m of HGT-links. These HGT-links are inserted randomly in the way described in Section 2.2.6. As for the mutations we assume initially that all HGT-links have the same HGT base rate c to simplify the analysis. We will study the effect of distributed HGT base rates c later.

How should we vizualize the dynamics of this system? It is not feasible to study the dynamics of all the population sizes k_i of the different genotypes i for larger l as there are 2^l different genotypes in the system. We therefore introduce a new variable $S(t)$ which measures how distributed the population is in the fitness landscape. We define $S(t)$ as an entropy-like variable

$$S(t) = -\sum_{i=0}^{2^l-1}\frac{k_i(t)}{N}\log\left[\frac{k_i(t)}{N}\right] \tag{6.2}$$

and call this the *population entropy*. This population entropy is $S = 0$ if the entire population is concentrated on one single genotype i ($k_i = N$). If the population is equally distributed among all genotypes ($k_i = N/2^l$), the population entropy takes its maximal value

$$S =: S_{\max} = l \cdot \log(2)\,. \tag{6.3}$$

For example, on the Fujiyama landscape defined above for low enough mutation probabilities and HGT base rates the population entropy will stay close to zero as the population is mainly concentrated around the fitness peak at $i = 0$ (cf. also Figure 6.1).

6.2. A transition in evolutionary dynamics

Depending on the individuals' competence for exchanging genetic material, how does HGT influence the population dynamics qualitatively? We studied the time evolution of the population entropy for different HGT base rates c in an otherwise unchanged fitness landscape as defined in Section 6.1. As we focus on the impact of HGT, we fix the mutation probability at a low value $\mu = 0.1N^{-1}$. Thus, mutations still occur in the population and the dynamics cannot converge to an absorbing state, yet the impact of the mutations on the overall evolution of the population remains small (cf. also Chapter 3). The population entropy reveals qualitatively different population dynamics for different HGT base rates c. Figure 6.1 illustrates the different dynamics for one example system. In general we find that for low c the population is mainly concentrated around the fittest genotype as expected. The dynamics of the population entropy remain close to zero, with minor fluctuations due to mutations and the stochasticity of the underlying population dynamics (cf. Figure 6.1a). This *low entropy state* is reached from arbitrary initial conditions after a transition period. Still, in this low entropy state the average fitness of the population may switch between different values (cf. Figure 6.1d) as the population is concentrated on different genotypes at different times. For higher HGT base rates c the dynamics occasionally reach a new state of high population entropy, which we call the *high entropy state* (cf. Figure 6.1b). In this state the population is distributed throughout the entire fitness landscape as the average rate of HGT events is highly increased. We observe that the dynamics switch stochastically between the low and the high entropy state, in each state fluctuating around an equilibrium value for many generations. The switching from one state to the other occurs on a much shorter time scale of only few generations. Furthermore, the average percentage of time spent in the low entropy and high entropy state respectively depends on c. The dynamics stay longer in the high entropy state for higher c (cf. Figure 6.1 and Figures 6.6 and 6.7 for more details). Thus, for high enough c the dynamics remains in the high entropy state for almost all times starting from any initial condition (cf. Figure 6.1c). Finally we note, that the value of the population entropy in the high entropy state is not the maximally possible population entropy $S_{\max}$, i.e. the population is not perfectly distributed among the genotypes in the high entropy state.

What drives the dynamics in the low and high entropy states? The examples in Panels d-f of Figure 6.1 illustrate that in the low entropy state the population's mean fitness $\langle f \rangle$ is relatively high, often even close to optimal as the population concentrates on the genotypes around the fitness peak. In the high entropy state the mean fitness $\langle f \rangle$ is rather close to the average fitness value

$$f_{\text{av}} = \frac{1}{2^l} \sum_{i=1}^{2^l} f_i = 1 \tag{6.4}$$

as the population is completely spread out in the genotype space. Consequently, as Pan-

els g-i of Figure 6.1 illustrate, the population's HGT rate

$$r_{\mathrm{HGT}} = \sum_{i=1}^{m} c \cdot k_A \frac{k_B}{N} \tag{6.5}$$

– where the sum goes over all m HGT-links in the system (cf. equation (2.15) in Section 2.2.6) – is close to zero in the low entropy state so that reproduction events occur much more often than HGT events. We remark that the HGT rate r_{HGT} is a dynamic variable of the system quantifying the rate at which HGT occurs in a certain state of the system while the HGT base rate c is a system parameter quantifying the individuals competence for HGT. In the high entropy state the HGT rate r_{HGT} is substantially higher than in the low entropy state. We observe that it becomes on average of the order of the reproduction rate or larger. For example, in Figure 6.1 the reproduction rate is always approximately $r_{\mathrm{Repr}} \approx 1000$, so that in the high entropy state in Figure 6.1h the average HGT rate approximately equals the reproduction rate and for higher base rates c as in Figure 6.1i it is larger than the reproduction rate. Thus, we conclude that in the low entropy state the dynamics are dominated by the selection process so that the population concentrates around the fitness peak, but in the high entropy state the dynamics are dominated by HGT and selection plays only a minor role.

To understand the transition and the emergence of the bistability of the population dynamics we developed a method – based on the population entropy – to study the forces induced by selection and HGT on the population dynamics. The evolution of the population is event driven such that at each event – be it a reproduction event of the Moran process or the exchange of genetic material in an HGT event – the setup of the population can slightly change, and thus also the population entropy S defined in (6.2). Therefore, at each event time there is a population entropy S_b directly *before* the event and a population entropy S_a directly *after* the event. The resulting change of population entropy

$$\Delta S = S_a - S_b \tag{6.6}$$

will in general depend on the type of event (reproduction or HGT event) and the actual state of the system. On average one of these events will induce a mean change $\overline{\Delta S}(S)$ of the population entropy if the system is in a state with population entropy S. Considering the rate $r(S)$ at which the events occur, for a given state S the population entropy will on average change with

$$\dot{S}(S) = r(S) \cdot \overline{\Delta S}(S) \tag{6.7}$$

through the reproduction and HGT events.

In our model system reproduction and HGT events are completely independent, so that we may study their effects on the population entropy independently of each other. Thus, the average change of the population entropy becomes

$$\dot{S}(S) = \dot{S}_{\mathrm{Repr}}(S) + \dot{S}_{\mathrm{HGT}}(S) = r_{\mathrm{Repr}}(S) \cdot \overline{\Delta S}_{\mathrm{Repr}}(S) + r_{\mathrm{HGT}}(S) \cdot \overline{\Delta S}_{\mathrm{HGT}}(S) \tag{6.8}$$

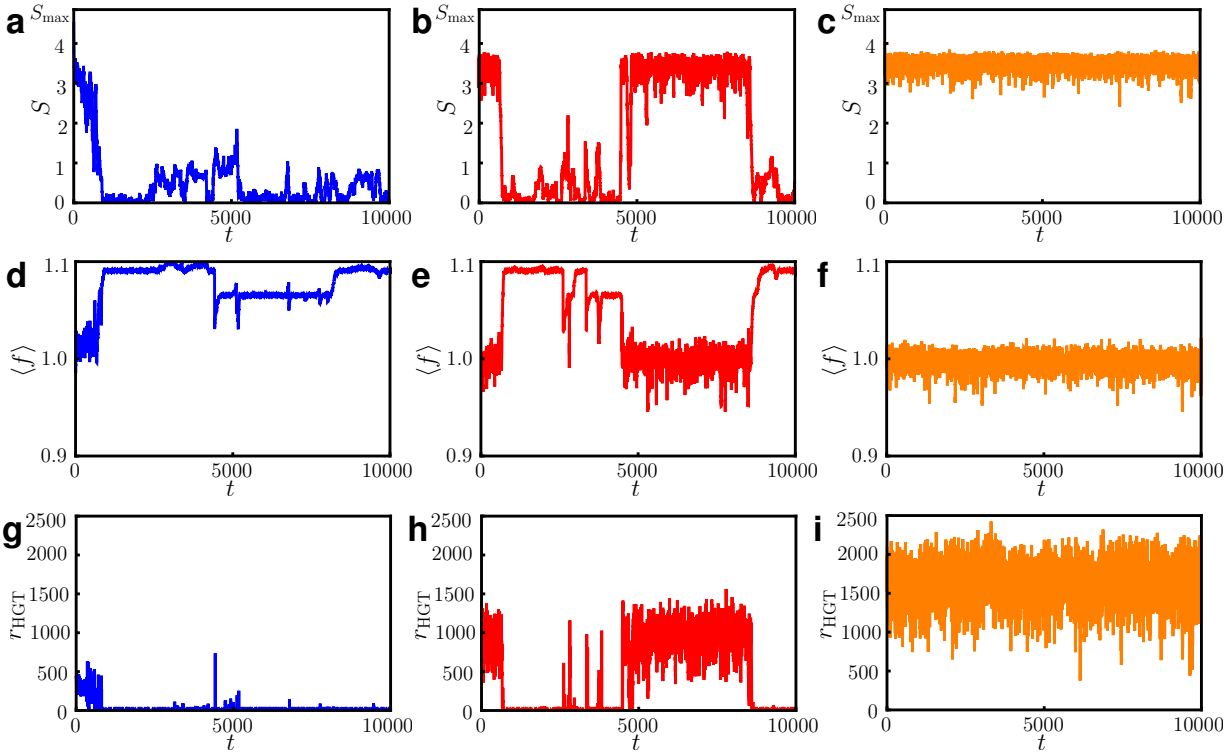

Figure 6.1.: For different HGT base rates c qualitatively different population dynamics emerge. Shown are example dynamics of the population entropy for $c = 1$ (a), $c = 3$ (b) and $c = 5$ (c) in an example system of a population of $N = 1000$ individuals with genome length $l = 7$, $m = 2000$ HGT-links and a fitness peak of height $A = 0.1$. For low HGT base rates c the population entropy fluctuates slightly above zero for all initial conditions (a), for high c the population entropy almost always fluctuates around a high value for all initial conditions (c) and for intermediate HGT base rates the dynamics switch stochastically between these two states (b). The maximal population entropy here is $S_{\max} = 7 \log(2) \approx 4.85$. In the low entropy state the population dynamics are driven by selection, in the high entropy state by HGT. Panels (d), (e) and (f) show the mean fitness $\langle f \rangle$ of the population corresponding to the entropy dynamics in (a), (b) and (c). The corresponding average HGT rates r_{HGT} are illustrated in Panels (g), (h) and (i). For low population entropies the fitness is high and HGT rate small and vice versa for high population entropies.

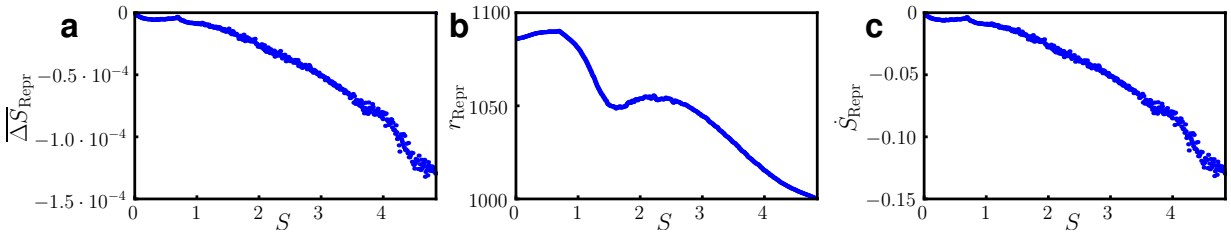

Figure 6.2.: The reproduction events on average decrease the diversity of a population. (a) shows the measured average effect $\overline{\Delta S}_{\mathrm{Repr}}(S)$ of the reproduction events in dependence of the population entropy for the example system studied in Figure 6.1. The genotypes' fitness is in the ranges $f \in [0.9, 1.1]$ so that the reproduction rate can only fluctuate between $r_{\mathrm{Repr}} \in [0.9 \cdot N, 1100 \cdot N]$. Therefore, the reproduction rate only slightly depends on the system state as (b) illustrates, so that the rate of change $\dot{S}_{\mathrm{Repr}}(S)$ due to reproduction shown in (c) has a similar shape as the effect $\overline{\Delta S}_{\mathrm{Repr}}(S)$ in (a). The results in (a)-(c) are almost independent of the HGT base rate in the system, so that only data for $c = 1$ is shown. Each dataset was obtained in simulations measuring the dynamics for a time $T = 10^7$.

with the reproduction rate $r_{\mathrm{Repr}}(S) = N\overline{f}(S)$ determined by the populations average fitness $\overline{f}(S)$, the HGT rate $r_{\mathrm{HGT}}(S)$, and the mean changes of population entropy $\overline{\Delta S}_{\mathrm{Repr}}(S)$ and $\overline{\Delta S}_{\mathrm{HGT}}(S)$ through reproduction and HGT events. In our simulations we measured these rates and the average changes in dependence of the actual population entropy S: At each event we recorded the population entropy before the event, the actual reproduction and HGT rates and the change of population entropy $\Delta S(S)$ caused by the event. Averaging over all recorded events that occurred at population entropy S thus yields the mean changes of population entropy $\overline{\Delta S}_{\mathrm{Repr}}(S)$ and $\overline{\Delta S}_{\mathrm{HGT}}(S)$ both for reproduction and HGT events, and the average reproduction and HGT rates $r_{\mathrm{Repr}}(S)$ and $r_{\mathrm{HGT}}(S)$ respectively.

How does the population entropy change through the reproduction events? Our results show that reproduction events have a mean effect $\overline{\Delta S}_{\mathrm{Repr}}(S) < 0$ and thus induce a rate of change $\dot{S}_{\mathrm{Repr}}(S) < 0$ for all population entropies S, i.e. on average reproduction events decrease the diversity of the population. This is illustrated in an example in Figure 6.2a. This was expected as selection alone should reduce the diversity of the population until all individuals are of the fittest genotype. Furthermore, we find that the effect of selection is stronger the more the population is distributed in genotype space, i.e. $\overline{\Delta S}_{\mathrm{Repr}}(S)$ is a monotonically decreasing function (cf. also Figure 6.2a). The actual shape of this function however depends on the shape of the actual fitness landscape. As we only assume relatively small fitness differences between the different genotypes, the reproduction rate $r_{\mathrm{Repr}} = N \cdot \overline{f}(S)$ is of the same order for all population entropies S (cf. Figure 6.2b). Thus, the resulting rate of change $\dot{S}_{\mathrm{Repr}}(S)$ illustrated in Figure 6.2c has a similar shape as the mean effect $\overline{\Delta S}_{\mathrm{Repr}}(S)$.

The population entropy's change due to HGT is more complex. We find that HGT events have a mean effect $\overline{\Delta S}_{\mathrm{HGT}}(S)$ which depends on the population entropy in a complex way (cf. Figure 6.3a). For intermediate population entropies S we find $\overline{\Delta S}_{\mathrm{HGT}}(S) > 0$,

i.e. HGT on average distributes the population in genotype space. However, close to the minimal population entropy $S = 0$ and the maximal population entropy $S_{\max}$ we observe $\overline{\Delta S}_{\mathrm{HGT}}(S) < 0$. This behaviour may be due to the HGT-link structure: Different genotypes can have different numbers of incoming HGT-links via which HGT brings new individuals to them. Thus, a genotype with many such incoming HGT-links will on average receive more individuals through HGT than another genotype with fewer HGT-links. Therefore, HGT will not distribute individuals perfectly in the system so that at high population entropies – where the individuals are equally distributed in genotype space – HGT will reduce the diversity in the population, i.e. $\overline{\Delta S}_{\mathrm{HGT}}(S) < 0$. At low population entropies the population is concentrated on the few, fittest genotypes which are similar in our system. Hence, a HGT event will often change an individual such that it will mutate to a genotype which is already present in the population. In this way, for low population entropies the HGT events could on average further reduce the population entropy. That the HGT-link structure causes these effects is suggested by our measurements where we found that the regions where $\overline{\Delta S}_{\mathrm{HGT}}(S)$ is negative depend on the HGT-link structure.

The HGT rate increases nonlinearly with increasing population entropy (cf. Figure 6.3b) which is caused by the nonlinear factor $k_A \cdot k_B/N$ in the rate equation (2.15) of the single HGT-links. Considering all HGT-links the HGT rate is small if the population is concentrated on few genotypes – as many HGT-links are inactive because $k_A = 0$ or $k_B = 0$ – and the rate increases nonlinearly with the population spreading in genotype space – as more and more links become activated as $k_A > 0$ and $k_B > 0$. Thus, the resulting rate of change $\dot{S}_{\mathrm{HGT}}(S)$ illustrated in Figure 6.3c increases nonlinearly from $\dot{S}_{\mathrm{HGT}}(0) = 0$ reaching a maximum at intermediate population entropies and becomes negative at high population entropies. The HGT base rate c mainly acts as a scaling factor for the HGT rate r_{HGT} (cf. Figure 6.4b) and thus also for the average rate of change $\dot{S}_{\mathrm{HGT}}$(cf. Figure 6.4c) so that the overall impact of HGT in the evolutionary dynamics is controlled by the HGT base rate c.

How do reproduction and HGT together drive the evolutionary dynamics? Adding up the average rate of change due to HGT and reproduction yields the overall rate of change $\dot{S}(S)$ (cf. equation (6.8)) which is illustrated in Figure 6.4c for different HGT base rates c. We observe that for low HGT base rates the impact of HGT is smaller than the impact of reproduction for all population entropies. Thus, the population is on average drawn towards $S = 0$ explaining the observed dynamics in Figure 6.1a. For higher HGT base rates c the average rate of change $\dot{S}_{\mathrm{HGT}}$ for HGT increases, so that it overcomes the negative rate of change $\dot{S}_{\mathrm{Repr}}$ for high population entropies. At the HGT base rate c_{cr} where this first occurs, a new stable state of the dynamics is created through a saddle-node bifuraction. The emerging fixed point is marked in the example in Figure 6.4c. The HGT base rate at which the saddle-node bifurcation occurs is the critical HGT base rate c_{cr} above which stochastically switching dynamics are observed (cf. Figure 6.1b), i.e. not only short stochastic excursions to high population entropy are observed but also dynamics which stay at high population entropy for longer times. We remark that we here only studied the *average* rate of change $\dot{S}(S)$ of the population entropy. Yet, as the underlying dynamics are

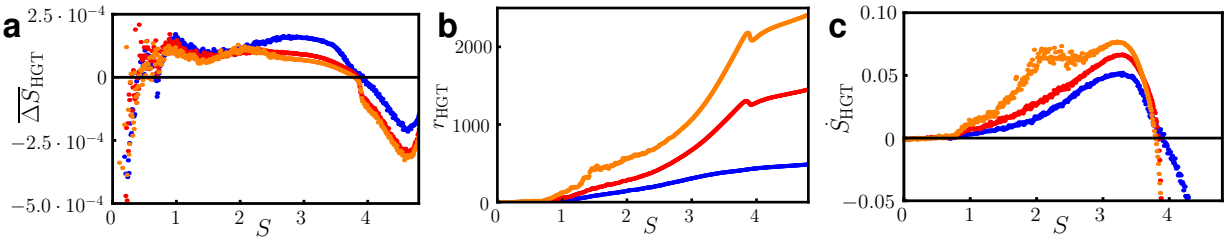

Figure 6.3.: The average rate of change $\overline{\Delta S}_{\mathrm{HGT}}(S)$ due to HGT depends nonlinearly on the population entropy. Similarly to Figure 6.2, (a) shows the measured average effect $\overline{\Delta S}_{\mathrm{HGT}}(S)$ of the HGT events for the example system studied in Figure 6.1. At low and high population entropies the effect $\overline{\Delta S}_{\mathrm{HGT}}(S)$ is negative while it is positive for intermediate population entropies. All Panels (a)-(c) show results for HGT base rates $c = 1$ (blue), $c = 3$ (red) and $c = 5$ (orange). The mean HGT rate $r_{\mathrm{HGT}}(S)$ increases nonlinearly with S (b) so that the rate of change $\dot{S}_{\mathrm{HGT}}(S)$ shown in (c) first increases nonlinearly with S before dropping down at high population entropies. Each dataset was obtained in simulations measuring the dynamics for a time $T = 10^7$.

stochastic, still fluctuations counteracting the observed average rate of change can occur. Therefore, the dynamics in Figure 6.1a do not converge to the state $S = 0$ but rather fluctuate above this value. Also, this explains why above the critical HGT base rate c_{cr} the dynamics stochastically switch between the stable states at high and low population entropy.

Why do the dynamics almost always remain in the high entropy state for high HGT base rates? Using the average rate of change $\dot{S}(S)$ we define a potential

$$V(S) = -\int_0^S \dot{S}(S')dS' \tag{6.9}$$

in which the dynamics move under additional stochastic forcing. This potential is shown in Figure 6.4d. According to reaction rate theory [33], the depths of the two stable states' potential wells determine the average time the dynamics stay close to each of the stable states. If there are two potential wells A and B at points a and b and a barrier at c, then the times $\tau_A \propto e^{E_A}$ and $\tau_B \propto e^{E_B}$ the dynamics remain in the wells A and B scale with the potential differences $E_A = V(a) - V(c)$ and $E_B = V(b) - V(c)$ if these are large compared to the diffusion speed ω ($E_A \gg \omega$ and $E_B \gg \omega$). As Figure 6.4d illustrates, for large HGT base rates c the stable state created in the bifurcation has a much deeper potential well than the state at $S = 0$. Therefore, for large HGT base rates the dynamics stay almost always in this potential well and the observed population dynamics are thus almost always in the high entropy state. From the dynamics shown in Figure 6.1 we have obtained the stationary probability density $\rho(S)$ to find the the population in a state with population entropy S. This probability density is illustrated in Figure 6.5 together with the potential from Figure 6.4d confirming that the dynamics stay the longer in the high entropy state the deeper the potential well of the high entropy state is. This is also illustrated by

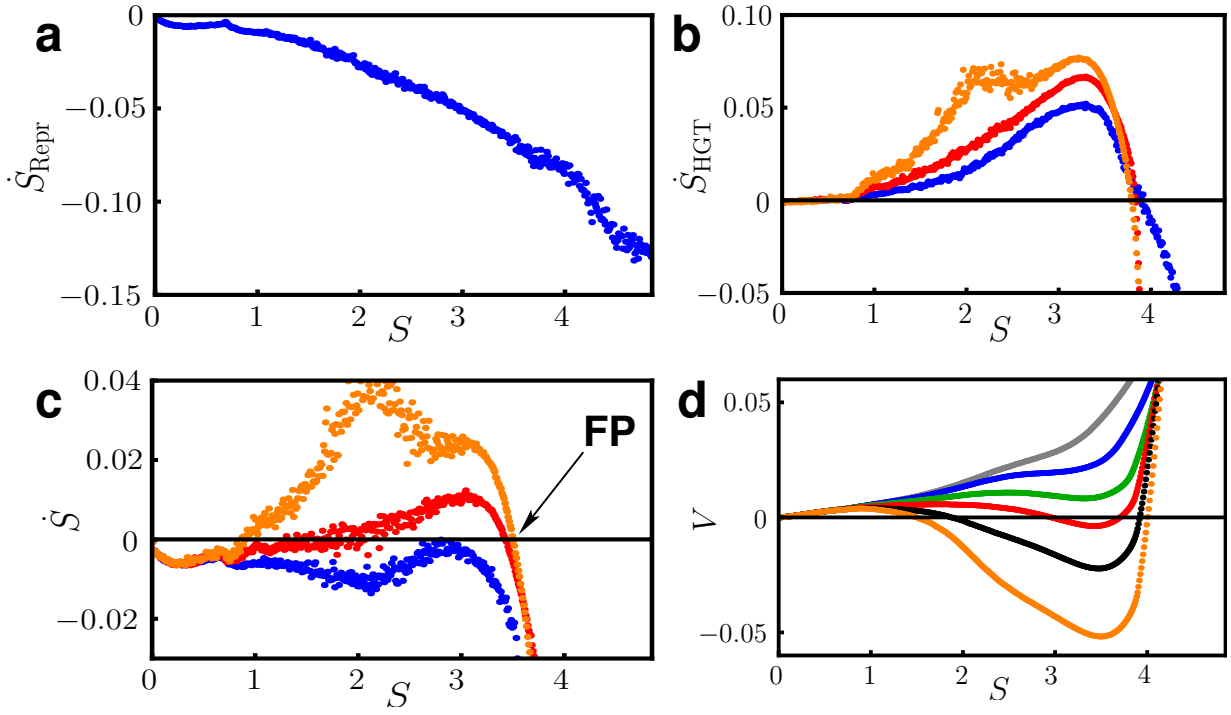

Figure 6.4.: At a critical HGT base rate c_{cr} a new fixed point at high population entropy emerges in a saddle-node bifurcation. Here the system is analyzed for which the dynamics are shown in Figure 6.1. Panels (a) and (b) show the measured rate of change of the population entropy due to reproduction and HGT from Figures 6.2 and 6.3 respectively. (b) shows results for HGT base rates $c = 1$ (blue), $c = 3$ (red) and $c = 5$ (orange). Adding the results from (a) and (b) according to equation (6.8) yields the overall rate of change $\dot{S}$ for the dynamics shown in (c). The arrow indicates the fixed point at high population entropies emerging through a saddle-node bifurcation. With equation (6.9) we define a potential $V(S)$ for the dynamics which is shown in (f) for the previous HGT base rates $c = 1$ (blue), $c = 3$ (red) and $c = 5$ (orange) and additionally for $c = 0.5$ (gray), $c = 2$ (green) and $c = 4$ (black). The potential valley at high population entropies emerges between $c = 1$ and $c = 2$ so that the critical HGT base rate must lie between these two values. Each dataset was obtained in simulations measuring the dynamics for a time $T = 10^7$.

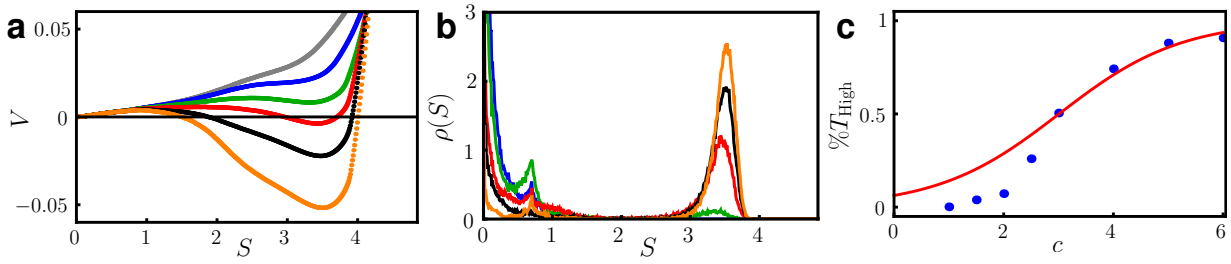

Figure 6.5.: For higher HGT-base rates c the population dynamics remain longer in the high entropy state corresponding to the depth of the potential well at high population entropies. The potential $V(S)$ from Figure 6.4d is shown again in (a). (b) illustrates the stationary probability density $\rho(S)$ to find the population in a state with population entropy S which we obtained from simulating the dynamics for a time $T = 10^4$ to let the initial condition fade away and then recording the population entropy of the population for a time $T = 5 \cdot 10^4$. The different colors in (a) and (b) indicate data sets for $c = 0.5$ (gray), $c = 1$ (blue), $c = 2$ (green), $c = 3$ (red), $c = 4$ (black) and $c = 5$ (orange). Only for $c \geq 2$ do we observe $\rho(S) > 0$ at high population entropies in (b) corresponding to the fact that only for $c \geq 2$ a potential well at high population entropies forms in (a). Integrating the data from (b) with $\int_2^{S_{\max}} \rho(S)dS$ yielded the percentage of time the dynamics remained on average in the entropy state which is illustrated by the blue dots in (c). The red line is a guide to the eye for the predictions of reaction rate theory leading to $\%T_{\text{High}} = e^{\beta\Delta}/(\alpha + e^{\beta\Delta})$ with $\Delta = E_B - E_A$ the difference of the potential wells' depths and the constants α and β determined by diffusion speed and the constraint $T_A + T_B = T$ [33]. As these constants are unknown in our system we only added this function to show the general form of the theoretical prediction. For low c this prediction fails anyway as the prerequisite for the theory does not hold as the potential well at high population entropy vanishes.

Figure 6.5c which shows the measured percentage of time the dynamics stayed in the high entropy state (cf. also Figures 6.6-6.9). We conclude that the potential we derived through the event-based analysis well fits to the observed population dynamics.

6.3. The transition's dependence on system parameters

How does the emergence of the stable high entropy state depend on the model system's parameters? Here we discuss this question in detail. We study the impact of different genome lengths l, population sizes N and the number m of HGT-links on the emergence of the high entropy state. Furthermore, instead of identical mutation probabilities and HGT base rates we use distributed mutation probabilities μ_{ij} as well as distributed competences c for different HGT-links. For all these studies we used the model from the previous section with the fitness landscape defined as above by equation (6.1) and only varied single parameters of this system.

How does the transition depend on the size of the system, i.e. the genome length l in our

model? To answer this question we need to consider different HGT-link structures for a given genome length l because the critical bifurcation point c_{cr} depends on the actual form of the HGT-link structure. Therefore, for a given genome length l we defined ten different systems by introducing ten different HGT-link structures in the way described in Section 2.2.6. As the number of genotypes for a genome length l is 2^l, we also inserted $m = m_0 \cdot 2^l$ HGT-links into a system of size l to make the systems of different size comparable. For each of these systems we then measured the average time the dynamics spent in the high entropy state in dependence of the HGT base rate c. We define the dynamics to be in the high entropy state whenever $S(t) > \log(2) \cdot l/2$. The obtained data show that the critical bifurcation point c_{cr} on average increases with the genome length l which is illustrated in Figure 6.6. The data also suggests that the variance of the dynamics over different systems decreases with increasing genome length l. However, we cannot find a scaling law describing how c_{cr} will increase with increasing l. This is mainly due to the fact that simulations for $l \gtrsim 8$ take very long, so that we only could simulate dynamics for systems up to genome length $l = 9$. Therefore, the data obtained are not sufficient to derive a scaling law and we may only conclude that the transition seems to occur at higher HGT base rates in systems with larger system size.

How does the number m of HGT-links influence the emergence of the high entropy state? To answer this question we successively added HGT-links to the model system and studied the transition of the the dynamics in dependence of the number of added links. Our results revealed that the high entropy state emerges only for a high enough number m of HGT-links, as illustrated by an example in Figure 6.7. If the number of HGT-links is small, the

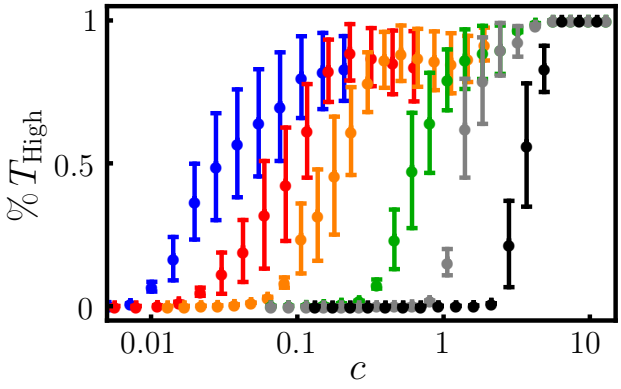

Figure 6.6.: For systems with larger genome lengths l the high entropy state emerges at larger critical HGT base rates c_{cr}. For different HGT base rates the figure shows the percentage of time T_{High} the dynamics stayed in the high entropy state averaged over systems of genome length l with ten different HGT-link structures. The error bars show the variance over the different systems. The genome lengths here were $l = 4$ (blue), $l = 5$ (red), $l = 6$ (orange), $l = 7$ (green), $l = 8$ (gray) and $l = 9$ (black). The number of HGT-links in a system with genome length l was $m = 3000 \cdot 2^{l-7}$, the fitness landscape was defined by equation (6.1) with $A = 0.1$ and the population size was $N = 1000$. Each datapoint was obtained in simulations with measurement time $T = 10^5$.

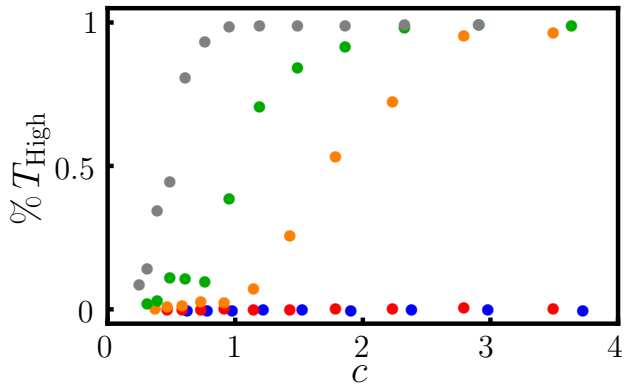

Figure 6.7.: The high entropy state only emerges if the number of HGT-links in the system is high enough. The figure shows the percentage of time the dynamics stayed in the high entropy state ($S > \log(2) \cdot l/2$) in dependence of the HGT base rate c. Different colors indicate systems with a different number m of HGT-links which are $m = 600$ (blue), $m = 800$ (red), $m = 1000$ (orange), $m = 1200$ (green) and $m = 1500$ (gray). Further system parameters were $l = 7$, $N = 1000$ and $A = 0.1$. The data were obtained in simulations of length $T = 10^5$.

dynamics remain at low population entropies for all HGT base rates. Only if the genotype space is well connected through many HGT-links the high entropy state emerges. This result may be understood considering the following argument. Assume, we place only one HGT-link into a given fitness landscape. This link moves individuals from genotype A to C. Then, for any given HGT base rate the population at A will be moved to genotype C where no further HGT events can occur to the individuals. Thus, only if all genotypes are connected by a sufficient number of HGT-links will the high entropy state emerge.

Does the impact of HGT scale with the population size similarly as fitness differences and mutations do (cf. Chapter 3)? We studied the dependence of the HGT-induced dynamics on the population size by varying the population size in a given fitness landscape with defined HGT-link structure. As we know from Section 3 fitness differences Δf have an impact on the dynamics according to $\Delta\tilde{f} = \Delta f \cdot N$ and the same applies for mutations with $\tilde{\mu}_{ij} = \mu \cdot N$. Therefore, on varying the population sizes we kept the values $\Delta\tilde{f}$ and $\tilde{\mu}_{ij}$ constant, i.e. we set the amplitude A in equation (6.1) to $A = \tilde{A}/N$ with $\tilde{A}$ fixed and similarly for the mutation probabilities $\mu_{ij} = \tilde{\mu}_{ij}/N$ in the fitness landscape. Studying the resulting dynamics for different N we found that for large N the impact of HGT scales similarly with N as fitness differences and mutations do. Therefore, we defined a rescaled HGT base rate $\tilde{c} = c \cdot N$. We then studied the system dynamics in dependence of this rescaled HGT base rate $\tilde{c}$. As an example in Figure 6.8 shows, by rescaling c to $\tilde{c}$ similar dynamics are obtained for different population sizes N and the emergence of the high entropy state always occurs at the same $\tilde{c}_{\mathrm{cr}}$. However, this only holds if the population size is large enough; for small population sizes the population cannot easily spread out in genotype space even for relatively high HGT base rates.

In all of the previous simulations we always used uniform HGT base rates. The results

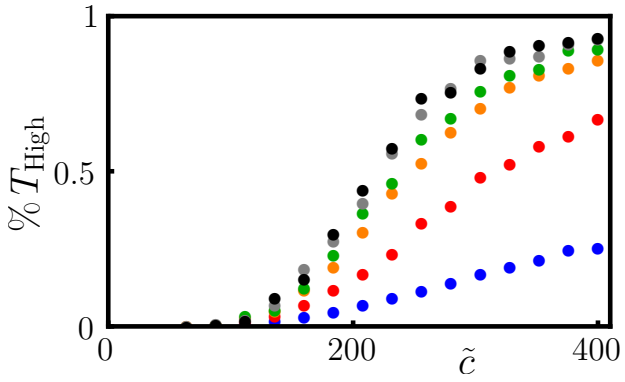

Figure 6.8.: For large enough populations the impact of HGT scales linearly in N. The figure shows the percentage of time the dynamics of an example system stayed in the high entropy state ($S > \log(2) \cdot l/2$) in dependence of the rescaled HGT base rate $\tilde{c} = c \cdot N$. The fitness landscape was the same in all simulations with $l = 6$, $m = 1500$ and $\tilde{A} = 100$, $\tilde{\mu}_{ij} = 0.1$ and only the population size was varied with $N = 100$ (blue), $N = 200$ (red), $N = 400$ (orange), $N = 600$ (green), $N = 1000$ (gray) and $N = 1500$ (black). For $N \gtrsim 600$ the dynamics are almost undistinguishable for different N, but fixed $\tilde{c}$. Each datapoint was obtained by simulating the dynamics for a time $T = 10^5$ and recording the time the population entropy fulfilled $S > \log(2) \cdot l/2$.

presented above do not change qualitatively when we use different HGT base rates for the different HGT-links. As long as there are enough HGT-links in the model system (cf. Figure 6.7) the diversity of HGT base rates for the different HGT-links seems to equal out due to the large number of HGT-links. We conclude, that qualitatively our model leads to similar results for uniform and non-uniform HGT base rates; to study the quantitative impact of distributed HGT base rates on evolutionary dynamics a more detailed model would be required. For example, each genotype could be assigned a different competence for HGT which would determine the HGT base rates of this genotype's HGT-links. In such a model the impact of the competence distribution on the dynamics could be quantitatively analyzed.

Distributed instead of uniform mutation probabilites also do not change the above results qualitatively. Actually, mutations only played a minor role for the dynamics as we kept the mutation rate low in the above simulations. Thus, the dynamics for distributed mutation probabilities also quantitatively are similar to the dynamics for uniform mutation probabilities. In the study by Jacobi and Nordahl [36] on HGT in a deterministic Eigen quasispecies model (cf. Section 2.2.4) a high entropy state for the dynamics is found similar to our results; however, there the authors show that the high entropy state vanishes for "extremely low" [36, p. 484] mutation probabilities, i.e. in their model the HGT process alone is not sufficient for the high entropy state to emerge. They analytically prove that in their model the high entropy state vanishes for mutation probabilities $\mu_{ij} = 0$. In our model however the analysis in Section 6.2 suggests that HGT by itself causes the emergence of the high entropy state and thus it should also exist for arbitrarily low mutation

probabilities. To check this we simulated the dynamics of the previous systems for very small mutation probabilities. As Figure 6.9 illustrates the high entropy state was still stable for vanishing mutation probabilities $\mu_{ij} = 0$. Naturally, if there are no mutations at all ($\mu_{ij} = 0$), the state $S = 0$ is an absorbing state of the dynamics and thus the system dynamics will eventually end up in this state. However, before reaching this absorbing state starting from an initial condition close to the high entropy state (as in the simulations for Figure 6.9) the dynamics can remain in this state for long times before being absorbed, i.e. the high entropy state is still metastable even for $\mu_{ij} = 0$. Furthermore, for arbitrarily low mutation probabilities the dynamics can reach the stable high entropy state from any initial condition for high enough HGT base rates. We conclude that the addition of a term representing HGT in the quasispecies equation only modifies the effect of mutations in the quasispecies model [36], while in our model HGT is an independent process that can drive the population dynamics towards a state of high diversity in the population.

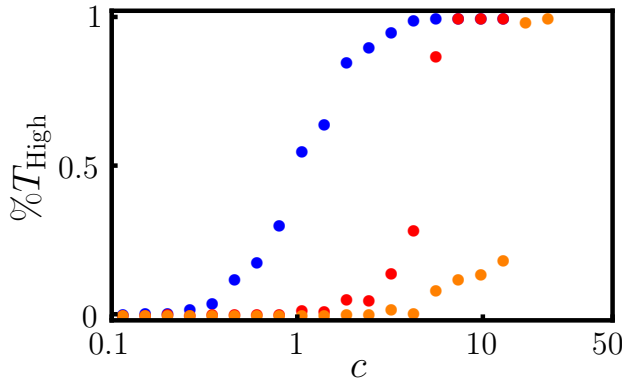

Figure 6.9.: The high entropy state is stable also for vanisishing mutation probabilities. Shown is the measured percentage of time a population stayed in the high entropy state similar to Figure 6.7 for a system with mutation probabilities $\mu_{ij} = 0.0001$ (blue) as in Figure 6.7, $\mu_{ij} = 10^{-12}$ (red) and no mutations at all $\mu_{ij} = 0$ (orange). Qualitatively the results are similar, only for higher mutation probabilities the critical transition occurs at a lower value c_{cr}. System parameters were $l = 7$, $A = 0.1$, $N = 1000$ and $m = 3000$. Each datapoint was obtained in a simulation of length $T = 10^5$ with the initial condition $S(0) = S_{\max}$.

6.4. Conclusion

Let us shortly summarize the above results. We have analyzed a high-dimensional system of many genotypes using an entropy-like variable S which we newly introduced. This approach reveals that in a static fitness landscape, where selection dynamics let a population converge to the genotype of highest fitness, sufficiently frequent HGT events cause the emergence of a new state at high population entropy S, i.e. the population is spread out in the entire genotype space. A detailed analysis – based on the population entropy S – of the average forces imposed on the population dynamics by reproduction and HGT revealed that this

new state emerges through a saddle-node bifurcation if the HGT base rate c is increased over a critical value. This transition is due to the nonlinear dependence of the HGT rate $r_{\mathrm{HGT}}(S)$ on the population entropy S of the system. With the emergence of the high entropy state the system becomes bistable stochastically switching between the high and low entropy states.

Similar bistability has been found already in deterministic models based on the quasispecies equation (2.13) with an additional term for HGT [7, 36, 68]. However, in these models the bistability vanishes for low mutation probabilities [36], i.e. the impact of HGT on the evolutionary dynamics depends on mutations to occur in the system. We conclude that in these studies HGT rather influenced the effects caused by mutations by lowering the model's mutational error threshold than providing an independent mechanism for the emergence of a stable state at high population entropy. Contrary to these results, in our stochastic model system the bistability arises through HGT alone as it will not vanish for arbitrarily small mutation probabilities. Thus, our findings suggest that HGT alone can drive a population to a state of high diversity, if the HGT rate is high enough.

Our study on how the system parameters influence the emergence of the high entropy state suggests that the critical HGT base rate c_{cr} increases with system size. However, in reality the genotype space is immense and thus in our model the high entropy state would only emerge at very high HGT base rates. Thus, further mechanisms such as spatial dimensions or special HGT-link structures may be important for the high entropy state to emerge in large systems. HGT is more probable between individuals that are close to each other in genotype space [46, 97]. We speculate that this may induce HGT-link structures for which the population will reach a state of high entropy at high HGT base rates, but the HGT-link structure may still contain the population in a certain range of genotype space. Thus, the effective genotype space "felt" by the population would not be as large as the real genotype space so that a relatively small HGT base rate is sufficient to drive the population to a high entropy state. Furthermore, in a model with spatial dimensions in different spatial regions different genotypes may be predominant. By diffusion in space and HGT these different genotypes may influence each other so that even at moderate HGT base rates a high entropy state emerges where the population is widely spread through genotype space. However, these considerations remain highly speculative until being incorporated into models for HGT.

The existence of a high entropy state for frequently occurring HGT and its disappearance for lower HGT base rates suggests a mechanism for the transition from a reactive soup dynamics in early evolution to distinct species (cf. Section 2.1.5 and [14, 96, 97]) if we assume that the competence of a population is a dynamic variable and not a fixed parameter as in our model system [12, 47, 48, 86]. Our results indicate that a population exhibiting a high competence for HGT may exhibit dynamics where the population occupies a large region of genotype space (cf. Figure 6.10). Through rare events the population may converge to a more concentrated state close to a fitness peak. Thus, this population has a higher fitness and could outcompete other populations spread out in genotype space. Therefore, geno-

types exhibiting lower competence are selected for as the population on average exhibits a higher fitness. As we assume competence to be a dynamically changing property of a population, the population could evolve to lower competence as this would on the long term increase the population's average fitness. Our results suggest that for a lower competence of the individuals the spread out state would become unstable and thus the population would evolve mainly according to the selection process in the fitness landscape. Thus, the disappearance of the high entropy state would correspond to the transition from the reactive soup to the first distinct species. This is illustrated in Figure 6.10. To understand how the transition from reactive soup to distinct species may occur, we thus need to incorporate dynamically changing competences [12, 47, 48, 86] into future more detailed models. Furthermore, we emphasize that our model may only explain the emergence of the first species from a previous evolutionary state where no species existed at all. To understand how the ensuing process of speciation under the influence of HGT proceeds, our model is not detailed enough. For such studies models would be needed where different species can coexist over large scales of time, i.e. that probably a more detailed fitness landscape including dynamic fitness would be a requirement for such models [65].

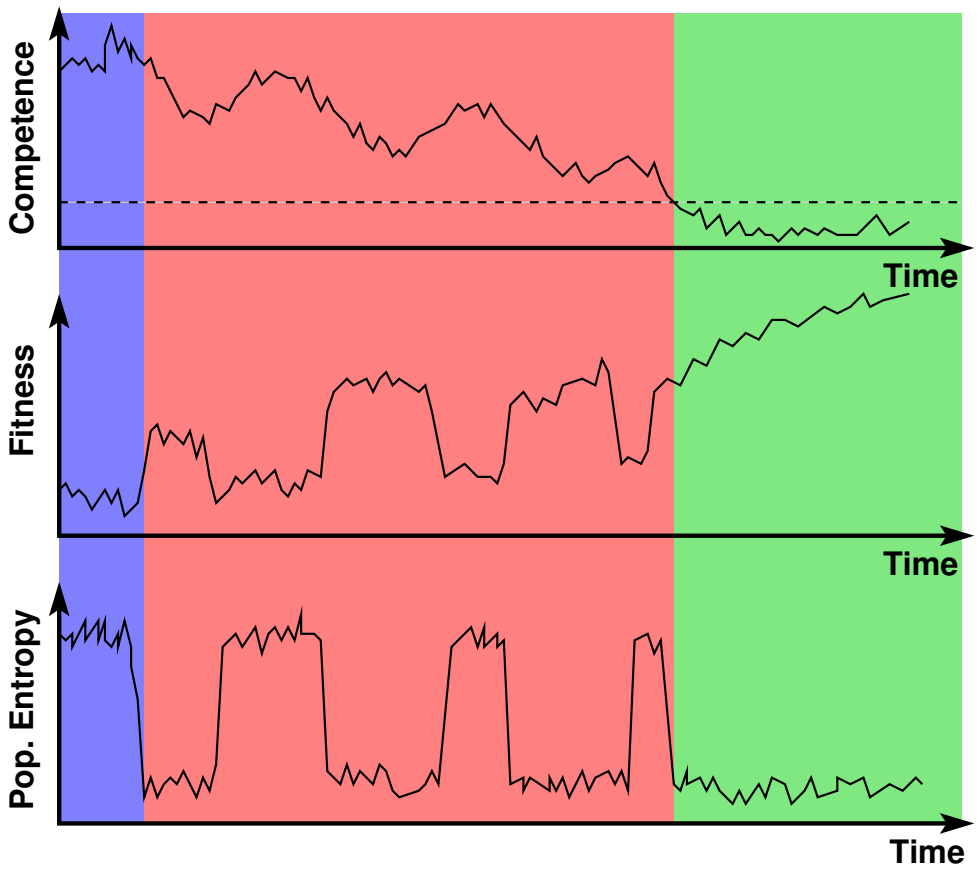

Figure 6.10.: A possible scenario for the evolution of distinct species from a reactive soup. The three graphs suggest how the average competence, the average fitness compared to an initial fitness and the population entropy of a population may evolve in the transition from a reactive soup to distinct species. In the initial state the competence is high, so that HGT drives the dynamics; the population exhibits a high population entropy and low average fitness. This state is marked in blue in this Figure. Through a stochastic switching the dynamics can reach a state of low entropy where the fitness is higher. Here the population could evolve slowly towards lower competence. Thus, the dynamics switch back and forth between the low and the high entropy state remaining longer and longer in the low entropy state as the competence decreases. This part of the evolutionary transition is marked in red. When the competence goes below a critical value (marked by the dashed line) the high entropy state becomes unstable and the dynamics remain in the low entropy state and, as distinct species evolve, the population's average fitness increases. This phase of the evolutionary dynamics is marked in green.

7. Summary and Conclusions

All life on earth has been and still is being shaped by evolutionary processes. Still, many aspects of evolutionary dynamics are far from being understood [20]. One such aspect of evolutionary dynamics is its underlying stochasticity, resulting for example from stochastic reproduction and death processes or random mutation events. Yet, the consequences of such stochasticity for the dynamics remain to be clarified. We therefore analyzed different aspects of stochastic evolutionary dynamics in this thesis. We developed and analyzed simple individual-based model systems where we considered finite-size populations of idealized individuals stochastically reproducing and dying. Such basic models focus on catching the essential features of evolution and yield qualitative conclusions about characteristic mechanisms in evolutionary dynamics [6]. As long as we have not even fully grasped the basic mechanisms driving evolution, such simple models seem more appropriate for studying evolutionary dynamics than more detailed quantitative ones. Our work focussed on two aspects of stochastic evolutionary dynamics: The impact of dynamic fitness and of horizontal gene transfer (HGT).

Dynamic Fitness

Fitness measures are a basic feature of many theoretical models of evolution because of their basic role in capturing how well individuals fare under specific environmental conditions [20,63]. Fitness therefore depends on both the (possibly changing) environment and, in particular, the interactions with other individuals in a population. To grasp the impact of such effects on the evolutionary dynamics we modelled fitness as a dynamic variable itself. In particular, interactions between individuals imply fitnesses that depend on the frequencies of the different genotypes present in a population.

Our knowledge of how a population will evolve in a dynamically changing fitness landscape due to such interactions remained yet incomplete. In fact, previous studies only systematically analyzed special linear and quadratic instances of such frequency-dependent fitness [65, 87–89] arising from game theoretic considerations for the interactions. This approach directly yields a linear dependence of fitness on genotype frequency, i.e. linear fitness functions (cf. equations (2.8)-(2.9) in Section 2.2.3). Normalizing these fitness functions causes the emergence of a specific quadratic dependence (cf. equations (2.11)-(2.12) in Section 2.2.4). Thus, a game-theoretic approach to interactions only yields linear or selected quadratic instances of fitness functions. Yet, experimental studies suggest that

fitness may depend nonlinearly on the genotype frequencies of a population [51]. In this thesis, we thus took a more general perspective on frequency-dependent fitness, considering a general class of arbitrary nonlinear fitness functions. Furthermore, mutation rates are often highly diverse for different genotypes [22, 77], but previous studies only considered mutation rates to be equal for all genotypes in the system [3, 65, 88, 100]. To complete our general approach, we here explicitly considered such diverse mutation rates.

Considering a population which exhibits such nonlinear frequency-dependent fitness and diverse mutation rates, we studied how these effects impact the population's dynamics. In Chapter 3, we analyzed the joint influence of frequency-dependent fitness, diverse mutation rates and genetic drift on the population dynamics in a model system of two genotypes. Here, we explicitly considered a general class of fitness functions that may depend on the genotypes' frequencies in an arbitrarily nonlinear way (cf. equations (3.20) and (3.21)). For this general setting we derived an analytic expression (cf. equation (3.19)) for the stationary probability distribution that a given genotype one exhibits a certain frequency in the population. This expression revealed that the population dynamics tend to switch stochastically between different metastable states that are induced by the interplay of the frequency-dependent fitness, mutations and genetic drift (cf. Figure 3.2). We found that fitness depending nonlinearly on genotype frequency may cause many of these metastable states (cf. Figure 3.3). Furthermore, if the genotypes exhibit different mutation rates, these states may be shifted and new such states may emerge (cf. Figures 3.5 and 3.6).

We conclude that frequency-dependent fitness together with heterogeneous mutation rates induces complex evolutionary dynamics, in particular if the interactions imply nonlinear fitness functions [2]. Previous studies had focussed on special linear or simple quadratic instances of the class of fitness functions presented here and considered mutation rates to be equal for all genotypes in the system [65, 87–89]. They revealed that genetic drift and mutations counteract each other [88] and that frequency-dependent fitness can cause the emergence of a metastable state where both genotypes coexist in the population [65, 88]. Going beyond these findings, our results indicate that frequency-dependent fitness may even cause the emergence of many of these metastable states. We further revealed how the mutation rates affect the genotype frequencies at which these stable states are located. Thus, if population dynamics in real environments are observed that repeatedly shift between different frequencies, even though environment remains stable, the observed dynamics may – according to our results – indicate interactions between the individuals that cause a nonlinear fitness dependence.

Dynamic fitness may also play an important role in stabilizing the dynamics of populations of dynamically varying size (cf. Chapter 4). So far, theoretical studies have often focussed on models of populations of fixed size described by reproduction processes such as, for example, the Moran or the Wright-Fisher process (cf. Section 2.2.2) [20, 63, 84, 88]. Experimental findings indicate that the genetic diversity of a population may depend on variations in population size [71, 83] and it remains an open question how evolutionary dynamics are influenced by dynamically changing population size. To study evolutionary dynamics with

changing population sizes we developed a model for a reproduction process based on independent birth and death events which we called IBD process (cf. Section 2.2.2). We studied populations evolving under this reproduction process in Chapter 4. We found that the population would either go extinct with high probability after only relatively few generations or it would grow infinitely if we considered individuals with static fitnesses. Yet, dynamic fitness stabilizes the population size, so that the probability of a rapid extinction is strongly reduced and infinite growth is impossible (cf. Figures 4.4 and 4.9). We conclude that dynamic fitness may be a stabilizing mechanism in evolutionary dynamics.

These novel models of evolution based on independent birth and death reproduction processes and dynamic fitness seem promising to gain valuable insights into evolutionary phenomena caused by the interplay of varying population sizes and individuals' interactions, e.g. quasicyclic behaviour in predator-prey systems [6, 53] or punctuated equilibrium dynamics [4, 31, 60, 67]. In Section 4.5 we demonstrated in an example system how predator-prey interactions [13, 50, 93] may be described using dynamic fitness which yields a relatively simple, individual-based model with variable population size. We found that this system exhibits a range of complex dynamics including quasi-cycles and punctuated equilibria (cf. Figure 4.11). Quasicyclic behaviour, the resonant amplification of the dynamics' underlying stochasticity, has only recently been studied, e.g. in [53, 70] where it was found that quasi-cycles occur "whenever the underlying deterministic population model exhibits damped oscillations towards an equilibrium" [70, p. 63]. Our study now confirms that quasi-cycles can be exhibited by the dynamics of finite populations.

It was suggested that evolution does not proceed gradually but rather in sudden steps [31, 72]. The extinction of one species through stochastic fluctuations can elicit an avalanche where after a short time further species go extinct as well [4]. Thus, large extinction events may be due to the population being in a self-organized critical state: The dynamics seem to be stable for some time, but then relatively small fluctuations can cause large extinction avalanches, i.e. many species go extinct in a short period of time. In a self-organized critical state the size of these avalanches is power law distributed. Previous studies suggested that punctuated equilibrium dynamics may describe extinction events in evolution, but to our knowledge until now there have been only very abstract population-level models such as e.g. the Bak-Sneppen model describing this effect [4, 31, 60, 67]. The analysis of the dynamics in our more detailed individual-based model system revealed that the population indeed is in such a state of punctuated equilibrium. As our model is more detailed than the abstract models in [4, 60, 67], it may help to understand how such punctuated equilibrium dynamics could emerge in real biological systems and how these dynamics are influenced by other features such as quasi-cycles. We conclude that our model system provides a promising approach to study in detail how interactions between individuals and the stochasticity induced by reproduction and death processes give rise to complex evolutionary dynamics.

Horizontal gene transfer

Regarding horizontal gene transfer, there are only very few theoretical studies that investigate its impact on evolutionary dynamics. However, it was proposed that HGT may have played an important role in early evolution [14, 45] and we therefore need to explore how HGT affects evolutionary dynamics. Here we developed a stochastic process modelling HGT (cf. Section 2.2.6) which captures the essence of HGT: In one HGT event an individual of genotype A takes up genetic material from an individual of genotype B with a certain competence to mutate to genotype C. Thus, this simple model is applicable to individual-based population models, while previous approaches considered HGT on the population level [7, 36, 68] or applied less general agent-based HGT-models with many degrees of freedom [92]. We conclude that our newly introduced model well balances between these approaches, as it is as simple as possible, but still captures the essence of stochastic HGT.

Can HGT be beneficial for evolving populations, i.e. can it increase the average fitness of the population? It was often proposed that HGT may confer a fitness advantage for populations adapting to new or changing environments [73, 86, 92]. Yet, recent studies on how HGT affects fitness focussed on fixed environments [73, 92]. Finding that HGT does not increase fitness in such fixed environments, the authors suggested that HGT might be beneficial for populations in changing environments. In Chapter 5 we have now confirmed that HGT can be beneficial for populations adapting to changing environments (cf. Figure 5.2). In addition, our results also reveal that the population better adapts to the changing environment through a high mutation rate than through increased HGT rates; if all genotypes exhibit high enough identical mutation rates, adaptation is not improved by HGT in the model (cf. Figure 5.3). Yet, mutation rates are often diverse for different genotypes in real biological systems [22, 77]. Our results indicate that in fitness landscapes including such diverse mutation rates, a population will adapt optimally to the changing landscape only if HGT occurs in the population (cf. Figure 5.3). Because of this finding we propose that when an adapting population gets stuck at genotypes of low mutation rate, HGT can help it to jump over such genotypes. We conclude that HGT can confer a fitness advantage for adapting populations. However, whether or not HGT yields a fitness increase depends also on the mutation rates in genotype space. Contrary to the simplistic proposition that HGT will be beneficial for adapting populations [73, 92], we conclude that the impact of HGT on a population's fitness in changing environments is more complex. In particular, our results reveal that the diversity of mutation rates may be an important factor in determining whether or not HGT is beneficial.

How do frequent HGT events influence evolutionary dynamics and how do selection-dominated dynamics at lower HGT rates emerge? It has been proposed that HGT played an important role in early evolution [45] where it may have dominated the evolutionary dynamics so that no distinct species formed, but rather all individuals exchanged genetic material at a high rate [14]. This state is sometimes referred to as a "reactive soup" and it is unclear how evolution could proceed from this state to form the first distinct species [14]. In our theoretical investigations (Chapter 6) we found that for frequent HGT there exists a

stable state where the population is spread out in the entire genotype space (cf. Figure 6.1). We identify this as a reactive soup state because no distinct species can be identified in it and the individuals exchange genetic material at a high rate. A mean field analysis of the forces induced by reproduction and HGT events revealed that this state emerges through a saddle-node bifurcation at a critical HGT base rate c_{cr} so that for all higher HGT base rates the reactive soup state is stable (cf. Figure 6.4). The dynamics then switch stochastically between two coexisting states, a selection-dominated state at low population entropy, where the population is concentrated around the fittest genotype, and an HGT-dominated reactive soup state at high population entropy. Thus, the system is bistable above c_{cr} and the dynamics remain in the reactive soup state longer when the HGT base rate is higher.

Our analysis demonstrates that mutations are not necessary for the reactive soup state to emerge (cf. also Figure 6.9), i.e. HGT alone can drive the dynamics towards a stable state where the population is spread throughout genotype space. Previous studies on HGT in quasispecies models had already found a bistability between two states of low and high population entropy [7, 36, 68], but it was attributed to the fact that HGT may lower the error threshold, i.e. the critical mutation rate above which too many mutations occur for a species to remain close to a fitness peak. The quasispecies model is designed for the analysis of how large mutation rates affect a population and HGT was introduced as an additional factor to this model. Thus, the bistability in this model vanished for low mutation rates since it mainly modifies the effect of mutations [36]. Contrary to this finding, our results highlight that HGT alone suffices to create the reactive soup state and mutations are not necessary for its emergence.

How could distinct species have evolved from a reactive soup where HGT dominated as proposed for example by Woese [96, 97]? Our findings suggest how this evolutionary transition could have occurred (cf. Figure 6.10): In an initial evolutionary state the individuals exchange genetic material at a high rate and thus form a reactive soup. After a long waiting time stochastic fluctuations could cause the dynamics to converge to the selection-dominated state. In this state the population would evolve to the peaks in the fitness landscape and thus obtain an increased fitness before switching back into the reactive soup state. Thus, after repeated switchings, the population could evolve to lower competences for HGT implying that the reactive soup state would vanish. After this transition the evolutionary dynamics are dominated by selection and thus distinct species form. To the ensuing process of further speciation under the influence of HGT our model is not applicable, as it lacks the details to model such a process. Still, we emphasize that the emergence of the first species from a reactive soup may be well understood with the model we applied. Of course, for now the sequence of events we propose for the emergence of the first species remains speculative as our model does not include competences which change explicitly with time. But recent first results by Vogan and Higgs [92] obtained in simulations of a simple model system indicate that a population may eventually evolve towards lower competence. In future research, more detailed models should include dynamically changing competences to check if and under which circumstances a transition from reactive soup dynamics to distinct species is possible in the way described above.

Outlook

To summarize, this thesis has contributed several steps towards our understanding of how basic mechanisms such as selection, mutation and HGT shape evolutionary dynamics. Our results suggest possible directions of future research: As the example system in Section 4.5 demonstrates, the IBD process (cf. Section 2.2.2) in combination with dynamic fitness may well be fit to study in greater detail how interactions between individuals and stochastic reproduction processes cause the emergence of phenomena such as punctuated equilibrium dynamics which previously were only analyzed in more abstract models [4, 60, 67].

Our simulations in Chapter 5 revealed that HGT can be beneficial for adapting populations. Still, the fitness advantage that the population obtained through HGT was relatively small. There are experimental studies suggesting that the individuals' competence for HGT may fluctuate over time and especially may depend on their fitness [47, 48]. Such a mechanism could increase the beneficial effect of HGT, for example if competence increases with decreasing fitness, as in such a setting HGT would preferentially drive populations towards new genotypes if they already exhibit a low fitness. Future research models should therefore include such dynamic competences to check how this mechanism may affect the impact of HGT on a population's fitness.

In Chapter 6 we showed how a transition from HGT-dominated reactive soup dynamics to selection-dominated dynamics may occur in evolutionary dynamics. However, our detailed model analysis also indicates that the critical HGT base rate – at which the reactive soup emerges – increases with system size. As in reality genotype space is very large [63, 78], within our model framework there should be no reactive soup in large systems. However, there may be further factors influencing the emergence of the reactive soup at lower HGT rates. For example, we assumed a random HGT-link structure, but there is evidence that HGT is more likely between closely related organisms [46, 97]. In large systems the resulting HGT-link structure may thus cause the emergence of a localized reactive soup where the individuals are spread out in a large but confined region of genotype space, i.e. the population does not need to spread out through the entire genotype space. The transition to such a state could thus occur at lower HGT rates. Also, spatial dimensions may play an important role as in such a setting different genotypes can be present at different locations of the system. It would thus be easier for the population to maintain a high diversity so that additional spatial dimensions may also cause the reactive soup state to emerge at lower HGT rates. We conclude that our model has revealed how HGT may drive a population into a reactive soup state, but more refined models are needed to gain a better understanding of the details of how this state may emerge.

Basic models catching the essence of evolutionary processes yield much insight into the mechanisms driving evolution. In this thesis, we exploited this approach obtaining qualitative predictions about dynamical features of evolution, in particular how dynamic fitness and HGT drive stochastic evolutionary dynamics. These predictions contribute to the growing knowledge of how evolution proceeds under different circumstances and also point towards open problems yet to be solved. We are confident that following the suggestions made here these questions will be answered in the near future.

A. Birth-death processes' absorption probabilities and mean time to absorption

Here, considering birth-death process with birth rates λ_k and death rates μ_k we derive the absorption probability p_k^H, i.e. the probability that the process will reach the state 0 from an initial state k , as well as the mean time to absorption $\overline{T}_k$ from state k (cf. Section 2.3.3). Our analysis here follows the arguments in [38].

Considering the probabilities with which the process will move from state k to $k+1$ and $k-1$ we obtain the recursion formula

$$p_k^H = \frac{\lambda_k}{\mu_k + \lambda_k} p_{k+1}^H + \frac{\mu_k}{\mu_k + \lambda_k} p_{k-1}^H \tag{A.1}$$

with the additional condition $p_0^H = 1$. Rewriting this equation yields

$$(p_{k+1}^H - p_k^H) = \frac{\mu_k}{\lambda_k}(p_k^H - p_{k-1}^H) \tag{A.2}$$

We apply this equation iteratively and obtain

$$p_{k+1}^H - p_k^H = (p_1^H - p_0^H) \prod_{j=1}^{k} \frac{\mu_j}{\lambda_j}. \tag{A.3}$$

Taking the sum of this equation and using $p_0^H = 1$, this becomes

$$p_{k+1}^H - p_1^H = (p_1^H - 1) \sum_{i=1}^{k} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}. \tag{A.4}$$

The variables p_k^H denote probabilities and are thus bounded by 1. Therefore, if the sum

$$\sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} \tag{A.5}$$

does not converge, we necessarily have $p_1^H = 1$ and consequently also $p_k^H = 1$ for all

$k \in \mathbb{N}$. This means, that if the sum (A.5) does not converge, the process will end up in the absorbing state with absolute certainty from any initial state k. Assuming that the sum (A.5) converges, we arrive at a solution for p_k^H with the following argument. First, we remark that $p_{k+1}^H \geq p_k^H$ as the process coming from state $k+1$ has to pass through state k to arrive in the absorbing state. Furthermore, if the sum (A.5) converges, the absorption probability has to fulfill $\lim_{k\to\infty} p_k^H = 0$. Thus, in the limit $k \to \infty$ equation (A.4) becomes

$$p_1^H = (p_1^H - 1) \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} \tag{A.6}$$

which we solve for p_1^H to obtain

$$p_1^H = \frac{\sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}}{1 + \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}}. \tag{A.7}$$

Using this result together with equation (A.4) yields

$$p_k^H = \frac{\sum_{i=k}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}}{1 + \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}} \tag{A.8}$$

for all $k \in \mathbb{N}$. In summary we have found that the absorption probability is

$$p_k^H = \begin{cases} 1 & \text{if } \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} = \infty \\ \frac{\sum_{i=k}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}}{1+\sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}} & \text{if } \sum_{i=1}^{\infty} \prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} < \infty \end{cases}. \tag{A.9}$$

In a similar way we calculate the mean time to absorption $\overline{T}_k$ from state k. As the mean waiting time in state k is $(\mu_k + \lambda_k)^{-1}$, we obtain the recursion formula

$$\overline{T}_k = \frac{1}{\mu_k + \lambda_k} + \frac{\mu_k}{\mu_k + \lambda_k} \overline{T}_{k-1} + \frac{\lambda_k}{\mu_k + \lambda_k} \overline{T}_{k+1} \tag{A.10}$$

with the condition $\overline{T}_0 = 0$. We rewrite this to

$$\overline{T}_k - \overline{T}_{k+1} = \frac{1}{\lambda_k} + \frac{\mu_k}{\lambda_k} (\overline{T}_{k-1} - \overline{T}_k) \tag{A.11}$$

and iterate this relation to obtain

$$\overline{T}_k - \overline{T}_{k+1} = \sum_{i=1}^{k} \frac{1}{\lambda_i} \prod_{j=i+1}^{k} \frac{\mu_j}{\lambda_j} - \prod_{j=1}^{k} \frac{\mu_j}{\lambda_j} \overline{T}_1 \tag{A.12}$$

where we applied $\overline{T}_0 = 0$. Here we define that $\prod_{j=k+1}^{k} \mu_j/\lambda_j = 1$. Further, we define

$$\chi_i = \frac{1}{\mu_i} \prod_{j=1}^{i-1} \frac{\lambda_j}{\mu_j} \tag{A.13}$$

so that equation (A.12) becomes

$$(\overline{T}_k - \overline{T}_{k+1}) \prod_{j=1}^{k} \frac{\lambda_j}{\mu_j} = \sum_{i=1}^{k} \chi_i - \overline{T}_1. \tag{A.14}$$

Due to the property $\overline{T}_{k+1} > \overline{T}_k$ of the mean waiting times, it immediately follows that $\overline{T}_1 = \infty$ if the sum $\sum_{i=1}^{\infty} \chi_i$ diverges. Suppose that $\sum_{i=1}^{\infty} \chi_i < \infty$, then we obtain the limit

$$\lim_{k\to\infty} (\overline{T}_k - \overline{T}_{k+1}) \prod_{j=1}^{k} \frac{\lambda_j}{\mu_j} = 0 \tag{A.15}$$

so that

$$\overline{T}_1 = \sum_{i=1}^{\infty} \chi_i. \tag{A.16}$$

This result together with equation (A.14) yields

$$\overline{T}_k = \sum_{i=1}^{\infty} \chi_i + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} \cdot \sum_{j=i+1}^{\infty} \chi_j \right]. \tag{A.17}$$

To summarize, we have found that the mean time to absorption from state k is

$$\overline{T}_k = \begin{cases} \infty & \text{if } \sum_{i=1}^{\infty} \chi_i = \infty \\ \sum_{i=1}^{\infty} \chi_i + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i} \frac{\mu_j}{\lambda_j} \cdot \sum_{j=i+1}^{\infty} \chi_j \right] & \text{if } \sum_{i=1}^{\infty} \chi_i < \infty \end{cases}. \tag{A.18}$$

B. The Fokker-Planck equation of the two-genotype system

In the limit of large population sizes N the master equation (3.4) is well approximated by a Fokker-Planck equation [74]. In the following we introduce the necessary transformation and derive the Fokker-Planck equation (2.40) corresponding to the master equation (3.4).

We use the transformation

$$x = \frac{k}{N}, \qquad s = t \cdot F(N), \qquad \tilde{\mu}_{ij} = \mu_{ij} \cdot G(N), \qquad \tilde{g}_i(x) = g_i(k) \cdot H(N) \tag{B.1}$$

where x represents the frequency of genotype A and $F(N)$, $G(N)$ and $H(N)$ are scaling functions to be determined in the following such that in the limit $N \to \infty$ the terms in equation (3.4) remain finite. With this transformation the probability distribution $p_k(t)$ becomes a probability density

$$\rho(x, s) = p_{Nx}(t)N|_{t=s/F(N)} \tag{B.2}$$

in the Fokker-Planck equation. Defining

$$x_+ = x + \frac{1}{N}, \qquad x_- = x - \frac{1}{N} \tag{B.3}$$

we substitute the above transformation into the master equation (3.4) and obtain

$$\begin{aligned}
\frac{d\rho(x,s)}{ds}F(N) \;=\; & \frac{N^2}{N+1}\Big\{\Big[(1-\mu_{AB})(1+g_A(x_-))x_-(1-x_-)+\mu_{BA}(1+g_B(x_-))(1-x_-)^2\Big]\rho(x_-,s) \\
& +\Big[(1-\mu_{BA})(1+g_B(x_+))(1-x_+)x_+ + \mu_{AB}(1+g_A(x_+))x_+^2\Big]\rho(x_+,s) \\
& -\Big[(1-\mu_{AB})(1+g_A(x))x(1-x)+\mu_{BA}(1+g_B(x))(1-x)^2 \\
& +(1-\mu_{BA})(1+g_B(x))x(1-x)+\mu_{AB}(1+g_A(x))x^2\Big]\rho(x,s)\Big\}
\end{aligned} \tag{B.4}$$

and find that by choosing $F(N) = N + 1$, $G(N) = H(N) = N$ in the limit $N \to \infty$ the terms remain finite. In the following calculation we drop the time argument s in the density $\rho(x, s)$ to make the notation more transparent. Further, we introduce the mean mutation rate

$$\tilde{\mu} := \frac{\tilde{\mu}_{AB} + \tilde{\mu}_{BA}}{2} = \frac{N}{2}(\mu_{AB} + \mu_{BA}) \tag{B.5}$$

and the mutation rate difference

$$\Delta\tilde{\mu} := \frac{\tilde{\mu}_{AB} - \tilde{\mu}_{BA}}{2} = \frac{N}{2}(\mu_{AB} - \mu_{BA}). \tag{B.6}$$

These definitions and a reordering the terms in equation (B.4) yields

$$\begin{aligned}
\frac{d\rho(x)}{ds} \;=\; & N^2\Big\{ -2x(1-x)\rho(x) + x_+(1-x_+)\rho(x_+) + x_-(1-x_-)\rho(x_-) \\
& \quad + \frac{\tilde{\mu}}{N}\left[-(1-2x)^2\rho(x) + \frac{1}{2}(1-2x_+)^2\rho(x_+) + \frac{1}{2}(1-2x_-)^2\rho(x_-)\right]\Big\} \\
& + N\Big\{ \tilde{g}_A(x_-)x_-(1-x_-)\rho(x_-) - \tilde{g}_A(x)x(1-x)\rho(x) \\
& \quad + \tilde{g}_B(x_+)x_+(1-x_+)\rho(x_+) - \tilde{g}_B(x)x(1-x)\rho(x) \\
& \quad + \frac{\tilde{\mu}}{2}\Big[(1-2x_-)\rho(x_-) - (1-2x_+)\rho(x_+)\Big] \\
& \quad + \Delta\tilde{\mu}\Big[x_+\rho(x_+) - x\rho(x) - (1-x_-)\rho(x_-) + (1-x)\rho(x)\Big] \\
& \quad + \frac{\tilde{\mu}}{N}\Big[\left(\tilde{g}_A(x_+)x_+^2 + \tilde{g}_B(x_+)(x_+^2 - x_+)\right)\rho(x_+) - \left(\tilde{g}_A(x)x^2 + \tilde{g}_B(x)(x^2 - x)\right)\rho(x) \\
& \quad + \left(\tilde{g}_A(x_-)(x_-^2 - x_-) + \tilde{g}_B(x_-)(1-x_-)^2\right)\rho(x_-) - \left(\tilde{g}_A(x)(x^2 - x) + \tilde{g}_B(x)(1-x)^2\right)\rho(x)\Big] \\
& \quad + \frac{\Delta\tilde{\mu}}{N}\Big[\left(\tilde{g}_A(x_+)x_+^2 - \tilde{g}_B(x_+)(x_+^2 - x_+)\right)\rho(x_+) - \left(\tilde{g}_A(x)x^2 - \tilde{g}_B(x)(x^2 - x)\right)\rho(x) \\
& \quad + \left(\tilde{g}_A(x_-)(x_-^2 - x_-) - \tilde{g}_B(x_-)(1-x_-)^2\right)\rho(x_-) - \left(\tilde{g}_A(x)(x^2 - x) - \tilde{g}_B(x)(1-x)^2\right)\rho(x)\Big]\Big\}.
\end{aligned}$$

The first terms with the factor N^2 in front in the limit $N \to \infty$ become the second order derivatives of $x(1-x)\rho(x)$ with respect to x. For example, for the first three terms in the above equation the limit is given by

$$\begin{aligned}
& \lim_{N\to\infty} N^2\left\{-2x(1-x)\rho(x) + x_+(1-x_+)\rho(x_+) + x_-(1-x_-)\rho(x_-)\right\} \\
= & \lim_{N\to\infty} \frac{-2x(1-x)\rho(x) + \left(x+\frac{1}{N}\right)\left(1-\left(x+\frac{1}{N}\right)\right)\rho\left(x+\frac{1}{N}\right) + \left(x-\frac{1}{N}\right)\left(1-\left(x-\frac{1}{N}\right)\right)\rho\left(x-\frac{1}{N}\right)}{1/N^2} \\
= & \lim_{h\to 0} \frac{-2x(1-x)\rho(x) + (x+h)\left(1-(x+h)\right)\rho(x+h) + (x-h)\left(1-(x-h)\right)\rho(x-h)}{h^2} \\
= & \frac{\partial^2}{\partial x^2}\left[x(1-x)\rho(x)\right]
\end{aligned}$$

where we used

$$\lim_{h\to 0}\left[f(x+h) + f(x-h) - 2f(x)\right]/h^2 = \frac{\partial^2}{\partial x^2}f(x) \tag{B.7}$$

in the last step [9]. Similarly, in the limit $N \to \infty$ the terms with N in front become first order derivatives with respect to x, where all terms with the factor $1/N$ vanish, so that the three terms $(\tilde{g}_A(x) - \tilde{g}_B(x))x(1-x)\rho(x)$, $\tilde{\mu}(1-2x)\rho(x)$ and $-\Delta\tilde{\mu}\rho(x)$ remain. Taken together we obtain the Fokker-Planck equation

$$\frac{\partial\rho(x,s)}{\partial s} = -\frac{\partial}{\partial x}\Big[\{(\tilde{g}_A(x) - \tilde{g}_B(x))x(1-x) + \tilde{\mu}(1-2x) - \Delta\tilde{\mu}\}\rho(x,s)\Big] + \frac{\partial^2}{\partial x^2}\Big[x(1-x)\rho(x,s)\Big] \tag{B.8}$$

approximating the master equation (3.4) in the limit of large population sizes N.

C. Time scales of the stabilized IBD process

To analyze the survival time distribution of the stabilized IBD process in Section 4.2 we need to calculate the time scale

$$\tau = \sum_{j=1}^{N_*} \frac{1 - \sum_{i=1}^{j-1} p_i^*}{\mu_j p_j^*} \tag{C.1}$$

derived in Section 2.3.4 using Kramers' method [29, 44]. Here p_j^* is the quasistationary distribution of the process given by

$$p_k^* = \frac{\prod_{j=1}^{k-1} \frac{\lambda_j}{\mu_{j+1}}}{\sum_{l=1}^{\infty} \prod_{j=1}^{l-1} \frac{\lambda_j}{\mu_{j+1}}} \tag{C.2}$$

(cf. equation (3.8)); λ_j and μ_j are the birth and death rates of the process in state j respectively. For the stabilized IBD process they are given by $\lambda_j = N_*$ and $\mu_j = j$ leading to (cf. equation (4.13))

$$p_k^* = \frac{N_*^k}{k!} \cdot \frac{1}{e^{N_*} - 1}. \tag{C.3}$$

According to equation (C.1) this quasistationary distribution yields a time scale

$$\begin{aligned} \tau &= \sum_{j=1}^{N_*} \frac{1 - \sum_{i=1}^{j-1} p_i^*}{\mu_j p_j^*} \\ &= \sum_{j=1}^{N_*} \left[1 - \sum_{i=1}^{j-1} \frac{N_*^i}{i!} \frac{1}{e^{N_*} - 1}\right] \left[j \cdot \frac{N_*^j}{j!} \frac{1}{e^{N_*} - 1}\right]^{-1} \\ &\approx \left(e^{N_*} - 1\right) \sum_{j=1}^{N_*} \frac{(j-1)!}{N_*^j} \left[1 - e^{-N_*} \sum_{i=1}^{j-1} \frac{N_*^i}{i!}\right] \end{aligned} \tag{C.4}$$

where we approximated $\left[e^{N_*} - 1\right]^{-1} \approx e^{-N_*}$ for $N_* \gg 1$ in the last step. The term

$$e^{-N_*} \sum_{i=1}^{j-1} \frac{N_*^i}{i!} \tag{C.5}$$

can be neglected for $N_* \gg 1$ as the sum only converges to e^{N_*} for $j \to \infty$. In particular, for small j the factor e^{-N_*} dominates and for large j there is the overall prefactor N_*^{-j} in equation (C.4). Thus, in first order approximation we only consider the first term $j = 1$ of the sum in equation (C.4) and obtain

$$\tau \approx \left(e^{N_*} - 1\right) \left[\frac{1 - e^{-N_*}}{N_*} + \mathcal{O}\left(N_*^{-2}\right)\right] \approx \frac{e^{N_*} - 1}{N_*} \tag{C.6}$$

which increases exponentially in the sustained population size N_*, so that the survival time distribution fulfills

$$p_S(t) \approx \exp\left(-\frac{N_*}{e^{N_*} - 1} t\right). \tag{C.7}$$

D. The mean extinction time of the scaled IBD process

Here we derive an estimate of the mean extinction time $\overline{T}$ in the scaled IBD process from Section 4.3. The scaled IBD process in state j exhibits a birth rate $\lambda_j = j(1-a) + aN_*$ and a death rate $\mu_j = j$. We apply these rates to equation (2.28) to calculate the mean extinction time:

$$\begin{aligned}\overline{T} &= \sum_{k=1}^{\infty} \frac{1}{\mu_k} \prod_{j=1}^{k-1} \frac{\lambda_j}{\mu_j} + \sum_{k=1}^{N_*-1} \left(\prod_{j=1}^{k} \frac{\mu_j}{\lambda_j} \right) \sum_{m=k+1}^{\infty} \frac{1}{\mu_m} \prod_{n=1}^{m-1} \frac{\lambda_n}{\mu_n} \\ &= \sum_{k=1}^{\infty} \frac{1}{k} \prod_{j=1}^{k-1} \frac{j(1-a)+aN_*}{j} + \sum_{k=1}^{N_*-1} \left(\prod_{j=1}^{k} \frac{j}{j(1-a)+aN_*} \right) \sum_{m=k+1}^{\infty} \frac{1}{m} \prod_{n=1}^{m-1} \frac{n(1-a)+aN_*}{n} \end{aligned} \tag{D.1}$$

Let us first focus on the first term in this equation. We will analyze the second term later. Using the definition of the Pochhammer symbol for the rising factorial

$$(x)_n = \prod_{i=0}^{n-1} (x+i) \tag{D.2}$$

we rewrite the first term $\overline{T}_1$ of equation (D.1) to

$$\begin{aligned}\overline{T}_1 &= \sum_{k=1}^{\infty} \frac{(1-a)^{k-1}}{k!} \left(1 + \frac{aN_*}{1-a}\right)_{k-1} \\ &= \sum_{k=1}^{\infty} \frac{(1-a)^k}{k! aN_*} \left(\frac{aN_*}{1-a}\right)_k \end{aligned} \tag{D.3}$$

where we used the identity

$$(x+1)_{n-1} = \frac{\Gamma(x+1+n-1)}{\Gamma(x+1)} = \frac{\Gamma(x+n)}{x\Gamma(x)} = \frac{1}{x}(x)_n \tag{D.4}$$

with the Gamma function $\Gamma(x)$. Furthermore, the Pochhammer symbol may be written as a generalized binomial coefficient [9]

$$\frac{(x)_n}{n!} = \frac{(x+n-1)(x+n-2)(x+n-3)\cdots x}{n(n-1)(n-2)\cdots 1} = \begin{pmatrix} x+n-1 \\ n \end{pmatrix} \tag{D.5}$$

for any $x \in \mathbb{R}$ so that we obtain

$$\overline{T}_1 = \frac{1}{aN_*} \sum_{k=1}^{\infty} (1-a)^k \begin{pmatrix} \frac{aN_*}{1-a} + k - 1 \\ k \end{pmatrix}. \tag{D.6}$$

The sum is evaluated using the binomial series [9]

$$\sum_{n=0}^{\infty} \begin{pmatrix} x+n \\ n \end{pmatrix} \alpha^n = \frac{1}{(1-\alpha)^{x+1}} \tag{D.7}$$

defined for $x \in \mathbb{R}$ and $|\alpha| < 1$ so that the first term finally becomes

$$\overline{T}_1 = \frac{1}{aN_*} \left(\left[(1-(1-a))^{\frac{aN_*}{1-a}-1+1} \right]^{-1} - 1 \right) = \frac{a^{-\frac{aN_*}{1-a}} - 1}{aN_*}. \tag{D.8}$$

The second term $\overline{T}_2$ of equation (D.1) becomes

$$\begin{aligned} \overline{T}_2 &= \sum_{k=1}^{N_*-1} \left(\frac{k!}{(1-a)^k} \prod_{j=1}^{k} \left[j + \frac{aN_*}{1-a} \right]^{-1} \right) \sum_{m=k+1}^{\infty} \frac{(1-a)^{m-1}}{m!} \prod_{n=1}^{m-1} \left[n + \frac{aN_*}{1-a} \right] \\ &= \sum_{k=1}^{N_*-1} \left(\frac{k!}{(1-a)^k \left(\frac{aN_*}{1-a} \right)_{k+1}} \right) \sum_{m=k+1}^{\infty} \frac{(1-a)^{m-1}}{m!} \left(\frac{aN_*}{1-a} \right)_m \end{aligned} \tag{D.9}$$

We now introduce the hypergeometric sum

$${}_2F_1(a,b,c;z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \cdot \frac{z^k}{k!} \tag{D.10}$$

which converges absolutely for $|z| < 1$ [9], so that we may rewrite the above formula for $\overline{T}_2$ to

$$\begin{aligned} \overline{T}_2 &= \sum_{k=1}^{N_*-1} \frac{k!}{(1-a)^k \left(\frac{aN_*}{1-a} \right)_{k+1}} \frac{(1-a)^k}{(k+1)!} \left(\frac{aN_*}{1-a} \right)_{k+1} \cdot {}_2F_1 \left(1, 1+k+\frac{aN_*}{1-a}, 2+k, 1-a \right) \\ &= \sum_{k=1}^{N_*-1} \frac{1}{k+1} \cdot {}_2F_1 \left(1, 1+k+\frac{aN_*}{1-a}, 2+k, 1-a \right) \end{aligned} \tag{D.11}$$

which we could not evaluate analytically. However, studying the single terms for $k = 1, 2, \ldots$ using MATHEMATICA, we find that each term is of the order of $\mathcal{O}\left((aN_*)^{-k-1}\right)$. Thus, for $aN_* \gg 1$ in first order approximation $\overline{T}_2$ may be neglected in comparison to $\overline{T}_1$ and we find, that

$$\overline{T} = \frac{a^{-\frac{aN_*}{1-a}} - 1}{aN_*} + \mathcal{O}\left((aN_*)^{-2}\right) \tag{D.12}$$

where the second term is always positive as all terms in equation (D.11) are positive.

Bibliography

[1] T. Antal, M. A. Nowak, and A. Traulsen, *Strategy abundance in 2x2 games for arbitrary mutation rates*, J. Theor. Biol. **257**, 340–344 (2009).

[2] H. Arnoldt, M. Timme, and S. Grosskinsky, *Frequency dependent fitness induces multistability in coevolutionary dynamics*, J. R. Soc. Interface **9**, 3387–3396 (2012).

[3] M. Assaf and M. Mobilia, *Large fluctuations and fixation in evolutionary games*, J. Stat. Mech., P09009 (2010).

[4] P. Bak and K. Sneppen, *Punctuated equilibrium and criticality in a simple model of evolution*, Phys. Rev. Lett. **71**, 4083–4086 (1993).

[5] A. A. Berryman, *Population cycles*, Oxford University Press, New York, USA, 2002.

[6] A. J. Black and A. J. McKane, *Stochastic formulation of ecological models and their applications*, Trends Ecol. Evol. **27**, 337–345 (2012).

[7] M. C. Boerlijst, S. Bonhoeffer, and M. A. Nowak, *Viral quasi-species and recombination*, Proc. R. Soc. Lond. B **263**, 1577–1584 (1996).

[8] C. Briones, E. Domingo, and C. Molina-Paris, *Memory in retroviral quasispecies: Experimental evidence and theoretical model for human immunodeficiency virus*, J. Mol. Biol. **331**, 213–229 (2003).

[9] I. N. Bronstein, K. A. Semendjajew, G. Musiol, and H. Mühlig, *Taschenbuch der Mathematik*, Wissenschaftlicher Verlag Harri Deutsch GmbH, Frankfurt, Germany, 2008.

[10] C. F. Camerer, *Behavioral game theory*, Princeton University Press, Princeton, USA, 2003.

[11] F. D. Ciccarelli, T. Doerks, C. von Mering, C. J. Creevey, B. Snel, and P. Bork, *Toward automatic reconstruction of a highly resolved tree of life*, Science **311**, 1283–1287 (2006).

[12] F. M. Cohan, M. S. Roberts, and E. C. King, *The potential for genetic exchange by transformation within a natural population of Bacilus subtilis*, Evolution **45**, 1383–1421 (1991).

[13] M. J. Crawley, *Natural enimies: The population biology of predators, parasites and diseases*, Blackwell Scientific Publications, Oxford, UK, 1992.

[14] T. Dagan and W. Martin, *The tree of one percent*, Genome Biol. **7**, 118 (2006).

[15] C. R. Darwin, *On the origin of species*, John Murray, London, UK, 1859.

[16] H. de Vries, *Die Mutationstheorie. Versuche und Beobachtungen über die Entstehung von Arten im Pflanzenreich*, Veit, Leipzig, Germany, 1901.

[17] W. F. Doolittle, *Phylogenetic classification and the universal tree*, Science **284**, 2124–2128 (1999).

[18] W. F. Doolittle, *Uprooting the tree of life*, Sci. Am. **February**, 90–95 (2000).

[19] P. Doreian and F. N. Stokman, *Evolution of social networks*, Gordon and Breach Publishers, Amsterdam, The Netherlands, 1997.

[20] B. Drossel, *Biological evolution and statistical physics*, Adv. Phys. **50**, 209–295 (2001).

[21] B. Drossel and A. J. McKane, *Modelling food webs*, Handbook of Graphs and Networks, Wiley-VCH, Weinheim, Germany, 2003.

[22] R. Durrett and D. Schmidt, *Waiting for two mutations: With applications to regulatory sequence evolution and the limits of Darwinian evolution*, Genetics **180**, 1501–1509 (2008).

[23] M. Eigen, *Selforganization of matter and the evolution of biological macromolecules*, Die Naturwissenschaften **58**, 465–523 (1971).

[24] M. Eigen and P. Schuster, *The hypercycle: A principle of natural self-organization*, Springer-Verlag, Berlin, Germany, 1979.

[25] S. F. Elena, L. Ekunewe, N. Hajela, S. A. Oden, and R. E. Lenski, *Distribution of fitness effects caused by random insertion mutations in escherichia coli*, Genetica **102/103**, 349–358 (1998).

[26] R. A. Fisher, *The general theory of natural selection*, Clarendon Press, Oxford, UK, 1930.

[27] S. J. Freeland and L. D. Hurst, *The genetic code is one in a million*, J. Mol. Evol. **47**, 238–248 (1998).

[28] D. J. Futuyma, *Evolution*, Sinauer Associates, Sunderland, USA, 2005.

[29] C. Gardiner, *Stochastic methods*, Springer-Verlag, Heidelberg, Germany, 2009.

[30] M. D. Giulio, *The origin of the genetic code: Theories and their relationships, a review*, Biosystems **80**, 175–184 (2005).

[31] S. J. Gould and N. Eldredge, *Punctuated equilibrium comes of age*, Nature **366**, 223–227 (1993).

[32] D. Haig and L. D. Hurst, *A quantitative measure of error minimization in the genetic code*, J. Mol. Evol. **33**, 412–417 (1991).

[33] P. Hänggi, P. Talkner, and M. Borkovec, *Reaction rate theory: Fifty years after Kramers*, Rev. Mod. Phys. **62**, 251–342 (1990).

[34] J. Hirsch, *Behavior genetics and individuality understood*, Science **142**, 1436–1442 (1963).

[35] J. Hofbauer and K. Sigmund, *Evolutionary games and population dynamics*, Cambridge University Press, Cambridge, UK, 1998.

[36] M. N. Jacobi and M. Nordahl, *Quasispecies and recombination*, Theor. Popul. Biol. **70**, 479–485 (2006).

[37] R. Jain, M. C. Rivera, and J. A. Lake, *Horizontal gene transfer among genomes: The complexity hypothesis*, Proc. Natl. Acad. Sci. U. S. A. **96**, 3801–3806 (1999).

[38] S. Karlin and H. M. Taylor, *A first course in stochastic processes*, Academic Press, New York, USA, 1975.

[39] M. Kimura, *Diffusion models in population genetics*, J. Appl. Probab. **1**, 177–232 (1964).

[40] M. Kimura, *Evolutionary rate at the molecular level*, Nature **217**, 624–626 (1968).

[41] M. Kimura, *The neutral theory of molecular evolution*, Cambridge University Press, Cambridge, UK, 1983.

[42] M. Kimura and G. H. Weiss, *The stepping stone model of population structure and the decrease of genetic correlation with distance*, Genetics **49**, 561–576 (1964).

[43] K. S. Korolev, M. Avlund, O. Hallatschek, and D. R. Nelson, *Genetic demixing and evolution in linear stepping stone models*, Rev. Mod. Phys. **82**, 1691–1718 (2010).

[44] H. A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica **7**, 284–304 (1940).

[45] C. G. Kurland, B. Canback, and O. G. Berg, *Horizontal gene transfer: A critical view*, Proc. Natl. Acad. Sci. U. S. A. **100**, 9658–9662 (2003).

[46] J. G. Lawrence, *Gene transfer in bacteria: Speciation without species*, Theor. Popul. Biol. **61**, 449–460 (2002).

[47] M. Leisner, J.-T. Kuhr, J. O. Rädler, E. Frey, and B. Maier, *Kinetics of genetic switching into the state of bacterial competence*, Biophys. J. **96**, 1178–1188 (2009).

[48] M. Leisner, K. Stingl, E. Frey, and B. Maier, *Stochastic switching to competence*, Curr. Opin. Microbiol. **11**, 553–559 (2008).

[49] M. Luther, *Die Bibel*, Deutsche Bibelgesellschaft, Stuttgart, Germany, 1985.

[50] R. M. May and W. J. Leonard, *Nonlinear aspects of competition between three species*, SIAM J. Appl. Math. **29**, 243–253 (1975).

[51] D. E. McCauley and M. T. Brock, *Frequency-dependent fitness in Silene Vulgaris, a gynodioecious plant*, Evolution **52**, 30–36 (1998).

[52] J. O. McInerney, J. A. Cotton, and D. Pisani, *The procaryotic tree of life: Past, present... future?*, Trends Ecol. Evol. **23**, 276–281 (2008).

[53] A. J. McKane and T. J. Newman, *Predator-prey cycles from resonant amplification of demographic stochasticity*, Phys. Rev. Lett. **94**, 218102 (2005).

[54] G. Mendel, *Versuche über Pflanzen-Hybriden*, Verhandlungen des naturforschenden Vereines in Brünn **4**, 3–47 (1866).

[55] P. A. P. Moran, *The statistical processes of evolutionary theory*, Clarendon Press, Oxford, UK, 1962.

[56] J. D. Murray, *Mathematical biology*, Springer-Verlag, Heidelberg, Germany, 1989.

[57] Y. Nakamura, T. Itoh, H. Matsuda, and T. Gojobori, *Biased biological functions of horizontally transferred genes in prokaryotic genomes*, Nature Genet. **36**, 760–766 (2004).

[58] S. Nee, *Birth-death models in macroevolution*, Annu. Rev. Ecol. Evol. Syst. **37**, 1–17 (2006).

[59] C. Neuhauser, *Mathematical challenges in spatial ecology*, Not. Am. Math. Soc. **48**, 1304–1314 (2001).

[60] M. E. J. Newman and R. G. Palmer, *Modeling extinction*, Oxford University Press, Oxford, UK, 2003.

[61] R. M. Nisbet and W. S. C. Gurney, *A simple mechanism for population cycles*, Nature **263**, 319–320 (1976).

[62] J. R. Norris, *Markov chains*, Cambridge University Press, Cambridge, UK, 1997.

[63] M. A. Nowak, *Evolutionary dynamics*, Harvard University Press, London, UK, 2006.

[64] M. A. Nowak, N. L. Komarova, and P. Niyogi, *Evolution of universal grammar*, Science **291**, 114–118 (2001).

[65] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg, *Emergence of cooperation and evolutionary stability in finite populations*, Nature **428**, 646–650 (2004).

[66] H. A. Orr, *The genetic theory of adaptation: A brief history*, Nat. Rev. Genet. **6**, 119–127 (2005).

[67] M. Paczuski, S. Maslov, and P. Bak, *Avalanche dynamics in evolution, growth and depinning models*, Phys. Rev. E **53**, 414–443 (1996).

[68] J.-M. Park and M. W. Deem, *Phase diagrams of quasispecies theory with recombination and horizontal gene transfer*, Theor. Popul. Biol. **70**, 479–485 (2006).

[69] H. Philippe and C. J. Douady, *Horizontal gene transfer and phylogenetics*, Curr. Opin. Microbiol. **6**, 498–505 (2003).

[70] M. Pineda-Krch, H. J. Blok, U. Dieckmann, and M. Doebeli, *A tale of two cycles – distinguishing quasi-cycles and limit cycles in finite predator-prey populations*, Oikos **116**, 53–64 (2007).

[71] J. E. Pool and R. Nielsen, *Population size changes reshape genomic patterns of diversity*, Evolution **61**, 3001–3006 (2007).

[72] D. M. Raup, *Biological extinction in earth history*, Science **231**, 1528–1533 (1986).

[73] Y. Raz and E. Tannenbaum, *The influence of horizontal gene transfer on the mean fitness of unicellular populations in static environments*, Genetics **185**, 327–337 (2010).

[74] H. Risken, *The Fokker-Planck equation*, Springer-Verlag, Heidelberg, Germany, 1996.

[75] M. J. S. Rudwick, *The meaning of fossils: Epsiodes in the history of palaeontology*, The University of Chicago Press, Chicago, USA, 1972.

[76] R. Sanjuan, A. Moya, and S. F. Elena, *The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus*, Proc. Natl. Acad. Sci. U. S. A. **101**, 8396–8401 (2004).

[77] A. Sasaki and M. A. Nowak, *Mutation landscapes*, J. Theor. Biol. **224**, 241–247 (2003).

[78] M. Singer and P. Berg, *Genes and genomes*, University Science Books, Mill Valley, California, USA, 1991.

[79] J. M. Smith and G. R. Price, *The logic of animal conflict*, Nature **246**, 15–18 (1973).

[80] J. R. Stone, *The evolution of ideas: A phylogeny of shell models*, Am. Nat. **148**, 904–929 (1996).

[81] M. Syvanen, *Cross-species gene transfer; implications for a new theory of evolution*, J. Theor. Biol. **112**, 333–343 (1985).

[82] M. Syvanen, *Horizontal gene transfer: Evidence and possible consequences*, Annu. Rev. Genet. **28**, 237–261 (1994).

[83] F. Tajima, *The effect of change in population size on DNA polymorphism*, Genetics **123**, 597–601 (1989).

[84] C. Taylor, D. Fudenberg, A. Sasaki, and M. A. Nowak, *Evolutionary game theory in finite populations*, Bull. Math. Biol. **66**, 1621–1644 (2004).

[85] P. D. Taylor and L. B. Jonker, *Evolutionary stable strategies and game dynamics*, Math. Biosci. **40**, 145–156 (1978).

[86] C. M. Thomas and K. M. Nielsen, *Mechanisms of, and barriers to, horizontal gene transfer between bacteria*, Nat. Rev. Microbiol. **3**, 711–721 (2005).

[87] A. Traulsen, J. C. Claussen, and C. Hauert, *Coevolutionary dynamics: From finite to infinite populations*, Phys. Rev. Lett. **95**, 238701 (2005).

[88] A. Traulsen, J. C. Claussen, and C. Hauert, *Coevolutionary dynamics in large, but finite populations*, Phys. Rev. E **74**, 011901 (2006).

[89] A. Traulsen, J. M. Pacheco, and L. A. Imhof, *Stochasticity and evolutionary stability*, Phys. Rev. E **74**, 021905 (2006).

[90] K. Uchiyama, *The first hitting time of a single point for random walks*, Electron. J. Probab. **16**, 1960–2000 (2011).

[91] K. Uchiyama, *One dimensional lattice random walks with absorption at a point/on a half line*, J. Math. Soc. Jpn. **63**, 675–713 (2011).

[92] A. A. Vogan and P. G. Higgs, *The advantages and disadvantages of horizontal gene transfer and the emergence of the first species*, Biol. Direct **6**, 1 (2011).

[93] V. Volterra, *Fluctuations in the abundance of a species considered mathematically*, Nature **118**, 558–560 (1926).

[94] C. O. Wilke, C. Ronnewinkel, and T. Martinez, *Dynamic fitness landscapes in molecular evolution*, Physics Reports **349**, 395–446 (2001).

[95] R. J. Williams and N. D. Martinez, *Simple rules yield complex food webs*, Nature **404**, 180–183 (2000).

[96] C. R. Woese, *Interpreting the universal phylogenetic tree*, Proc. Natl. Acad. Sci. U. S. A. **97**, 8392–8396 (2000).

[97] C. R. Woese, *On the evolution of cells*, Proc. Natl. Acad. Sci. U. S. A. **99**, 8742–8747 (2002).

[98] S. Wright, *The evolution of dominance*, The American Naturalist **63**, 556–561 (1929).

[99] S. Wright, *Evolution in mendelian populations*, Genetics **16**, 97–159 (1931).

[100] B. Wu, C. S. Gokhale, L. Wang, and A. Traulsen, *How small are small mutation rates?*, J. Math. Biol. **64**, 803–827 (2012).

[101] D. A. Young, *The contemporary relevance of Augustine*, Perspectives on science and Christian faith **40**, 42–45 (1988).

Acknowledgements

Here, I would like to express my gratitude towards the many people who have supported me during the development of this thesis. My first thanks go to Marc Timme for providing me with an excellent supervision, at the same time giving me the freeness to research what I found interesting and handing out invaluable advice whenever needed. Also, thank you for letting me be part of the Network Dynamics Group with its wonderful working atmosphere. I also thank Theo Geisel for welcoming me in the Department for Nonlinear Dynamics as well and for agreeing to referee this thesis.

Furthermore, I would like to thank Stefan Grosskinsky and Stephan Eule for helpful discussions on stochastic systems and Oskar Hallatschek and all the members of his group for lots of input on evolutionary systems. Thank you very much for your constant support.

My special thanks go to the members of the Network Dynamics Group for providing a relaxed working atmosphere and for so many discussions on and off topic. Looking back, we spent so much time together which was always lots of fun. I think of retreats, table tennis, cakes, barbecues and much more. In particular, I want to thank Martin Rohden for bearing the burden of being my eternal office mate and Christian Bick for awesome comments on mathematics, style and more.

Johannes Klinglmayr gave me the opportunity to visit the Mobile Systems group in Klagenfurt. Thank you for the idea of this cooperation, the wonderful time in Klagenfurt and many nice evenings in Klagenfurt, Göttingen and Sestri Levante.

I want to thank our technical staff. Without you my work would not have been possible. In particular, I thank Sven Jahnke for fast and uncomplaining responses whenever a technical problem arose. Likewise, I thank our administrative staff, especially Barbara, for constant support on the endless fight with bureaucracy. Without you, no one could even think about writing a thesis.

For me, a good work-life balance is the basis to find inspiration for research and in the end for the motivation to write a thesis. I am indebted to my friends for providing so many nice ways of distraction. Thank you for your hospitality, drinks, bycicle tours, games, parties, holidays, ... – and thank you for the music.

Last, but not least, I would like to thank my family for their unconditional support in every respect. And naturally, thank you, Julia. For all, and for letting me be Mr. Arnoldt.