Quantifying Evolutionary Dynamics

Dissertation
for the award of the degree
“Doctor rerum naturalium”
of the Georg–August–Universität Göttingen

in the program
International Max Planck Research School
“Physics of Biological and Complex Systems”
(IMPRS PBCS)
of the Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences
(GGNB)

submitted by

Lukas Geyrhofer

from St. Pölten

Göttingen, 2014

Thesis committee

Prof. Dr. Oskar Hallatschek, Biological Physics and Evolutionary Dynamics Group, Max Planck Institute for Dynamics and Self-Organization

Prof. Dr. Annette Zippelius, Institute for Theoretical Physics, Georg-August-Universität Göttingen

Prof. Dr. Burkhard Morgenstern, Institute of Microbiology and Genetics, Department of Bioinformatics, Georg-August-Universität Göttingen

Thesis referees

Prof. Dr. Oskar Hallatschek, Biological Physics and Evolutionary Dynamics Group, Max Planck Institute for Dynamics and Self-Organization

Prof. Dr. Annette Zippelius, Institute for Theoretical Physics, Georg-August-Universität Göttingen

Examination committee

Prof. Dr. Oskar Hallatschek, Biological Physics and Evolutionary Dynamics Group, Max Planck Institute for Dynamics and Self-Organization

Prof. Dr. Annette Zippelius, Institute for Theoretical Physics, Georg-August-Universität Göttingen

Prof. Dr. Burkhard Morgenstern, Institute of Microbiology and Genetics, Department of Bioinformatics, Georg-August-Universität Göttingen

Prof. Dr. Jürgen Vollmer, Department of Complex Fluids, Max Planck Institute for Dynamics and Self-Organization

Eleni Katifori, PhD, Physics of Biological Organization, Max Planck Institute for Dynamics and Self-Organization

Prof. Dr. Marc Timme, Network Dynamics Group, Max Planck Institute for Dynamics and Self-Organization

Examination date

June 24th, 2014

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

Göttingen, 2014

Contents

Abstract 7

1. Introduction 9
  1.1. Historical overview of experimental evolution 9
  1.2. Microbes as model organisms in experimental evolution 10
  1.3. Modeling adaptation in microbial populations 16
  1.4. Quantifying evolutionary dynamics 22

2. Simulations of models of adaptation 23
  2.1. Simulation model 23
  2.2. Simulation results 30

3. Constrained BRWs as minimal models for adaptation in asexual populations 35
  3.1. Traveling wave models 35
  3.2. Tuned models - employing a special population size constraint 40
  3.3. Interpretation of the tuned constraint 45
  3.4. Interpretation of tuned models 56
  3.5. The distribution of fixating clones 57
  3.6. Higher moments in tuned models 61

4. Simulations of tuned models 67
  4.1. Stochastic simulation of tuned models 67
  4.2. Fluctuations in population size 68

5. Oscillations in traveling wave models of adaptation 77
  5.1. Introduction 77
  5.2. Time-dependent dynamics of BRWs 78
  5.3. Fitness measurements in tuned models 85
  5.4. Experimental evidence 86

6. Applications of tuned models: oasis model 93
  6.1. Motivation 94
  6.2. Simulations 95
  6.3. Analytical approach 100
  6.4. Discussion 103

7. Discussion 105
7.1. Traveling waves are simplified models of adaptation 105
7.2. Tuned models exhibit exactly solvable dynamics 107
7.3. Fixation probabilities help to discern effects of clonal competition 109
7.4. Fluctuations in the adaptation process 114
7.5. Experimental evolution revisited in the light of theoretical results 118
7.6. Concluding remarks 119

Bibliography 120

Appendix 132

A. Simulation code 133

B. Numerical methods 135
B.1. Multi-dimensional Newton-Raphson iteration to solve the non-linear constraint equation 135
B.2. Numerical methods for the mean stationary population density 137
B.3. Solving the 2-point-correlation-function 139

C. Analytical addendum 141
C.1. Integral transforms for exponential mutation kernels 141
C.2. Fixation probability for moment closures at higher levels 143

D. Oscillations compared to fitness scales in tuned models 149
D.1. Fitness distributions in traveling wave models 149
D.2. Comparing numerical profiles to relevant timescales 150
D.3. Evaluation of fitness scales 153

E. Experimental data 155
E.1. Additional tests to examine experimental data 155

Acknowledgments 157

CV 158

Abstract

Adaptation is a central process in evolution: populations adapt to their environment by accumulating beneficial mutations over the course of generations. Microbial populations present an excellent experimental system that allows to observe the adaptation process on reasonable timescales. Recent results of evolutionary experiments with microbes expose an intricate dynamics: beneficial mutations appear continuously and often many of them are present simultaneously. These different clones interfere on their way to fixation or extinction in the mostly asexually reproducing populations. This dynamical regime, termed clonal competition, has been observed in various laboratory populations.

In recent years traveling wave models gained considerable attention as possible description of the ensuing dynamics in the experimental microbial populations. There, the population is represented as a density in fitness space. Individuals are grouped by their growth rate. The size of these groups, called fitness classes, changes over time via mutation and selection. In many implementations genetic drift is considered to correctly account for the dynamics of very fit clones that are still low in number. If the influx of beneficial mutations is large enough, the bell-shaped density moves to higher fitness as soliton-like traveling wave: the population adapts.

In this thesis, we investigate the dynamical behavior of these traveling wave models. In particular, we try to quantify the fluctuations in the adaptation process that are inherent due to the discreteness of individuals and individual mutations. Even very fit clones can go extinct due to genetic drift. Successful clones sweep to fixation and eradicate much of the genetic diversity. In order to describe these fluctuations, we use a special version of traveling wave models. Based on branching random walks, we couple the dynamics to a specific tuned constraint that allows to construct exactly solvable equations. As a particularly interesting extension, we show that slightly modifying this formalism yields all moments of the population density, giving access to the complete fluctuation spectrum.

Based on simulation data, we extract relevant timescales of the adaptation process in our model. An almost universal scaling is indicated: the time of large scale oscillations in fitness variance appears to depend largely on the product of adaptation speed and mutational scale alone. When applied to experimental data, we can infer the mutational scale, a value that is usually obscured by the dynamics of interfering mutations.

1. Introduction

Among the first evidence of life on earth are microfossils, some of them dated to about 3.7 to 3.5 billion years ago [WKS+11, OKI+14]. Since the origin of life, organisms have conquered almost every niche on this planet. During that process, life shaped our planet as we know it today.

Evolution is an inherently dynamical process. Populations experience a constant turn-over of individuals and on larger timescales populations grow, shrink and go extinct. Variations in the heritable information, *i.e.* information that can be passed to the offspring, can affect phenotypic traits of these individuals. In turn, this can alter their reproductive success. Over generations, variations accumulate in populations.

Adaptation is one of the central processes in evolution: a particularly well suited trait to deal with environmental conditions can emerge and natural selection favors the individuals featuring such traits. Then, the trait can spread through the population. Over time the population adapts.

In this thesis, we focus on a specific scenario of adaptation: microbial populations that evolve under laboratory conditions find their environment virtually unchanged over long periods of time. This allows them to adapt to these conditions. In the following, we first present a brief historical overview of evolutionary experiments, before going into details concerning microbes in section 1.2. Based on insights from these experiments, we will turn our attention to modeling adaptation, which is the main topic of this thesis.

1.1. Historical overview of experimental evolution

Humans have been conducting evolutionary "experiments" for several millenia: the beginning of the domestication of plants (and animals) is known as the neolithic revolution. Dogs have been among the first animals that were domesticated. Sources vary about the date and it seems that the dog has been domesticated several times. Earliest proposed dates are around 40000 BCE [GSS+09]. Other animals, like sheep, pigs, goats and cows followed at different places around the globe. The first precursors of plant domestication have been dated to about 9500 BCE [HHM+01]. Cereal crops appear to have been domesticated around 7000 BCE in the Fertile Crescent [LYGA00]. Domestication is insofar an evolutionary experiment, as humans select continually and controlled for favorable traits among individuals by breeding or cultivating this particular lineage over many generations. Darwin's thinking was influenced by the observation that domestication is evidence for evolution: he started his "Origin of species" with a chapter on this very same topic [Dar59].

Despite these early "experiments", dealing with evolution has been largely a "historical" science since the emergence of the concept in the 19th century: only the contemporary end-product of a lineage of descendants can be investigated and described. Sometimes, the fossil record helps to understand the different steps two lineages have taken while separating from each other. Unfortunately, the existence of intermediate steps in the fossil record is often scarce[1]. Experimental evolution on the other hand, tries to observe and quantify patterns in populations over many generations with organisms that reproduce reasonably fast.

Some of the first "modern" quantitative experiments on heredity and evolution were conducted by Gregor Mendel. He hybridized peas, and scored the resulting traits [Men66]. His results have not been recognized by the scientific community until they were rediscovered independently in 1900 by different scientists [Hen00]. Mendel's results were a major constituent of evolution's *Modern Synthesis* in the first half of the 20th century. The Modern Synthesis combined Darwin's theory of natural selection with knowledge about genetics at that time [Hux42].

Experimental evolutionary research in the 20th century used often fruitflies of the *Drosophila* genus as model organisms. These flies reproduce reasonably fast, with a generation time of roughly 10 days. Thus, large stocks of flies can be bred simultaneously in the lab. Many population genetic models were tested using these organisms [Gil10]. Nowadays, fruitflies still remain common model organisms in many areas of biology.

The first attempts to use microbes in evolutionary experiments date back the late 19th century. William Dallinger evolved unicellular organisms over the course of several years to grow at progressively elevated temperatures [Dal87]. Other scholars also proposed the usage of microbes in evolutionary experiments around that time [Rus93a, Rus93b]. In the 20th century microbes were used in many biological experiments, but rarely to test evolutionary hypotheses. A lot of what we know about molecular biology today was found first using microbes as model organisms [Zim12], by inactivating some genes to decipher cellular processes. Experimental evolution, on the other hand, allows to investigate gains of function instead of functional loss [EL03]. Only in the 1990s, when Richard Lenski's experiment with *Escherichia coli* [Len15] showed first promising results and became well known, did scientists started to use commonly microbes as model organisms in experimental evolution.

These microbial experiments are the main setup we want to describe with our theoretical approach. In the next section, we summarize results of microbial evolution experiments obtained by several different groups over the last two decades.

1.2. Microbes as model organisms in experimental evolution

The term *microbes* is used to refer to microscopic organisms, such as bacteria and archea, as well as unicellular eukaryotes, *e.g.* algae and fungi. Occasionally, even viruses are included. Microbes can exhibit different features: eukaryotes can be single-celled or form multicellular aggregates [RDBT12, RHH+13] and bacteria can also live in larger, interacting colonies as biofilms [LVK10]. Some single-celled eukaryotes can perform meiosis and mate in addition to multiplying asexually via mitosis. Some bacteria encode proteins that allow them to exchange genetic material in ”horizontal gene transfers”. However, for the sake of simplicity we will consider single-celled mitotically reproducing microbes in our models.

---

[1] Even the mere existence of intermediate fossils is doubted [Gou02]. We will discuss the accompanying theory of *punctuated equilibrium* in section 7.4.3.

Microbes have several advantages when used in evolutionary experiments, as compared to other organisms. Several recent reviews deal with microbial evolution experiments [EL03, BMBC09, Bur12, KLE+12, Kus13]. Here, we will list some of these advantages. We are mainly concerned in modeling the growth rate of populations, hence our review of literature will be biased in that direction.

Microbes are important model organisms in *quantitative* experimental evolution, because they usually need little laboratory space. Moreover, their generation time is short, ranging from approximately 20 minutes for the bacterium *Escherichia coli*, and up to 90 minutes for the yeast *Saccharomyces cerevisiae*. Both are among the best characterized organisms so far in biology. Hence, experiments with a duration of several hundreds up to a few thousands of generations can be managed in a normal lab [Des13b]. Because microbes use “little” time and “little” space *replicate experiments* are possible. With enough replicates, quantitative statements on evolution can then be made.

Another important feature of evolutionary experiments in the laboratory is that environmental conditions can be measured and controlled. Experiments can be run for long times with well defined conditions. Often, microbes do not need complex environments and few chemical compounds suffice for their growth. In turn, this simplicity can also be used to introduce additional features, that allows to test basic ecological interactions [MNNM14, WPH+14].

Asexual reproduction allows to define a lineage as all the descendants of one cell. When a mutation occurs in one individual all the descendants of this individual will inherit this particular mutation[2]. This simplifies the ensuing reconstruction of genetic kinship. Previously, investigators relied on specific, discrete genetic markers to track the evolution of each lineage. However, for the last 10 years DNA sequencing has becoming more and more affordable, both time-wise and money-wise. Deep sequencing of several whole genomes within a population, often at multiple time points of the experiment, is almost a standard technique today [BYY+09, LRH+13, HD13, KRJD14]. The small genome size of microbes, compared to many higher organisms, contributes to the value of genomic sequencing in evolutionary research.

Moreover, microbes can be stored deep-frozen (at −80°C). When thawed, cells regain functionality and grow again. This allows to keep a “fossil record” of the complete experiment. After an interesting observation has been made, the occurrence and trajectory of a particular mutation can be tracked back to its origin [BBL08, BBDL12]. Evolution can also be replayed with similar starting conditions, allowing to replicate experiments, necessary for quantitative statements.

---

[2] Note that recombination and horizontal gene transfer mentioned above can introduce additional complications. On timescales of experiments this could possibly happen, but is neglected in many theoretical approaches.

Further advantages, also with respect to specific experiments, are listed in the review articles [EL03, BMBC09, Bur12, KLE+12, Kus13].

1.2.1. Experimental observations

Here, we describe recurrent observations that come out of many microbial evolution experiments, independent of the particular laboratory or even particular (single-celled) organism. Nevertheless, most of the results have been obtained with various strains of *E. coli* and *S. cerevisiae*. Details of one particular relevant experimental setup will be explained afterwards, in section 1.2.2.

The most striking observation is that populations *adapt fast* to their laboratory conditions: after only 2000 generations their growth rate can increase by 25% or more [Len15] (see Figure 1.1A below). Following the publication of Kimura's book on the *neutral theory of molecular evolution* [Kim84] most mutations were considered neutral. Deleterious mutations that reduced the reproductive capacity of the cell were thought to be less common and even less common than deleterious mutations were beneficial ones that actually increase the fitness of the microbe. Still, beneficial mutations appear to be prevalent enough that the mean fitness of the population increases. Those mutations survive genetic drift and eventually reach fixation, *i.e.* 100% frequency, in the population.

In experiments that run long enough, *diminishing returns* are observed [VZG+99, CCD+11, WRL13b, KRJD14]: adaptation is fastest at the beginning of the experiment, when the microbes experience a new environment, then decelerates.

Reported mutation rates of beneficial mutations are sometimes as high as 10^{-4} per generation per genome, see [Des13b] and references therein. In a typical microbial population of size $\gtrsim 10^6$, often several mutations arise in the population per generation. Lineages with these different mutations compete for fixation: only one of them can finally take over a population. This effect has been termed *clonal competition* (or also *clonal interference*[3]) [GL98]. Clonal competition has been observed experimentally in various setups with microbes [RDVG02, LBD11, LRH+13, BYY+09, KRJD14].

In some cases, clonal competition can lead to stable coexistence of different strains within the same population [BBL08, LGPH+12, HD13, PHLG+14]. This occurs when mutations drive sub-populations into slightly different ecological niches, still within the boundaries set by the experiment. Then, each of those sub-populations can often be characterized by the fact that they utilize parts of the provided resources better than

---

[3] The distinction between "clonal interference" and "clonal competition" is only technical. Among the first theoretical approaches to describe this effect has been the publication [GL98]. There, Gerrish and Lenski coined the term "clonal interference". They considered the simultaneous occurrence of multiple beneficial mutations that interfered for fixation. Only *after* the fixation of one of these mutations, another set of beneficial mutations is allowed to occur. In reality, this process is more involved and the genetic background of these multiple beneficial mutations will receive additional mutations *before* any of them fixates. To distinguish this more general model from the one in [GL98], we call it "clonal competition".

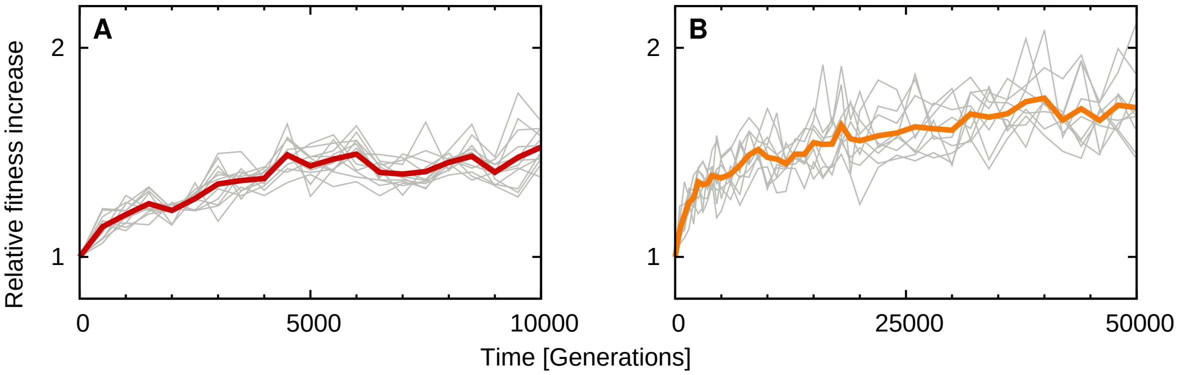

Figure 1.1.: **Fitness increase over time in the Lenski experiment.** Data for (A) are taken from the webpage of the project [Len15], data for (B) are from a recent publication [WRL13b]. The first striking observation, immediately visible, is that populations adapt to their laboratory conditions. Even after 22 years (equivalent to 50000 generations) novel mutations increase the fitness of the 12 strains, although adaptation has speed slowed down already. Gray lines depict the relative fitness of the single populations, colored lines are their average.

the other one. However, more than two different lineages are unlikely to evolve from a single ancestor and continue to coexist [Des13a].

Analyzing the genome of evolved microbial populations can lead to more detailed insights on the adaptation process. First, the scenario of clonal competition can be observed directly when sequencing multiple clones within a population at multiple time points. Evidence suggests that mutations sweep in *cohorts*, *i.e.* functionally unrelated mutations are linked together on the same genetic background [LRH+13]. Often, the fitness advantage of a single mutation is not enough to trigger a sweep, especially if multiple mutations with similar fitness effects are present. The exact cause(s) of the fitness increase can be determined by genomic analyses. For instance, the effect of a mutation on fitness can be directly measured when inserted into an isogenic construct [BYY+09].

Moreover, DNA sequencing has revealed another intriguing fact. Some genes have been found to mutate independently in multiple replicate experiments [BBL08, BYY+09, KRJD14]. Thus, convergent evolution seems pervasive, at least at gene scale [TRVG+12]. However, when directly investigating the DNA sequence, it becomes evident that these genes are often mutated at different positions: on the sequence level, evolution appears highly stochastic [KRJD14].

1.2.2. The Lenski experiment

The *long-time evolution experiment* under the supervision of Richard Lenski [Len15] has produced massive data on experimental evolution. Many results cited in the last section arose in the context of this experiment. Here, we present additional fascinating details, that (so far) only occurred there. Today, the amount of data produced is orders of magnitude larger than comparable experiments [Pen13][4].

The experiment started in 1988 and continues up until today[5]. As of writing this thesis (May 2014) more than 60000 generations have passed. A colony of the ancestor cell was separated into twelve different populations that are grown independently in liquid minimal media. Every day, populations are diluted 1:100 into new nutrients. Glucose is the limited carbon source for growth. On average cells divide 6 to 7 times each day, with the amount of glucose supporting a stationary population size of about 10^7 to 10^8 cells [Len15]. Every 500 generations (75 days), a sample of the population is frozen and stored.

The (relative) fitness increase of the twelve populations is shown in Figure 1.1. Initially, adaptation to the "new“ environment was fast, but it decelerated over time. Fitness is measured by direct competition with the original strain: samples from the original and evolved strains are mixed, and plated on agar before and after 24h of growth under standard experimental conditions. Then, the difference in the number of colonies from evolved and original strain before and after their growth yields the fitness as the Malthusian growth parameter [LRST91, WRL13b, Len15], as shown in Figure 1.1.

In six of the twelve populations *mutator* alleles fixed in the population [WRL13b]. These mutations introduced defects in proteins of the DNA repair system, increasing the mutation rate up to a factor of 140 [WBT⁺13]. However, these mutations did not only occur, they fixated in half of the populations. Apparently, mutator cells have a considerable advantage over non-mutator ones. This might indicate that enough beneficial mutations exist. If deleterious mutations are too abundant an effect of *genetic meltdown* would set in, as considered in many theoretical studies [LBBG93]. It has also been reported that at least one of the hypermutator populations subsequently experienced fixation of additional mutations that reduced the mutation rate again [WBT⁺13]. However, the resulting mutation rate was still higher than the original rate.

One particularly remarkable event was the evolution of a key innovation in one of the twelve populations [BBL08]. A duplication close to a gene for an ancient citrate-succinate transporter increased its expression [BBDL12]. Citrate was used as chelating agent in the minimal media, and was therefore present in abundance. In turn, those cells were not limited by glucose as their single carbon source for growth. With the citrate transporter in their membrane, the microbes could take up citrate and use it as additional nutrient after glucose has been used up.

---

[4] All experiments cited in the last section lasted for at least several hundred generations. Most common were 1000 generations, a few experiments lasted 2000 generations. The Lenski experiment is the single experiment with published data up to generation 50000.

[5] N.B.: The ancestry of the strain of *E. coli* was traced back to the Institute Pasteur in the early 20th century [DSL⁺09].

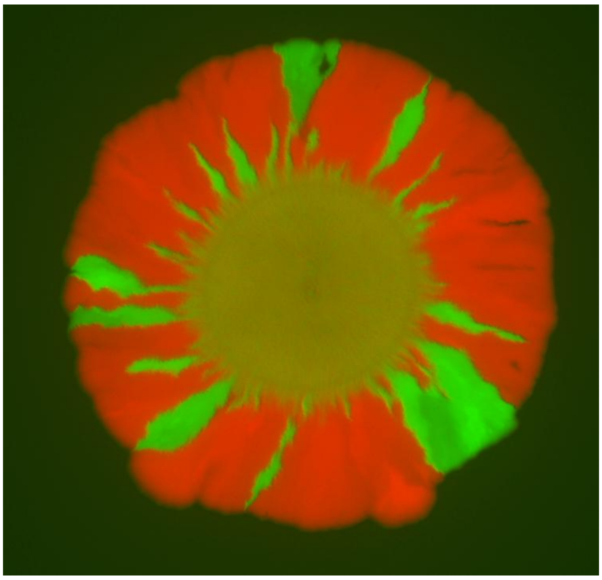

Figure 1.2.: **Spatial distribution of clones after three days of growth.** A plasmid encoding one of two different color variants of a fluorescent protein was inserted into *E. coli*. Each variant emits light at different wavelengths. No other difference exists between the two strains and they both grow at the same rate (except for mutations occurring during growth). A mixture of both variants was inoculated on an agar plate. The speckled pattern in the center is the remnant of this original mixture. After inoculation, cells only grow only within a small boundary layer. The colony extends and genetic drift at the edge creates the observed sector pattern: once a lineage lost contact to the front, it does not influence the further extension at the boundary. Experimental setups of this type have been pioneered in [HHRN07]. Image by Fabian Stiewe.

While the lineage carrying the mutation for citrate uptake was still growing, other beneficial mutations occurred in the remaining population. In particular, some cells acquired mutations that made them use glucose faster [BBL08]. Ultimately, both lineages formed a stable coexistence, each within its ecological niche: one lineage is faster in utilizing glucose and can use a larger share of it than the lineage taking up citrate. However, the latter continues to grow on citrate.

1.2.3. Other experimental setups

Most examples so far have evolved microbes in liquid culture, often under shaking conditions. There, the habitat can be considered isotropic, which simplifies theoretical approaches enormously. However, real habitats are often structured. *E. coli* originally inhabited the intestine tract of humans before it became the ubiquitous "pet animal" of biologists in their labs. When microbes are grown on agar-plates, the surface is a two-dimensional habitat. Adding spatial dimensions usually complicates the experimental setup, but on the other it can hand also introduce additional interesting effects.

Various such effects are emphasized in spatial evolution, in particular during range expansions. Chance effects play a crucial role in determining the individuals that generate successful offspring at the edge of a colony [HHRN07]. Simple microbial experiments can be used to investigate the effects of range expansions on genetic diversity (see Figure 1.2). The rim of reproducing organisms is only a few cells wide at the edge, leaving a "frozen record" of largely quiescent cells behind. On the boundary a constantly reproducing lineage of cells exists. Their lineage accumulates generations faster than in liquid culture [FH15]. In the spatial case the nutrients are only accessible to microbes at the front, while in liquid culture all cells use up nutrients. These experiments are not only faster, but also allow to examine spatial effects that could not be seen in liquid culture, *e.g.* [MNNM14, WPH$^+$14].

1.3. Modeling adaptation in microbial populations

When trying to craft these experimental results into mathematical models, we can build on an almost century long history of population genetics [Fis30, CK70, Gil10]. Population genetics is instrumental in our current understanding of evolutionary processes.

1.3.1. Population genetics

Population genetics is the theoretical discipline trying to describe the change of *allele frequencies* within a population. Here, an *allele* is defined as a mutational variant on a given position on the genome. This dynamical process is influenced by several different "forces". We mainly focus on three "forces" ultimately relevant for our modeling purposes. These are *mutation*, *genetic drift* and *selection*. Each will be elucidated in its own paragraph. Other concepts, that are often associated with population genetics, will be briefly mentioned, but are not instrumental to this thesis.

Mutations

Mutations are changes affecting the genome. These changes can either be substitutions of single bases, then called SNPs (single nucleotide polymorphisms), or larger scale rearrangements of extended stretches of DNA, or addition or removal of genetic material (called insertions or deletions). Depending on the mutation, effects vary. When within a coding region, the original function of the encoded protein can be perturbed or altered. Mutations can also appear in regulatory regions of the DNA. Then, the dynamics of gene expressions could be distorted, and protein expression up- or down-regulated. Recently, a consortium tried to map all regulatory interactions in the human genome, but the analysis of the data is far from complete [ENC04, RCR$^+$11]. However, some of the conclusions have been questioned [GZP$^+$13].

These complex interactions can have an impact on the growth capacity of the cell. In population genetics, however, the effect of a mutation is often reduced to a single number, that quantifies how much faster or slower the cell grows and can produce offspring (in the environment used for the experiment). This number, the *selection coefficient*, is used to classify mutations into various groups: *neutral* mutations do not have an effect on fitness. Usually, those are considered the most common mutations. *Deleterious* mutations reduce the reproduction capacity of the individual, while *beneficial* mutations increase it.

Mutations ultimately create the molecular variation upon which selection can act.

Genetic drift

Genetic drift refers to the stochastic processes occurring in a population of finite size, because of random variation in the number of offspring of each individual. Genetic drift is very strong for small populations. It also encompasses external factors in the life cycle of organisms. These include, for instance, mortality of individuals before they can reproduce.

A dilution step is part of many experimental protocols explained above. This dilution is necessary to provide new nutrients. For instance, in the Lenski experiment (*cf.* section 1.2.2) cells are diluted each day by a factor 1:100. This resampling process of cells is almost independent of any other of their traits, like fitness.

The amount of stochasticity, *i.e.* genetic drift, can vary to some extent with different experimental (or natural) conditions. For instance, the spatial microbial experiments mentioned above in section 1.2.3 have an increased amount of stochasticity in determining the successful lineages at the front [SE12], compared to growth in liquid culture. Some animals also follow a different evolutionary strategy by producing an enormous amount of fertilized eggs. Frogs are an example for this strategy: only a minute fraction of tadpoles grow and live long enough to reproduce. There, chance effects are important in determining the genetic material destined to be carried to the next generation.

Often, an *effective population size* N_e is introduced in population genetics for formal reasons. It can incorporate these different amplitudes of fluctuations in offspring number [Gil10]. However, we do not use effective population sizes here: the parameter N is always considered the *census population size*.

Selection

Selection refers to process of changing the frequency of a *heritable* trait in a population that affects the reproduction capacity of individuals carrying this trait. If the trait increases the reproduction capacity, it can accumulate in the population over the course of generations, and the population adapts, which was Darwin's seminal insight [Dar59]. However, selection does not act on only a single individual. It depends on the existence of multiple individuals (or a population) with perceptible differences in a particular trait.

In population genetics, the concept of *fitness* is tightly linked to selection. Fitness defines the rate at which individuals reproduce on average, while all stochastic effects in reproduction number are attributed to genetic drift, as explained above. Although one can measure an absolute value for fitness, it is often used as relative number between clones carrying different alleles.

A selective advantage crucially depends on the environment: an allele could be beneficial in one environment, while deleterious in another. In the microbial experiments above, some mutations affect proteins involved in metabolic pathways that are not associated with glucose, but other carbohydrates [BBL08]. In the experimental conditions used, such mutations are neutral or even beneficial as glucose is abundant. Under other conditions, growth of the different experimental population would be highly variable [BBL08].

As we have illustrated above, when describing observations in microbial experiments, in some situations selection does not only unconditionally favor one allele over another. Then, two (or rarely, more) alleles can coexist, or even reinforce each others coexistence. This type of selection is called *frequency dependent selection*, as the scenario can be described by selection coefficients that depend on the frequency of the allele in the population.

Other concepts

The frequencies of alleles in a population are not only determined by selection, mutation and genetic drift. We briefly mentioned spatial distributions of organisms already, with an example shown in Figure 1.2. In extended habitats, *migration* becomes another crucial factor in the dynamics of allele frequencies.

In general, spatial structure implies a more elaborate demographic scenario than just using the single number that indicates the population size in the well-mixed (or zero-dimensional) case. Then, it is often implemented by discretizing the whole population into *demes*, *i.e.* spatial sub-populations, that are arranged on a (regular) lattice. The stepping-stone-model is the best example of such an approach [KW64, WK65].

Spatial dimension will not be used in our model of adaptation. However, in chapter 6, we deal with a model that explicitly treats an extended habitat when investigating the localization of a population at a nutrient source.

For completeness, we also mention *recombination* as an important aspect in the dynamics of alleles. Through recombination, mutations in different individuals (or lineages) can combine into a new individual. Sexual reproduction is the main example where recombination happens. Other mechanisms, like *horizontal gene transfer* also exist [GDL02]. In horizontal gene transfer, genetic material is taken up from the environment and the information on this "new" DNA is utilized. The emergence of sexual reproduction is often considered as a possible solution to the evolutionary problem that arises in clonal competition scenarios: many beneficial mutations are wasted, as they occur on different genetic backgrounds. Recombination can bring those mutations together, thereby speeding up evolution in certain parameter regimes. Research on these effects of sexual reproduction dates back to founders of the field of population genetics [Fis30, Mul32, HR66]. However, here we will not follow on these ideas, as it would be an extended route on its own.

1.3.2. Dynamical regimes in mutation-selection models

Several different dynamical regimes can be distinguished in population genetic models. Those depend on the relative strength of the aforementioned forces to each other.

A main observable in so-called mutation-selection models is the adaptation speed. It is defined as the rate of the increase of mean fitness.

Periodic selection – weak-mutation, strong-selection regime

The periodic selection regime can be considered the classical dynamical regime in population genetics. Mutations are rare, while selection is strong, *i.e.* the selection coefficient of these mutations is large. Only a single mutation is considered at a time: the waiting time until a novel mutation occurs is much longer than the time of a *sweep*. Here, a sweep denotes the process of increasing the frequency of an allele from low values to fixation via selection.

Therefore, the limiting factor of adaptation speed in this regime is the mutational supply. Novel mutations occur in individuals with a constant, small rate. After the genome of one individual has mutated, the number of its descendants has to reach a certain threshold before selection can increase the sub-population exponentially. Until then, genetic drift affects the mutant sub-population significantly and can drive the mutation to extinction again. This threshold is inversely proportional to the selection coefficient, and independent of population size. Moreover, the probability that the sub-population reaches the threshold is proportional to the selection coefficient [Gil10]. Hence, the population size enters adaptation speed only via the rate of occurrence of mutations:

$$v \sim N \; . \tag{1.1}$$

The proportionality coefficient of this linear scaling (1.1) depend on the actual mutation rate and the selection coefficients of the fixating mutations.

Muller’s ratchet – abundance of deleterious mutations

If the mutation rate is increased, often multiple clones with different fitness exist in the population simultaneously. Mutation-selection models treating such a diverse population were first considered in the context of *Muller’s ratchet* [Mul64, Fel74]. There, a gradual loss of mean fitness occurs due to the lack of beneficial mutations and extinction of the highest fitness class as consequence of genetic drift.

Figure 1.3 illustrates the main concepts in this dynamical regime. When mutations are mostly deleterious, they are in general only relevant at the lower end of the fitness distribution in the population. There, they create individuals with an even lower fitness. Stochasticity, *i.e.* genetic drift, is implied by the finite size of the population. In the context of Muller’s ratchet, it is only important for the loss of the highest fitness class. Once it is lost it cannot be regained, as beneficial mutations do (usually) not occur. This ”separation“ of mutations, selection and genetic drift reduces mathematical complexity.

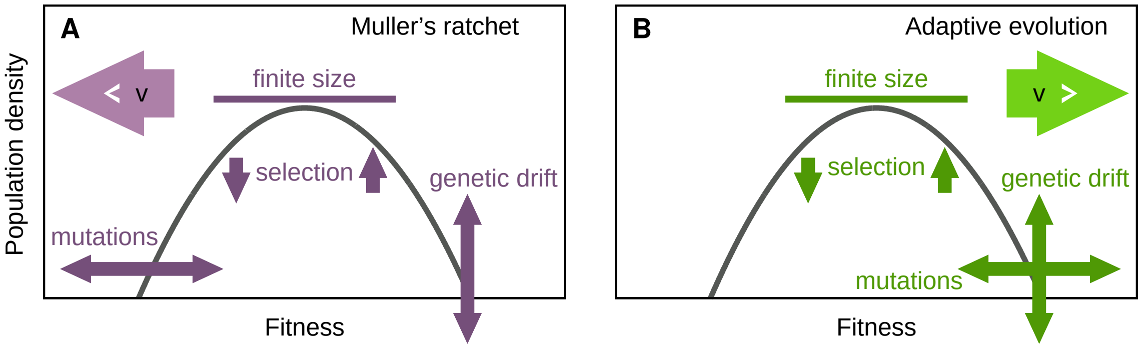

Figure 1.3.: **Schematic representation of major concepts in models of Muller's ratchet (A) and Adaptive Evolution (B).** A population is represented by a density in fitness space (dark line). Population genetic models often include mutations, selection and genetic drift (that implies a finite size of the population). Muller's ratchet represents a dynamical regime where mutations are predominantly deleterious, hence the noise of mutations and genetic drift can be separated. In general the density moves to lower fitness. Adaptive evolution (and clonal competition in particular) on the other hand, requires sufficiently many beneficial mutations. In turn, mutations, genetic drift and selection create intricate dynamics in the nose of the wave. Overall, the density moves to higher fitness with speed v.

Extensive literature is analyzing this situation [HRWB02, RBW08, Jai08, GBJ+12, NS12, ME13].

Clonal competition – strong-mutation, strong-selection regime

The dynamical regime of clonal competition is the main focus of this thesis. As illustrated in section 1.2 on recent experiments, many microbial populations experience dynamics of this type: various beneficial mutations compete for fixation simultaneously. One among the first publications to investigate this dynamical regime in the light of those microbial populations has been [GL98]. There, the probabilistic dynamics of concurrent beneficial mutations with different selection coefficients was analyzed. In recent years this idea has been generalized in many directions, as will be outlined below.

A central question is how fast these microbial populations adapt [GL98, RWC03, CKL05, DF07, PK07, BRW08, RBW08, SG10, PSK10, Hal11b, GRB+12]. The main concordant result is that adaptation speed v depends *logarithmically* on population size N:

$$v \sim \left(\log N\right)^{\alpha} . \tag{1.2}$$

The exponent $\alpha \sim \mathcal{O}\left(1\right)$ and the proportionality coefficient in expression (1.2) depend on the model implementation. Most of the variation among cited publications can be attributed to how mutations are incorporated in the models.

The *logarithmic* dependence in (1.2) emphasizes the fact that adaptation depends only on a small, highly fit sub-population. Additional mutations occurring in one of these individuals will increase their fitness beyond currently present fitness values. However, while small in number, these new sub-populations are highly influenced by genetic drift (see Figure 1.3). If the sub-population survives, these individuals (and their descendants) will drive adaptation to ever higher fitness. Mutations on a low fit genetic background are often wasted for the adaptation process in these models. We will outline the derivation of the various relations condensed into (1.2) later, in the next chapter.

To simplify the analysis, the dynamics of clonal competition is often investigated in a steady state. Mutations entering the population increase genetic diversity, while selection counteracts this tendency and decreases variation[6]. Clones with a higher fitness will increase in frequency, while clones with smaller fitness will dwindle. In the clonal competition regime where mutations are common, however, diversity within the population remains. This diversity is modeled as a *traveling wave* in fitness space: the whole population consists of multiple clones with varying fitness. The dynamics of selection, mutations and genetic drift creates a soliton-like behavior, *i.e.* a moving, bell-shaped density in fitness space (*cf.* Figure 1.3).

1.3.3. Fluctuations in the adaptation process of asexual populations

The mean quantitative outcome of this traveling wave approach seem to be understood quite well. Adaptation speed can be calculated by various models and different approaches [GL98, DF07, GRB$^+$12]. The distribution of mutational effects that fixate in the population was also computed recently [GRB$^+$12]. Despite these successes on calculating the mean behavior, a full theory explaining the fluctuations of traveling waves is still absent. Recently, an attempt to describe various universality classes in the fluctuation behavior of traveling wave models was published [Fis13].

Adaptation depends crucially on fluctuations in small sub-populations carrying extraordinarily beneficial mutations. Such sub-populations are always small at the beginning and heavily influenced by genetic drift. Slight deviations from the expected behavior could have tremendous impact on the overall dynamics at a later time. Hence, fluctuations in the fitness variance will inevitably occur.

In addition to these (small) fluctuations, another process affects the fitness variance of asexual populations on larger scales: if a very fit clone emerges and sweeps through the population, often the majority of the variance is lost and it takes a while until diversity of the population is restored. However, even then, the next sweep might repeat this process, leading to large scale oscillations in fitness variance during adaptation.

A first step towards the quantification of these (large) fluctuations will be proposed later in chapter 5.

---

[6] Here, selection only decreases variation: this implies that no frequency-dependent selection occurs.

1.4. Quantifying evolutionary dynamics

Quantifying the evolutionary dynamics of adaptation in large, asexually reproducing populations is the main aim of this thesis.

Chance effects play a crucial role in these populations: during the resampling only a fraction of the population is transferred to the next step. Mutations alter the genotype of organisms. Therefore, the population turn-over inherently depends on these random events. While mean effects are explained in recent literature, the fluctuations of important observables remain poorly understood. We will ask – and partially answer: what role do fluctuations play in the adaptation process?

Quantification and modeling are ultimately linked to a mathematical description. Without proper mathematical tools, many quantitative questions remain unanswered. To this end, we aim to develop, extend and explain appropriate methods to tackle the previous question of explaining fluctuations of crucial quantities in the adaptation process.

1.4.1. Structure of this thesis

The structure of this thesis is as follows. In the next chapter, we introduce a basic simulation model that tries to incorporate the essential features of the evolution experiments with microbes that were explained in section 1.2. At first, we only present elementary phenomenology. After that, in chapter 3, we build up the mathematical model. It is based on branching random walks. Of crucial importance is section 3.2, where we introduce tuned models. Tuned models have been originated in [Hal11b] and are a special version of traveling wave models where the governing equations are *exactly* solvable. The rest of the thesis is devoted to explaining details and peculiarities of asexual evolution by means of tuned models. In particular, we are interested in the fluctuations exhibited by traveling waves. Chapter 4 investigates the behavior of population size fluctuations. Chapter 5 deals with fluctuations in the fitness variance, or – as will be seen later, equivalently – with fluctuations in adaptation speed.

The contents of chapter 6 have already been published as [GH13]. There, we employed the same method of using tuned models in order to analyze a localization/delocalization problem. Although the interpretation of the model is different, the underlying equations are quite similar. We conclude with the discussion in chapter 7.

2. Simulations of models of adaptation

In this chapter we present a simulation model of a diverse population, considering selection, mutation and genetic drift. Our model is designed to be applicable to the large asexual populations, that adapt to laboratory conditions. Experimental evidence for these populations was summarized in the last chapter. Individuals with different growth rates (fitness) exist simultaneously. A large influx of novel beneficial mutations induces a population that is spread out in fitness space. The labeling as “traveling wave models” arises from the phenomenology in the model: different clones grow and mutate. Mutations change their fitness. If a clone with additional beneficial mutations survives genetic drift, it will grow faster than the rest of the population. An additional constraint keeps the population finite. The enveloping, bell-shaped density of the complete population forms a soliton-like traveling wave with increasing mean fitness over time.

After the model has been introduced in section 2.1, we present results in section 2.2. These simulation results are compared to analytic predictions, obtained in recent years. Moreover, these simulations form the basis on that we will build our (exactly solvable) theoretical model in chapter 3. Results presented in the current chapter will be revisited in the light of these new theoretical considerations in chapter 4.

Here, we focus only on a minimal phenomenology. Other (more detailed) aspects of observable phenomena will be described and explained later on.

2.1. Simulation model

A population of N asexually reproducing individuals is represented by a density $\mathbf{n} \equiv n_i$ ($i \in \mathbb{Z}$) in fitness space[1]. The number of individuals in a (discrete) fitness class is denoted by n_i. At all times, we demand a constant (finite) population size,

$$\sum_i n_i = N \; . \tag{2.1}$$

Membership in a fitness class is the only property of an individual. Within class i, individuals reproduce with rate $1 + x_i - \overline{x_t}$ and die with rate 1. This sets the unit of time to one generation. The expected net reproduction rate is $\left(x_i - \overline{x_t}\right)$ with x_i the (relative) fitness of individuals in class i, while $\overline{x_t} = \left(\sum_j x_j n_j\right)/N$ is the mean fitness of the whole population at time t. Growth of the number of individuals within a fitness class is exponential, hence we understand $\left(x_i - \overline{x_t}\right)$ as Malthusian (log-)fitness[2]. The occupancy n_i in fitness classes declines when $x_i < \overline{x_t}$. All fitness classes are evenly spaced with distance dx to each other, such that $x_{i+1} = x_i + dx = (i+1)dx$.

---

[1] In this thesis bold symbols, *i.e.* $\mathbf{n}$, will always denote elements of an (appropriate) vector space. This notation is used extensively in later chapters.

[2] Every time we use the term *fitness* in this thesis, it could be replaced by *Malthusian (log-)fitness*. Due to readability we will not do that.

The time evolution of the population is illustrated in Figure 2.1. A single step of duration ϵ in the simulation consists of several sub-steps in the algorithm. In each of those sub-steps the occupancies (can) change. The upper indices in the following equations denote the corresponding panel in the graphical explanation (Figure 2.1). In the first sub-step the mean (deterministic) change of occupancies due to selection and mutations is applied,

$$\Delta^{(A)} n_i/\epsilon = \big(x_i - \overline{x_t}\big) n_i + \mu \left(\sum_j m_j n_{i-j} - n_i \right) . \tag{2.2}$$

The first term represents effective growth due to selection, while the second term accounts for the modification of fitness due to mutations. Possible choices for the mutation kernel $\mathbf{m} \equiv m_j$ are explained below. In the next sub-step, the randomness due to birth and death events (*i.e.* genetic drift) further modifies all occupancies n_i,

$$\Delta^{(B)} n_i/\sqrt{\epsilon} = \sqrt{2}\big(\text{Poisson}(n_i) - n_i\big) , \tag{2.3}$$

where Poisson(n_i) is a Poisson-distributed random number with parameter n_i. This particular form of the noise term (with n_i already updated from selection and mutation) ensures that (i) occupancies n_i do not drop below zero, (ii) the mean value of the noise term is zero and (iii) the variance within a single fitness class amounts to $2\epsilon n_i$.

Notice, that additional variance due to randomness in the mutation process can be neglected. Many mutations enter the population, but only few of them will establish a new fitness class. With that choice we implicitly assume that the fitness effects of novel mutations are tiny, such that all the occurrence and subsequent establishment of novel mutations tends to average out: we model this process with a constant rate. Thus, the main source of stochasticity is genetic drift.

In the last sub-step the population is scaled uniformly to fulfill the constraint, (2.1):

$$\Delta^{(C)} n_i = \left(\frac{N}{\sum_j n_j} - 1 \right) n_i . \tag{2.4}$$

Although the deterministic sub-step, (2.2), does not change the number of individuals, and also the expected change due to genetic drift is zero, (2.3), tiny deviations from this expected behavior will add up. Without this normalization, the stochastic changes in (2.3) lead either to extinction or infinite growth of the population. In order to execute the uniform scaling, we allow fractional occupancies n_i of fitness classes. This is similar to a diffusion approximation valid for large populations [Gar09, Kam92].

After these three sub-steps, one full time step is concluded and the mass of the population density is (usually) shifted to higher fitness classes. The rate of this shifting defines the adaptation speed v. As we are mainly concerned with adaptation in this thesis, we choose our parameters such it is unlikely that the population shifts to lower fitness (although it can happen in the stochastic algorithm).

The simulation algorithm is similar to the one already published in [GH13]. In this thesis we will revisit the results from this particular publication in chapter 6.

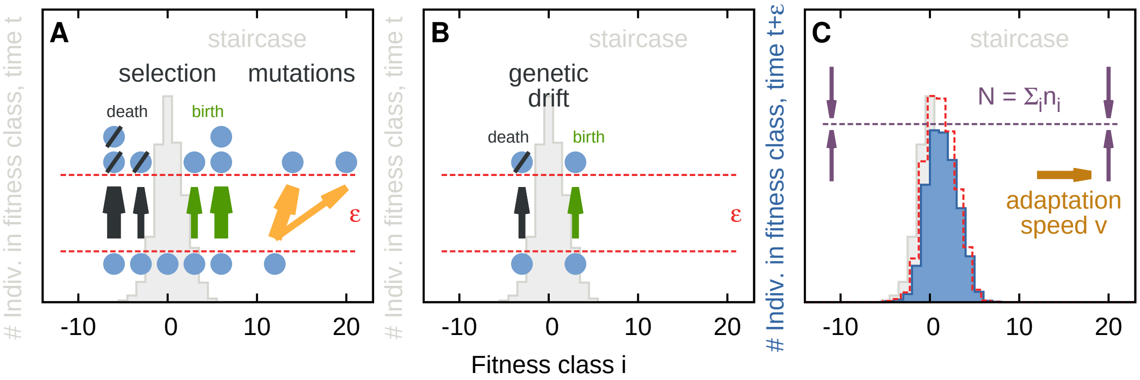

Figure 2.1.: **Stochastic dynamics of the traveling wave model.** The time evolution of the population density $\mathbf{n} \equiv n_i$ (gray profile) during a time step ϵ consists of several sub-steps: (A) Selection adds or subtracts individuals from the fitness class n_i. Mutations increase (or decrease) the fitness of the affected individual, (2.2). (B) Stochastic changes to n_i are applied due to the randomness in the reproduction of individuals (genetic drift), (2.3). (C) After the (stochastic) population turn over from (B) has been applied, usually slight deviations from the constraint $\sum_i n_i = N$ exist (the red dashed line depicts an exaggerated version). A (uniform) global scaling ensures that the population size is kept constant, (2.4), resulting in the final population density $\mathbf{n}$ (blue profile) after time ϵ has elapsed. This uniform scaling forces us to allow fractional occupancies n_i.

2.1.1. Mutation models

Mutations ultimately drive the adaptation process. Without them, evolutionary innovation would not be possible. In this section we describe (three) models, *i.e.* mutation kernels, which have already been used throughout recent literature to model the effect of novel mutations in traveling waves models of adaptation. As already used in (2.2), mutations that shift an individual from fitness class $i - j$ to the new fitness class i occur at rate μm_j. Moreover, novel mutations are independent of the fitness class $i - j$ where they occur. The probability of a fitness increase by $j\,dx$ is the same for all fitness classes. This is equivalent to the infinite sites model of mutation [Gil10, Wat75]: the effect of a new mutation does not depend on the genetic background it appears on. We also neglect any epistatic interaction between mutations. The mutation kernel $\mathbf{m} \equiv m_j$ in our simulations is a discrete approximation to the *distribution of fitness effects* (DFE).

The choice of a DFE is a crucial ingredient in modeling adaptation. However, the exact distribution of effects of possible mutations is hard to map experimentally. It has been measured for subsets of possible SNPs (single nucleotide polymorphisms) in single genes. Antibiotic resistance is a prime example, where effects of SNPs on growth can be measured directly [WDDH06]. Other experiments with single traits have been surveyed in [SSF$^+$13]. A major problem is that the mutation process cannot be turned off: mutations will enter the population perpetually, due to the faulty nature of the DNA replication mechanism. Observing mutations is conditioned on the fact that subpopulations carrying the mutation rise to appreciable numbers in the population, *i.e.* the descendants of the mutated individual have to overcome genetic drift. Even though experimental techniques improved a lot from simple markers, like microsatellites, to full genome sequencing nowadays, it wont be feasible (in the near future) to sequence every single clone in a microbial population of 10^7 or more cells. Larger genome rearrangements might have even a larger and more unpredictable impact on the fitness compared to SNPs.

The DFEs (or used equivalently here, *mutation kernels*) are based on either theoretical considerations, experimental evidence or simply analytical tractability. A review on the statistics (and dynamics) of DFEs can be found in [EWK07], a more recent one is [SSF$^+$13].

Some general assumptions on all mutation kernels will be made throughout this thesis. First, we demand a normalization of the kernel, $\sum_j m_j = 1$, any numerical prefactor will be absorbed in the (global) mutation rate μ. Furthermore, we will use the symbol σ to denote the mean effect of an mutation,

$$\sigma = \sum_j x_j m_j \ . \tag{2.5}$$

Most of the time, the dynamics of traveling waves can be stated with parameters in units of σ. For instance in [DF07, GRB$^+$12] the authors use the reduced parameters $N\sigma$ and μ/σ to fully disclose their results. These two quantities describe the strength of selection against genetic drift and the strength of selection against the influx of mutations, respectively [DF07]. In figures we will usually measure fitness in units of σ. Although time can also be declared in units of σ^{-1}, we will mostly keep the more intuitive unit of one generation.

Two different regimes of the dynamics can be distinguished. In the first regime mutations occur at a low rate, but have a large effect on fitness, $\mu/\sigma \ll 1$. Such an approximation is valid when treating the dynamics of “driver mutations” in cancer models, which enable different stages of cancer progression [BAD$^+$07, BAO$^+$10, BST14]. The other extreme case is given by $\mu/\sigma \gg 1$. Here, almost neutral mutations enter the population at a high rate. This scenario might apply to fast mutating viral populations [DH99] or also (but less likely) to microbial colonies [SGL97]. In the following we describe three different mutation kernels, chosen to capture either regime in the dynamics (though by different kernels). Of course, one can use parameters for the opposite regime in all kernels, but then parameter values start to deviate significantly to those one would expect from experiments. In an intermediate regime, however, parameters for all three kernels can be matched, using reasonable values. A major aim of this thesis is to compare results from different kernels and regimes.

Staircase mutation kernel

The "staircase" mutation kernel is the simplest mutation kernel one can think of when trying to model adaptation. Here, all mutations are beneficial, and all mutations have the same effect σ. Adaptation proceeds by increasing fitness step after step in mutations, like in a staircase. It can be stated as

$$m_j = \delta_{j,1} \ , \tag{2.6}$$

when fitness classes are spaced with distance $dx = \sigma$ to each other[3].

At first, the two assumptions made with this mutation kernel seem like an enormous oversimplification. However, it has been argued that deleterious mutations can be neglected when trying to model adaptation [DF07]. An individual receiving one of those deleterious mutations will not (or at least unlikely) participate in the competition to fixation anymore. As long as the mutational process is not completely dominated by deleterious mutations occurring at a large rate, we can focus on the "trajectory" of clones that only receive beneficial mutations. We will come back reviewing the impact of deleterious mutations at the end of the thesis, when the whole formalism, and results obtained with it, have been presented (see section 7.3.4).

The second simplification, treating all mutations to have an identical effect σ, simplifies the analysis of adaptation (enormously). In some sense, the dynamics generated by this mutation kernel appears to be generic for adaptation. Many (more complicated) DFEs can be mapped back to this case by using a set of "effective parameters" [GRB+12]. Mean quantities obtained by this this kernel coincide with these effective parameters arising from these more elaborate calculations. It has been argued that the dynamical behavior of all short tailed mutation kernels (with a tail decaying faster than exponential) is in the same universality class [Fis13].

Exponential mutation kernel

In an exponential mutation kernel "large" jumps in fitness are possible. The probability for an increase in fitness decays exponentially,

$$m_j = c \, \exp\bigl(-j \, dx/\sigma\bigr) \ , \quad j \geq 0 \ , \tag{2.7}$$

while for $j < 0$ we set $m_j = 0$ to incorporate only beneficial mutations in the mutation kernel. The constant c is adjusted to comply with the normalization of the mutation kernel, and depends on dx and σ[4].

---

[3] Any other choice of dx would not make sense: technically, $dx < \sigma$ would create σ/dx (independent) populations with an unnatural coupling of the overall population size in the algorithm of section 2.1.1.

[4] The normalization is given by $c = \bigl(\exp\bigl(dx/\sigma\bigr) - 1\bigr)$. In the continuous limit, $dx \to 0$, we converge to $c \to dx/\sigma$, and arrive at $\int dx/\sigma \ \exp\bigl(-x/\sigma\bigr) = 1$ (with $x = j \, dx$). Strictly speaking, in the discrete case, σ is not the mean effect of an mutation in the exponential kernel. However, deviations are tiny. For $dx/\sigma = 1/10$ (a value we use in simulations) the real mean mutational effect is smaller by a factor 0.951, while for $dx/\sigma = 1/20$ (a value we use in the numerics later on) we acquire a correction factor of 0.975. The *stochastic* simulations would have to be run for a very long time to distinguish such differences.

Exponential DFEs have received considerable attention in the population genetics literature, both in experiments [IS01, KB06], and also theoretical considerations based on extreme value theory yield an exponential DFE [Orr98, Orr03, Orr06, EWK07]. In the context of traveling wave models, those exponential mutation kernels have been incorporated for instance in [GL98, GRB+12].

The exponential kernel, together with the staircase mutation kernel discussed above, are main examples of the dynamics in the regime $\mu/\sigma \ll 1$: even a single mutation has a significant impact on fitness.

The dynamics of exponential DFEs results in a much more fragmented population density than the staircase mutation kernel explained above and the diffusion mutation kernel explained below (see also Figure 2.2). However, individuals are still only spread over a small number of (almost) consecutive fitness classes. Heavier tails than exponential in the distribution can lead to the creation of separate sub-populations far ahead in fitness space of the current population [HF14]. The exponential DFE seems to be on the transition to this much more erratic behavior [Neh13].

Diffusion mutation kernel

The diffusion mutation kernel is a prime example of the second regime of the adaptation dynamics, where (meaningful) parameters are such that $\mu/\sigma \gg 1$. Its mathematical description is given by

$$m_j = \mu^{-1}\,\frac{D}{dx^2}\big(\delta_{j,1} - 2\delta_{j,0} + \delta_{j,-1}\big) + \delta_{j,0}\ . \tag{2.8}$$

The terms in (2.8) are constructed in a way to modify the original general form of mutations in (2.2) and scale out the mutation rate μ. In turn, the first (sub-)step in the simulation algorithm is given by $\Delta^{(A)} n_i = \big(x_i - \overline{x_t}\big) n_i + D/dx^2\big(n_{i+1} - 2n_i + n_{i-1}\big)$. The mutation term is reminiscent of a (discrete) second derivative with respect to fitness, hence the name “diffusion”.

Due to symmetry in this second derivative we cannot define the mean effect of a mutation by (2.5): the terms would cancel and $\sigma = 0$. The analysis done in [TLK96, CKL05, Hal11b, NH13] suggests that $D^{1/3}$ sets the scale in the diffusion kernel. Thus, the diffusion constant D takes the role of both, the mean mutational effect, $\sigma \sim D^{1/3}$, and also of the mutation rate $\mu \sim D^{1/3}$ which we explicitly scaled out with the prefactor μ^{-1} in (2.6). In figures we present fitness in units $D^{1/3}$ instead of σ for this mutation kernel. In order to conduct simulations, we usually choose the width dx of fitness classes much smaller than the corresponding value used in the staircase (and exponential) mutation kernels.

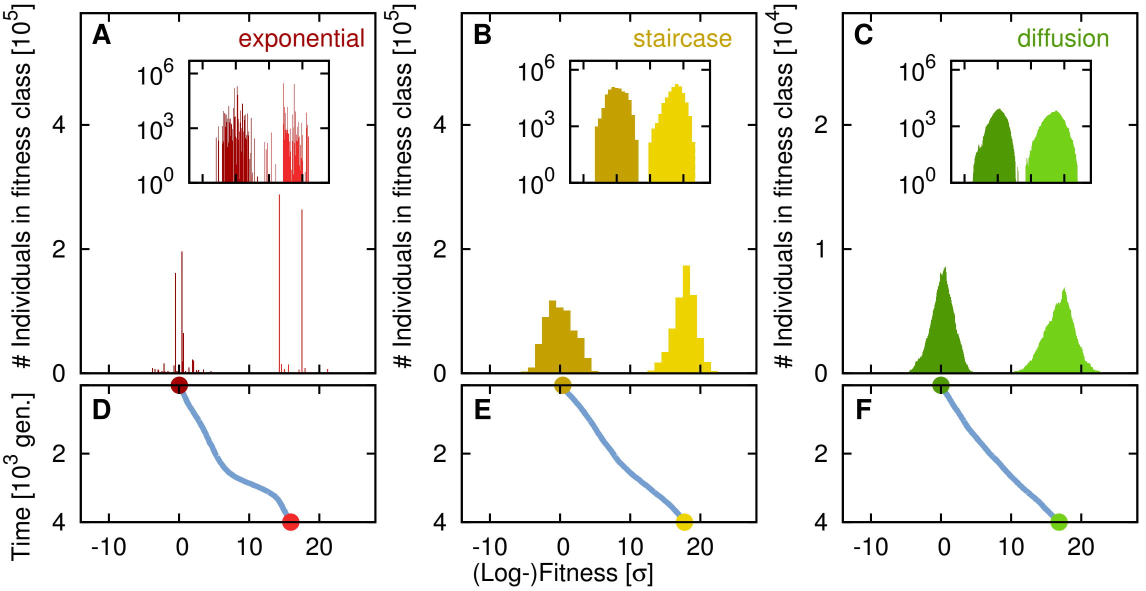

Figure 2.2.: **Snapshots of occupancies in different fitness classes of traveling waves with different mutation kernels.** (A) - (C) In the upper part of the figure we show the occupancies n_i in normal and logarithm scale (insets). The occupancy n_i is defined as the number of individuals growing (on average) with (log-)fitness $\left(x_i - \overline{x_t}\right)$. The population size $N = \sum_i n_i = 6.3 \cdot 10^5$ is kept constant. Moreover, the (mean) adaptation speed is adjusted to $\overline{v} \approx 4 \cdot 10^{-6}$ by varying the mutation parameters. We find $dx^{\text{diffusion}} = 5 \cdot 10^5$, $D = 10^{-9}$ for the diffusion kernel. In the exponential and staircase mutation kernel we set $\sigma = dx^{\text{staircase}} = 10^{-3}$, $dx^{\text{exp}} = 10^{-4}$ and varied the mutation rate μ such that we obtain a matching adaptation speed $\overline{v}$. This resulted in $\mu^{\text{exp}} = 4.1 \cdot 10^{-5}$ and $\mu^{\text{staircase}} = 6.1 \cdot 10^{-4}$. Both snapshots (for each mutation kernel) are taken with the dynamics exhibiting a steady state, when (on average) the adaptation speed reached its desired value. Notice that the vertical scaling in the diffusion mutation kernel, (C), is reduced by a factor of 20 in order to make up for the factor of 1/20 in the width of fitness classes. That the average bell-shaped profile is roughly Gaussian can be estimated from snapshots in logarithmic scale. (D) - (F) In the lower part of the figure we display the time evolution of the mean fitness in the population, connecting the two snapshots that are 4000 generations apart.

2.2. Simulation results

2.2.1. Snapshots of population densities

In the following we present simulation results obtained with the algorithm explained in section 2.1 above. We will focus on the *mean* adaptation speed $\overline{v}$ resulting from the parameters population size N, mutation rate μ and mutational scale σ (resp. diffusion constant D). A more detailed analysis of $\overline{v}$, including the fluctuation behavior of adaptation speed, is postponed until chapters 4 and 5.

A first impression of the noisy time-evolution in simulations is shown in Figure 2.2. In the depicted snapshots, the population size was kept constant at $N = 6.3 \cdot 10^5$ for the staircase, exponential and diffusion mutation kernels. All (mutation) parameters were adjusted such that the resulting adaptation speed, $\overline{v} \approx 4 \cdot 10^{-6}$, was the same.

Figure 2.3 shows the time evolution of mean fitness. A primary observation is evident: fitness does not increase linearly over time. The mean adaptation speed $\overline{v}$, however, falls within a narrow range of values for simulations with identical parameters. Irregularities exist in all mutation kernels, but they are most pronounced in the exponential mutation kernel. Fluctuations around the linear fitness have intermediate amplitudes in the staircase mutation kernel, while in the diffusion mutation kernel the advance to higher fitness is rather smooth.

In the exponential mutation kernel it becomes evident that the large effect mutations seeding fitness classes far ahead of the bulk population leads to these large fluctuations in fitness increase: after a seeding event genetic drift can boost the occupancy of the fitness class to a dynamical regime where the influence of selection is stronger than genetic drift. Then, it grows deterministically and soon constitutes a major fraction of the whole population. Mean fitness increases fast during this exponential growth. Thereafter, the fitness increase is slower, because occupancies in fitness classes ahead of this newly dominating class are low, and it takes a time until the process repeats itself. The phenomenology of these large (almost regular) fluctuations will be explored in detail in chapter 5. Here, we just want state that the amplitudes of irregularities in fitness increase decline from exponential to staircase to diffusion mutation kernels.

2.2.2. Mean adaptation speed with constant population size

In figure 2.4 we show the dependence of the (mean) adaptation speed $\overline{v}$ on the population size N. All three cases have already been treated in analytical approximations, such that the dependence on the (two) mutation parameters was found. In the following, we present these approximations and outline the main steps in their derivation.

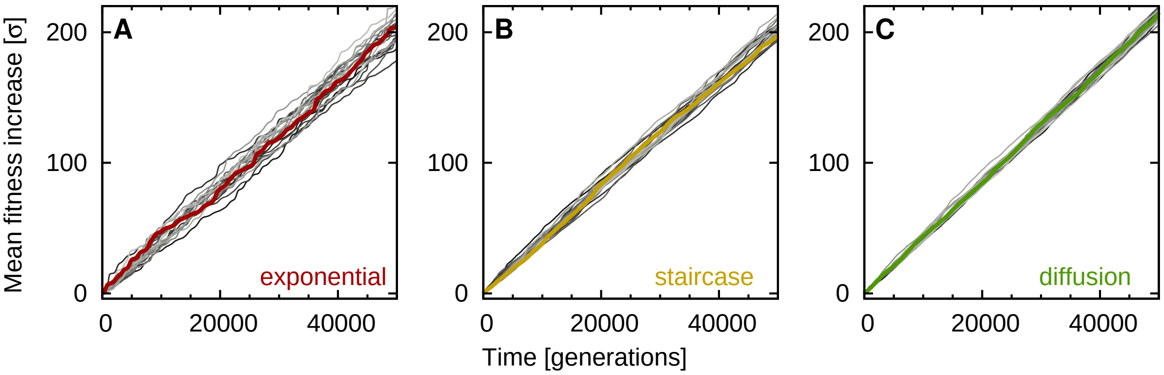

Figure 2.3.: **Trajectories in fitness space measured in stochastic simulations.** 20 sample trajectories for each mutation kernel are shown: (A) exponential, (B) staircase and (C) diffusion.The increase to higher mean fitness is only on average linear. Irregularities and fluctuations are clearly observable in the stochastic simulation model. One trajectory, chosen at random, is drawn in color to emphasize a single trajectory in contrast to the 19 others, indicating the variance in fitness increase. Investigating the fluctuation behavior in detail will be the main focus in chapter 5. Simulation parameters are population size $N = 6.3\ 10^5$ and the mean adaptation speed $\overline{v}/\sigma^2 \approx 4$. Mutation parameters are $\sigma = 10^{-3}$, $\mu^{\text{exponential}} = 4.5\ 10^{-5}$, $\mu^{\text{staircase}} = 6.1\ 10^{-4}$, $D = 10^{-9}$.

Staircase mutation kernel

In the staircase model of mutations (each mutation has the same effect σ), the adaptation speed was obtained in [DF07]:

$$\overline{v}/\sigma^2 \approx \frac{2\log\big(N\sigma\big)}{\log^2\big(\sigma/\mu\big)}\ . \tag{2.9}$$

Similar to our simulation model, the authors use a discrete number of evenly spaced fitness classes. The analysis is based on a single stochastic (sub-)population in the nose of the wave, while all other fitness classes grow (or shrink) deterministically[5]. The adaptation speed is obtained in a steady state, where the highest fitness class of the population (the *nose* of the wave) and the bulk of the population advance at the same speed. Their description of the nose is based on a detailed branching process model, from which they derive the *establishment time*. A new fitness class (with a *lead* of q mutations ahead of the mean population) takes (on average) $\overline{\tau_q}$ generations to reach an occupancy large enough that it is in the deterministic regime of exponential growth, and hence is established. For a second condition, they estimate the time $\overline{\tau_{sw}}$ for this newly established

---

[5] Here is a difference to the staircase model as we use it: we allow stochastic changes in *all* fitness classes.

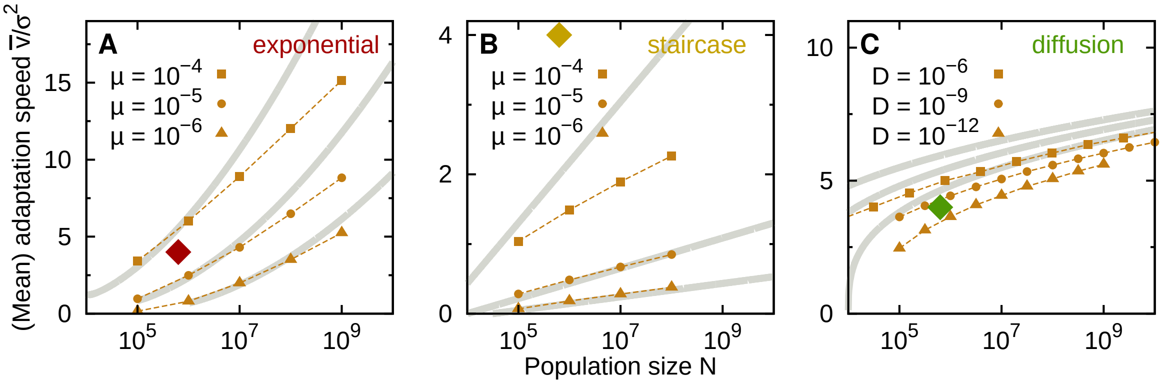

Figure 2.4.: **Mean adaptation speed dependence on population size.** The overall logarithmic dependence on the population size is clearly indicated by the results on different mutation kernels: (A) exponential, (B) staircase and (C) diffusion. Whereas in the exponential mutation kernel (A) the dependence appears quadratic (2.10a), $\overline{v} \sim \left(\log N\right)^2$, in the staircase mutation kernel the dependence is only linear in the logarithm (2.9), $\overline{v} \sim \log N$. In both kernels, the agreement with analytic predictions increases with smaller mutation rates ($\mu = 10^{-4}, 10^{-5}, 10^{-6}$ from top to bottom curves, $\sigma = 10^{-3}$), indicating the approximation $\sigma/\mu \gg 1$ used in the calculations. (C) In the diffusion mutation kernel, the scaling $\overline{v} \sim \left(\log N\right)^{1/3}$ is suggested, (2.11). Different sets of results correspond to diffusion constants $D = 10^{-6}, 10^{-9}, 10^{-12}$ (top to bottom). In the absence of a well-defined mutation scale σ, the adaptation speed is now scaled with $D^{-2/3}$. The diffusion mutation kernel is a prime example of the $\sigma/\mu \ll 1$ limit. Simulation results are presented as brown symbols. The scaling in equations (2.10a), (2.9) and (2.11) are shown as gray lines. Diamond shaped symbols correspond to parameter values used in Figures 2.2 and 2.3.

class to “sweep” through the population and become the fitness class with the largest occupancy. From these two arguments, they arrive at $\overline{v} = \sigma/\overline{\tau_q}$ and $\overline{v} = (q-1)\sigma/\overline{\tau_{sw}}$, where both times still depend on the lead q. After elimination of q, the relation for the adaptation speed $\overline{v}$ in (2.9) is obtained.

Comparing (2.9) to simulation data (see Figure 2.4) shows that the agreement increases significantly with a smaller parameter μ/σ. We show results for three different mutation rates, $\mu = 10^{-4}, 10^{-5}, 10^{-6}$, and the mutational scale $\sigma = 10^{-3}$. While for $\mu/\sigma = 0.1$ deviations from the predicted value are apparent, with the smaller value $\mu/\sigma = 0.001$ prediction and simulations are within a small margin. This is not surprising, as in [DF07] the authors use μ/σ as small parameter and neglect higher orders when stating (2.9). This condition is used as separation condition for their Ansatz of a (single) stochastic population in the lead and the deterministically growing populations in the bulk.

Exponential mutation kernel

The analysis of an exponential mutation kernel has been carried out in [GRB⁺12] and yields

$$\begin{aligned}\overline{v}/\sigma^2 &\approx \frac{\log^2\left(2N\mu\log(\sigma/\mu)\right)}{2\log\left(\frac{\sigma}{\mu}\frac{1}{\sqrt{\pi}}\frac{\log(2N\mu)}{\log^{1/2}(\sigma/\mu)}\right)} && (2.10a)\\ &\sim \frac{\log^2(N\mu)}{2\log(\sigma/\mu)}\,. && (2.10b)\end{aligned}$$

Here, the authors use a continuous description of the population density (as we will do in the next chapter). Similar as before, two (transcendental) equations in two unknowns are stated. One of these unknowns is the adaptation speed $\overline{v}$. The scaling (2.10a) recovered from the original equations (see [13] and [14] in [GRB⁺12]) deviates slightly from simulations, nevertheless yields a correct trend and correct order of magnitude. A numerical solution for the adaptation speed results in an almost perfect agreement with the simulation data presented in Figure 2.4.

The first of the two equations to determine $\overline{v}$ arises from an integral over the distribution of fixating mutations (see section 3.5.3). This yields the rate at which fixating mutations increase fitness, and the result is equated to the adaptation speed $\overline{v}$. The second equation is an integrated version of the equation of motion for the population density itself (integrating over the continuous version of (2.2), (2.3) and (2.4)). Besides the adaptation speed $\overline{v}$, the *stochastic edge* x_c is the second unknown. This stochastic edge is related to the lead q (in the staircase mutation kernel) by $x_c \approx \sigma q$, and denotes the fitness where only few individuals are expected, such that stochastic effects play a crucial role in the dynamics.

The authors also show that fixating mutations arise from a narrow range of fitness effects in this exponential DFE. The narrow range is small compared to the mean effect. This result has already been found earlier in extensive simulations [FND08]. Therefore, the approximation of the whole DFE with single steps of size σ, as in the staircase mutation kernel (2.9), with an appropriate rescaling of mutation rate and mutational effect, appears justified.

Diffusion mutation kernel

The diffusion mutation kernel has been treated in [TLK96, CKL05, Hal11b, NH13]. The results within those publications suggest a scaling of

$$\overline{v}D^{-2/3} \approx \left(24\log\left(ND^{1/3}\right)\right)^{1/3}\,. \qquad (2.11)$$

This result has been obtained with a governing (continuous) equation that is similar to the FKPP (Fisher-Kolmogorov-Petrovsky-Piscounoff)-equation [Fis37, KPP37], but with an additional cut-off on the growth term. The cut-off prohibits growth and sets it to zero if the density falls below a value of $1/N$. Matching the approximate solutions left and right of x_c (the stochastic edge, where the density reaches $1/N$) leads to the adaptation speed.

Comparing the result above to our simulation results for the adaptation speed shows that (2.11) overestimates the measured value by an additive constant. The relation (2.11) is strictly valid only for large population sizes. The offset seen in Figure 2.4 could be canceled using $\log\bigl(ND^{1/3}/10\bigr)$. This factor of $1/10$ inside the logarithm results in a good agreement of the scaling with simulation data (although it is only backed by this numerical estimate).

3. Constrained branching random walks as minimal models for adaptation in asexual populations

After the brief introduction of the simulation model and the basic phenomenology of traveling wave models in the last chapter, we turn our attention to a theoretical description of the adaptation process in this chapter. In order to describe the population we use *branching random walks* (BRWs).

BRWs combine two the concepts: first, they incorporate branching processes [All03, HJV05], that investigate the stochastic dynamics of a population of discrete replicating entities. The second concept are random walks in space. In BRWs entities replicate *and* move in space. The equations below will illuminate the particular choice of steps in the simulation algorithm from section 2.1.

3.1. Traveling wave models

We model adaptation as a branching random walk (BRW). The population is represented by a density in fitness space, where at time t the number of individuals in the fitness range $\left[x, x+dx\right]$ is given by $c_t(x)dx$. For small population sizes one can think of the population density $c_t(x)dx$ as counting individuals in different fitness classes x, which are intervals $\left[x, x+dx\right]$ with finite dx. The occupancies from the simulation algorithm in the last chapter are an example for this counting measure, $n_i = c_t(x_i)dx$. The continuous approximation $\mathbf{c}_t \equiv c_t(x)$ is justified when the population size is large enough, an approach which we will pursue here for our analytical treatment.

The description of the dynamics is given by a Langevin-equation [Gar09, Kam92], with a deterministic and a stochastic part. At first, time proceeds step-wise with steps of duration ϵ, only later we take the continuous limit $\epsilon \to 0$. The time evolution of the population density $\mathbf{c}_t$ is governed by selection $\mathcal{S}$, mutation $\mathcal{M}$ and genetic drift:

$$\mathbf{c}_{t+\epsilon} = \mathbf{c}_t + \epsilon \mathcal{S} \mathbf{c}_t + \epsilon \mathcal{M} \mathbf{c}_t + \sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t \ . \tag{3.1}$$

A single individual in the branching process dies with rate 1 and reproduces with rate $\left(1 + x - \overline{x_t}\right)$. $\overline{x_t} = \left(\int dx\ x c_t(x)\right)/N_t$ denotes the current mean fitness of the population (with N_t the current population size). The unit of time is one generation. During the time ϵ the population density changes (deterministically) with the difference in death and birth rates, $\left(x - \overline{x_t}\right)$, represented by the selection operator $\mathcal{S}\mathbf{c}_t \equiv \left(x - \overline{x_t}\right)c_t(x)$. Other mappings from the fitness class to (exponential) growth rate would be possible, but we restrict ourselves here on the simplest case, where the population is infinitely crawling up a *linear* fitness gradient.

The mutation operator $\mathcal{M}$ represents all changes in fitness due to novel mutations. A mutation of effect s changes the overall fitness of a clone in fitness class x to the new fitness class $(x+s)$. Epistasis is absent, no interaction between different mutations occurs. Hence, the growth rate is the only property of an individual: if two individuals are in the same fitness class, they are indistinguishable in terms of our model. The mutation operator is a continuous version of the mutation dynamics from the simulation model, details are given below in section 3.1.1.

Genetic drift is a purely stochastic force, proportional to $\sqrt{\epsilon}\boldsymbol{\eta}_t$ instead of the linear relation ϵ in the selection and mutation terms[1]. In branching processes the variance of reproduction noise in a fitness class x is proportional to the sum of death and birth rates of all individuals in the class, $\big(2+x-\overline{x_t}\big)c_t(x) \approx 2c_t(x)$. The explicit dependence on the fitness x can be omitted, as the spread of the population density in fitness space is very small compared to 1, a fact we have used when postulating (3.1). Typical fitness effects in the considered microbial populations are orders of magnitude below 1. $\boldsymbol{\eta}_t \equiv \eta_t(x)$ is a white noise (vector field), uncorrelated over time t and over different fitness classes x.

As a first consistency check with previous theoretical considerations, we can derive an expression equivalent to “Fisher’s fundamental theorem” [Fis30, Pri72]. In short, it relates the adaptation speed in a population to the variance in fitness. In order to derive the statement, we multiply the equation of motion (EOM) (3.1) with x/N_t, and subsequently integrate over fitness:

$$\partial_t \int dx\; x\big(c_t(x)/N_t\big) = \int dx\; x\left(x-\overline{x_t}\right)\big(c_t(x)/N_t\big) + \int dx\; x\mathcal{M}\big(c_t(x)/N_t\big) + \mathcal{O}\big(1/N\big) \tag{3.2}$$

Here, N_t denotes the current size of the population. The last term of order $\mathcal{O}\left(1/N\right)$ originates from genetic drift. Evaluating all terms, we have the “theorem”:

$$v = \mathrm{Var}\big[x|\mathbf{c}_t\big] + \mu\sigma + \mathcal{O}\big(1/N\big)\ . \tag{3.3}$$

Thus, adaptation speed v (change in mean fitness over time) is equal to the fitness variance of the population, $\mathrm{Var}\left[x|\mathbf{c}_t\right]$, with an additional possible offset due to mutations. The mutational load $\mu\sigma$ is given as the product of the mutation rate μ and the average effect of a new mutation σ, which can be either positive or negative due to sufficiently many beneficial mutations or an excess of deleterious mutations, respectively[2]. The expression $\mu\sigma$ for the mutational load will become apparent below, in section 3.1.1. In this general form, (3.3) is equivalent to either the *Price equation* [Pri70, Pri72] or also the *breeder’s equation* [NH13]. It should be noted that “Fisher’s fundamental theorem” is basically neither fundamental, nor a (mathematical) theorem [Hal13], that can be rigorously derived and is valid for arbitrary populations. As the term “Fisher’s fundamental theorem” is widespread, so we will continue to refer to it that way. However, even if (3.3) is not a fundamental theorem, it reflects an evolutionary principle: a more diverse population might be able to accustom to new environments better in the long run, compared to monoclonal populations.

---

[1] This proportionality to $\sqrt{\epsilon}$ represents the square root of the identity for the Wiener measure, $dW_t^2 = dt$, as used in stochastic calculus [Gar09, Kam92]. Here, the noise is considered in an Itō representation.

[2] In real populations the mutational load is most likely expected to be negative.

So far, the model has a major shortcoming: the only two possible long time behaviors are either extinction of the population or infinite growth. While the first one can still be argued to be an interesting feature of the model, even worth investigating, the latter one is clearly unnatural: resources in natural systems are limited. Often, the standard approach to this problem is setting the population size to a constant value,

$$\int dx \; c_t(x) \overset{!}{=} N \; , \tag{3.4}$$

which has been used throughout the population genetic literature in form of the ubiquitous Wright-Fisher- and Moran-models. We also employed this fixed population size constraint in the simulation model of chapter 2. This constraint will also serve as null model, to which we compare our generalized constraint below, in section 3.2.

Including the constraint (3.4), the solution to (3.1) takes the form of a bell-shaped *traveling wave*. Snapshots have already been shown in chapter 2. The traveling wave moves to either higher or lower mean fitness $\overline{x_t}$, depending on the delicate balance between the selection, mutation and genetic drift terms. An advance to higher fitness is fed by a sufficient supply of beneficial mutations, which grow faster than the bulk of the population, due to selection. It is exactly in this *nose* of the traveling wave, where the crucial processes of adaptation happen: newly occurring mutations in the fittest (or one of the fittest) individuals have to survive long enough to leave the regime where genetic drift dominates and reach a certain number of individuals such that they can grow exponentially via selection and reach fixation in the population. However, even very fit clones can go extinct due to genetic drift.

Our model is insofar a minimal treatment of adaptation (as advertised in the title of this chapter), as it includes the essential ingredients of the adaptation process: novel mutations enter the population, have to overcome genetic drift, thereafter grow faster than the bulk population due to selection, and (if they reach high enough numbers) their growth saturates due to limited resources. There are several aspects of real populations we have omitted here. A crucial simplification is that we do not distinguish different individuals with identical fitness. We also do not track their past lineage how they acquired this particular fitness value. This limitation implies that we neglect epistasis, *i.e.* interactions between different mutations in their effect on fitness. It has been observed that epistatic interactions occur in populations [KDBP11], but on a statistical level, these interactions that increase or decrease fitness tend to average out [KRJD14]. Another aspect that we ignore in our model, but is observed in real populations, are “diminishing returns” of novel mutations [VZG+99, CCD+11, KDS+11, Len15]. Over time, the effect of novel mutations declines when populations adapt to static environments. Our model could be adapted to incorporate such effects by either treating the mutation operator $\mathcal{M}$ or the selection operator $\mathcal{S}$ time-dependent.

Without the stochastic term, (3.1) is a standard equation considered for instance already in [CK70]. It has also been known as a form of the *replicator equation* [Now06].

In the context of *quasi-species theory* [Eig71, Wil05] it has been used to treat the dynamics of mutationally coupled genotypes. However, including the noise, a major problem arises: the combined equations, (3.1) and (3.4), are not analytically solvable, and solution schemes often rely on unjustifiable assumptions [Hal11b]. The non-linearity implicit in (3.4) leads to an unclosed hierarchy of n-point-correlation functions: the dynamics of the first moment $\overline{\mathbf{c}_t}$ couples to the second moment, $\overline{\mathbf{c}_t\mathbf{c}_t}$. In turn, the dynamics of $\overline{\mathbf{c}_t\mathbf{c}_t}$ itself couples to $\overline{\mathbf{c}_t\mathbf{c}_t\mathbf{c}_t}$, and so on. We present our solution to this problem in section 3.2.

3.1.1. Mutation models

Here, we will briefly revisit the three mutation kernels introduced for the simulation model in section 3.1.1. Contrary to before, we treat a continuous version here.

Mutations occur at a rate μ. The distribution of fitness effects s of a new mutation is given by the mutation kernel $\mathbf{m} \equiv m(s)$. As we ignore epistatic effects, a single mutation of effect s changes the fitness x of the individual to $\big(x+s\big)$. Hence, the "flow" of individuals from other fitness classes is given by the convolution of the population density $\mathbf{c}_t$ with the DFE $\mathbf{m}$:

$$\begin{aligned} \mathcal{M}\mathbf{c}_t &= \mu\left(\mathbf{m}-\boldsymbol{\delta}\right) * \mathbf{c}_t \\ &= \mu \int\limits_{-\infty}^{\infty} ds\; m(s)\big(c_t(x-s)-c_t(x)\big)\ . \end{aligned} \tag{3.5}$$

The singular contribution $\boldsymbol{\delta} \equiv \delta(s)$ accounts for all outgoing mutations of a fitness class x. A new mutation can occur on any of the available genetic backgrounds. The actual number of mutations on a given background is proportional to the number of individuals with this background. The mutation kernel is normalized,

$$\int ds\; m(s) = 1\ . \tag{3.6}$$

The overall rate of mutations is absorbed in the constant μ. The mean effect of new mutations will always be denoted by

$$\sigma \equiv \int ds\; s m(s)\ . \tag{3.7}$$

The expression for the mutational load in (3.3) follows as a consequence of (3.5) and (3.7): $\int dx\; x\mathcal{M}c_t(x) = \mu\sigma N_t$.

In the following, we will employ a Kramers-Moyal-expansion in the mutation term [NH13], which allows to classify and relate different models to each other. This approximation is justified, when most of the mass in the DFE is at small s values, such that

$c_t(x-s)$ can be expanded into its Taylor series:

$$\begin{aligned}
\mathcal{M}\mathbf{c}_t &= \mu \int ds\; m(s)\big(c_t(x-s) - c_t(x)\big) \\
&= \mu \int ds\; m(s) \left(\sum_{n=0}^{\infty} \frac{(-s)^n}{n!} \partial_x^n c_t(x) - c_t(x) \right) \\
&= \mu \sum_{n=1}^{\infty} \left(\frac{(-1)^n}{n!} \int ds\; s^n m(s) \right) \partial_x^n c_t(x) \; . \qquad (3.8)
\end{aligned}$$

The expression in brackets correspond to (scaled) moments of the DFE.

Staircase mutation kernel

In the staircase mutation kernel, only beneficial mutations of identical size occur [DF07]. The continuous version of this DFE $\mathbf{m}$ is given by

$$m(s)\, ds = \delta\big(s - \sigma\big)\, ds \; . \qquad (3.9)$$

For small σ we can approximate the mutation term

$$\mathcal{M}\mathbf{c}_t \equiv \mu\big(c_t(x-\sigma) - c_t(x)\big) = \mu\sigma \underbrace{\frac{c_t(x-\sigma) - c_t(x)}{\sigma}}_{\xrightarrow{\sigma\to 0} -\partial_x c_t(x)} \approx -\mu\sigma\, \partial_x c_t(x) \; , \qquad (3.10)$$

and arrive at first order derivative that represents the mutation process, which enhances the solvability of our equations, as will be seen later (see section 3.2.1).

Exponential mutation kernel

As mentioned in section 3.1.1, exponential DFEs have received considerable attention in the population genetics literature, both in experiments [IS01, KB06], but also theoretical considerations based on extreme value theory seem to yield an exponential DFE [Orr98, Orr03, Orr06, EWK07].

Here, the effect s of a beneficial mutation is distributed exponentially with parameter σ,

$$m(s)\, ds = \Theta(s) \exp\big(-s/\sigma\big)/\sigma \, ds \; . \qquad (3.11)$$

Deleterious mutations do not occur, indicated by the Heaviside-Theta function $\Theta(s)$. A Kramers-Moyal expansion yields

$$\mu \int ds\; m(s)\big(c_t(x-s) - c_t(x)\big) = \mu \sum_{n=1}^{\infty} (-\sigma)^n \partial_x^n c_t(x) \; . \qquad (3.12)$$

For small mean effects of mutations $\sigma \ll \sqrt{v}$, higher order terms can be neglected and the exponential DFE (3.11) can be approximated by the single peaked DFE in (3.9).

A whole family of exponential distributed DFEs and its effect on adaptive dynamics has been investigated in [GRB+12].

Diffusion mutation kernel

When mutations occur at high rates, but are almost neutral, the mutation process can be modeled by a diffusion process in fitness space,

$$\mu \int ds\ m(s)\big(c_t(x-s)-c_t(x)\big) \overset{!}{=} D\partial_x^2 c_t(x)\ . \tag{3.13}$$

In the Kramers-Moyal expansion above, the diffusion constant can be identified as $D = \mu/2 \int ds\ s^2 m(s) \approx \mu\sigma^2/2$. The diffusion kernel has a zero mean effect, $\int ds\ sm(s) = 0$, and all higher moments ($n \geq 3$) of the DFE are negligible. An example of such a DFE $\mathbf{m}$ would be a centered Gaussian distribution with (small) variance σ^2, *i.e.* $m(s) \sim \exp\big(-s^2/\left(2\sigma^2\right)\big)$, although others exists that would yield similar results. The diffusion model for mutation has been considered for example in [TLK96, CKL05, Hal11b, NH13].

3.2. Tuned models - employing a special population size constraint

A major problem of (unconstrained) BRWs is that they end in either extinction or infinite growth. One can use a non-linearity in the dynamics to keep populations from reaching these states. The fixed population size constraint in the chapter on simulations 2 is a prime example. However, such non-linearities induce another problem in stochastic analysis. The moment hierarchy of *correlation functions* $\overline{\mathbf{c}_t}$, $\overline{\mathbf{c}_t\mathbf{c}_t}$, $\overline{\mathbf{c}_t\mathbf{c}_t\mathbf{c}_t}$, … does not close: the time-evolution of the first moment $\overline{\mathbf{c}_t}$ couples to the second moment $\overline{\mathbf{c}_t\mathbf{c}_t}$, while the dynamics of $\overline{\mathbf{c}_t\mathbf{c}_t}$ depends on $\overline{\mathbf{c}_t\mathbf{c}_t\mathbf{c}_t}$, and so on.

In this section we present an approach that allows to circumvent the problem of this infinitely coupled moment hierarchy. To this end, we note that in [Hal11b] the constraint (3.4) has been generalized by using a (not yet specified) function $\mathbf{u}_t$ to keep the population from reaching either extinction or infinite population sizes (using Quantum mechanical notation to represent integrals):

$$\int dx\ u_t(x)c_t(x) = 1 \equiv \big\langle \mathbf{u}_t \big| \mathbf{c}_t \big\rangle\ . \tag{3.14}$$

Clearly, setting $u_t(x) = 1/N$ recovers the fixed population size constraint (3.4). For other (possibly time-dependent) $\mathbf{u}_t$, the population size N_t will fluctuate due to the stochasticity in the model. *A priori*, different choices of $\mathbf{u}_t$ can give rise to different models. It has been argued, however, that the specific form of a non-linearity, which is implicit in (3.14), does not influence the front dynamics of an advancing traveling wave too much [Saa03]. Furthermore, effects of the non-linearity influence predominantly the bulk population, where the large number of individuals has to be constrained and their growth saturates. For processes in the nose of the wave, which are crucial for the adaptation dynamics, the saturation effects are almost irrelevant. It should also be noted that a fluctuating population size is no less realistic than a fixed population size constraint.

In the following, we describe the dynamics of *constrained* BRWs, where the constraint (3.14) is already coupled to the time-evolution. In order to ensure the constraint (3.14) at all times, the population density $\mathbf{c}_t$ evolves during a time step ϵ as

$$\mathbf{c}_{t+\epsilon} = \frac{\mathbf{c}_t + \epsilon\big(\mathcal{M}+\mathcal{S}\big)\mathbf{c}_t + \sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t}{\big\langle \mathbf{u}_t + \epsilon\partial_t\mathbf{u}_t \big| \mathbf{c}_t + \epsilon\big(\mathcal{M}+\mathcal{S}\big)\mathbf{c}_t + \sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\big\rangle} \,, \tag{3.15}$$

where we used $\mathbf{u}_{t+\epsilon} = \mathbf{u}_t + \epsilon\partial_t\mathbf{u}_t$. $\big\langle \mathbf{u}_{t+\epsilon}\big|\mathbf{c}_{t+\epsilon}\big\rangle = 1$ follows immediately. From now on we define the infinitesimal generator of the deterministic part of the dynamics as $\mathcal{L} \equiv \big(\mathcal{M}+\mathcal{S}\big)$, a Liouvillean. An expansion of (3.15) to orders $\mathcal{O}\big(\epsilon\big)$ and $\mathcal{O}\big(\sqrt{\epsilon}\big)$ yields

$$\begin{aligned}\mathbf{c}_{t+\epsilon} - \mathbf{c}_t \;\;=\;\; & \epsilon\Big[\mathcal{L}\mathbf{c}_t - \sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\big\langle \mathbf{u}_t\big|\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\big\rangle - \mathbf{c}_t\big\langle\partial_t\mathbf{u}_t\big|\mathbf{c}_t\big\rangle - \mathbf{c}_t\big\langle\mathbf{u}_t\big|\mathcal{L}\mathbf{c}_t\big\rangle + \mathbf{c}_t\big\langle\mathbf{u}_t\big|\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\big\rangle^2\Big] \\ & +\sqrt{\epsilon}\Big[\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t - \mathbf{c}_t\big\langle\mathbf{u}_t\big|\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\big\rangle\Big] + \mathcal{O}\big(\epsilon^{3/2}\big) \,. \end{aligned} \tag{3.16}$$

Averaging over the noise, $\overline{\eta_t(x)} = 0$ and $\overline{\eta_t(x)\eta_s(y)} = \delta_{ts}\delta(x-y)$, the term $\mathcal{O}\big(\sqrt{\epsilon}\big)$, representing the noise, cancels in the equation above. Dividing by ϵ, we arrive at the differential equation for the mean population density $\overline{\mathbf{c}_t}$,

$$\partial_t\overline{\mathbf{c}_t} = \big(\mathcal{L} - 2\mathbf{u}_t\big)\overline{\mathbf{c}_t} - \big\langle\big(\partial_t + \mathcal{L}^\dagger - 2\mathbf{u}_t\big)\mathbf{u}_t\big|\overline{\mathbf{c}_t\mathbf{c}_t}\big\rangle \,, \tag{3.17}$$

where we used the adjoint operator $\big(\mathcal{L}^\dagger - 2\mathbf{u}_t\big)$ under the integral. Note that the dynamics of the first moment couples to the second moment of the population density, $\overline{\mathbf{c}_t\mathbf{c}_t}$[3].

However, as has been originated in [Hal11b], we can "tune" the model by choosing a specific constraint $\mathbf{u}^\star_t$. When choosing $\mathbf{u}_t = \mathbf{u}^\star_t$ such that

$$-\,\partial_t\mathbf{u}^\star_t \overset{!}{=} \big(\mathcal{L}^\dagger - 2\mathbf{u}^\star_t\big)\mathbf{u}^\star_t \,, \tag{3.18}$$

the moment hierarchy in (3.17) decouples or closes (at the first moment) and we arrive at a linear equation for the mean population density $\overline{\mathbf{c}_t}$,

$$\partial_t\overline{\mathbf{c}_t} = \big(\mathcal{L} - 2\mathbf{u}^\star_t\big)\overline{\mathbf{c}_t} \,. \tag{3.19}$$

When comparing (3.19) with the original expansion (3.16), it becomes apparent that the additional term $-2\mathbf{u}^\star_t$ is a result from the influence of stochasticity (genetic drift) and the finite population size constraint in the model. In particular, the second term in (3.16) is of order $\sim \boldsymbol{\eta}_t\boldsymbol{\eta}_t$ and averages to cancel the integral denoted by angular brackets.

In the remainder of this chapter we use this formalism to derive crucial features of the adaptation dynamics, as the population density is governed only by a *linear* operator. We argue that employing this tuned constraint does not change the overall dynamics qualitatively. Even quantitative predictions obtained with tuned models comply with predictions obtained by other methods. This will be shown in chapters 4 and 5, when we compare simulations of tuned models with the fixed population size simulations from chapter 2. However, it has been speculated that the fluctuations in tuned model are in a different universality class [Fis13].

---

[3] A comment on notation is in order here: multiplication of vector quantities in bold (*e.g.* the functions $\mathbf{u}_t$ or $\overline{\mathbf{c}_t}$) is considered point-wise: $\mathbf{u}_t\overline{\mathbf{c}_t} \equiv u_\star(x)\overline{c_t(x)}$. However, correlation functions denoted with a joint overbar (*e.g.* $\overline{\mathbf{c}_t\mathbf{c}_t}$) have *independent* fitness variables in each dimension: $\overline{\mathbf{c}_t\mathbf{c}_t} \equiv \overline{c_t(x)c_t(y)}$. In (3.17), the second variable of the correlation function is used in the integral denoted by angular brackets, while the first variable x is identical for all terms.

3.2.1. Solutions to the non-linear constraint equation

The equation for the dynamics of the constraint, (3.18), is a non-linear partial differential equation. Such equations are usually incredibly hard to analyze. Nevertheless, some approximations can be made, and in this section we present asymptotic solutions to different ranges in fitness. A global solution to (3.18) can be obtained numerically. In appendix B.1 we present algorithms for this task. Equations similar to (3.18) (or also in the form (3.21) below) have also been considered in [NSF10, Hal11b, GRB$^+$12, Fis13, NH13]. There, approximations explore solutions $\mathbf{u}_t^\star$ locally for different ranges of the fitness, as we will do below. We use the obtained shape of $\mathbf{u}_t^\star$ in the next section, 3.3, to interpret this solution, providing insight into the mechanisms underlying the concepts of adapting populations.

Knowing that (3.18) admits traveling wave solutions, we introduce a comoving frame to reduce the PDE (in variables x and t) to an ODE. We fix the adaptation speed v of the population. The single coordinate in fitness space is then

$$X = x - vt \;, \tag{3.20}$$

where X describes the relative fitness of individuals with respect to the population mean, *i.e.* X replaces $\left(x - \overline{x_t}\right)$ in the averaged equations. A solution in this comoving frame will always be denoted with a lower asterisk, $\mathbf{u}_\star \equiv u_\star(x - vt) = u_t^\star(x)$. In general there are more valid solutions to (3.18) than those employing the coordinate transform (3.20), each of them closing the moment hierarchy. However, in the following we will mostly be concerned with solutions in the comoving frame. Notice that our restriction allows us to perform a detailed investigation of the steady state in traveling waves, but we cannot describe transients in this simplified setup. The time-derivative transforms to $\partial_t = -v\partial_X$, and (3.18) becomes

$$v\partial_X \mathbf{u}_\star = X\mathbf{u}_\star + \mathcal{M}^\dagger \mathbf{u}_\star - 2\mathbf{u}_\star^2 \;. \tag{3.21}$$

Here, the selection operator reduces to a simple multiplication with X, the (relative) fitness in the comoving frame. The next step to obtain an approximate solution is to apply the Kramers-Moyal expansion on the mutation term[4]:

$$\begin{aligned}
\mathcal{M}^\dagger \mathbf{u}_\star &= \mu \int ds\; m(s)\big(u_\star(X+s) - u_\star(X)\big) \\
&= \mu \int ds\; m(s)\Big(\big(u_\star(X) + s\partial_X u_\star(X) + \mathcal{O}\big(s^2\big)\big) - u_\star(X)\Big) \\
&= \mu\sigma\, \partial_X u_\star(X) + \mathcal{O}\big(\sigma^2\big)
\end{aligned} \tag{3.22}$$

---

[4] Note that in the *adjoint* operator $\mathcal{M}^\dagger$, fitness is increased in the argument of $u_\star(X)$ instead of decreased, like in (3.5).

A truncation of the Kramers-Moyal expansion after the first order is valid as long as the mean effect of a mutation σ is small enough. When the variance of the population, $\mathrm{Var}\big[x|\mathbf{c}_t\big]$, is large compared to σ^2, this condition is met. Defining $V \equiv v - \mu\sigma$, we arrive at

$$V\partial_X \mathbf{u}_\star \approx X\mathbf{u}_\star - \alpha \mathbf{u}_\star^2 \,. \tag{3.23}$$

Using Fisher's theorem, (3.3), we see that V is the fitness variance $\mathrm{Var}\big[x|\mathbf{c}_t\big]$. (3.23) is solved by

$$u_\star(X) = \frac{\exp\big(X^2/(2V)\big)}{N + \alpha\sqrt{\pi/(2V)}\mathrm{Erfi}\big(X/\sqrt{2V}\big)} \;, \quad \alpha = 2\,. \tag{3.24}$$

Erfi denotes the imaginary error function. The parameter α introduced here will be used later in a generalized context. N is an integration constant[5]. Note that we get a whole family of solutions to (3.23) due to its non-linearity, $\mathbf{u}_\star = u_\star(X, N)$. Any choice of N is a valid solution to (3.23). In order to determine a unique solution $\mathbf{u}_\star$ we have to include higher orders of the DFE, $\int ds\; s^n m(s),\;\; n \geq 2$, in the Kramers-Moyal expansion.

Next, we turn our attention to various approximation schemes, which have been used throughout literature cited above to solve equations equivalent to (3.18) (or (3.21)). In order to do so, we define three different regimes of the solution. These regimes are separated by the blue lines in Figure 3.1. Increasing in fitness, we have regime (I) for negative values of X up to a (small) positive value. Then, regime (II) describes an intermittent range of fitness, until for large (positive) fitnesses in regime (III) we have an asymptotic linear solution. In each of these regimes, different terms in (3.21) are small and can be neglected. The approximations below use this fact.

In regime (III), the dominant terms in (3.21) are $0 \approx X\mathbf{u}_\star - 2\mathbf{u}_\star^2$. This allows a (local) characterization of the solution as $u_\star^{(III)}(X) \sim X/2$. Our solution, (3.24), also exhibits this behavior: in the asymptotic expansion, $\exp\big(Y^2\big)/\mathrm{Erfi}\big(Y\big) \approx Y/\sqrt{\pi}$ for $Y \gg 1$, we also obtain a linear solution. The transition to regime (II) occurs, when both terms in the denominator are of the same order of magnitude, *i.e.* $N \approx 2\sqrt{\pi/(2V)}\mathrm{Erfi}\big(X_c/\sqrt{2V}\big)$.

---

[5] We have to anticipate later results to interpret N: note that $u_\star(0) = 1/N$. In classical theories [Gil10], the fixation probability of a neutral allele is proportional to the inverse population size. We will see the next section 3.3 that $\mathbf{u}_\star$ is in fact a fixation probability. In addition to the approximation (3.24), we also have numerical solutions to the full equation, (3.21). Fitting the two parameters V and N in (3.24) to exact solutions shows that there is an excellent agreement of the adaptation speed – or, population variance – parameter. The major contribution to the integration constant *is* the population size N, but there are deviations: $\log u_\star(0) \approx -\log N + c$. c is much smaller than $\log N$ but still of order $\mathcal{O}\big(1\big)$. The value of c depends on the other parameters population size N, mutation rate μ, mutational scale σ and the choice of the mutation kernel. However, it is still reasonable to interpret the integration constant N as population size in the model, indicated by the choice of its symbol.

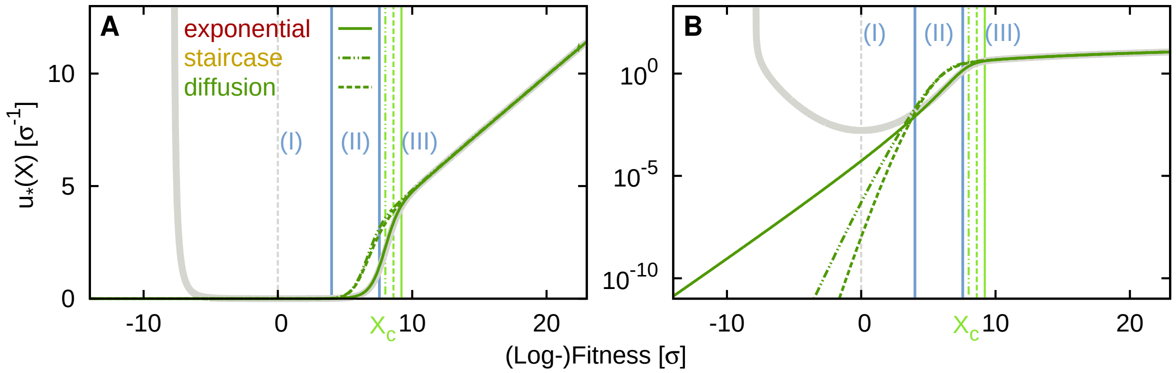

Figure 3.1.: **General shape of the solutions $\mathbf{u}_\star$ in (A) normal and (B) logarithmic scale.** Dark green lines depict exact numerical solutions to (3.21), while the gray line is the approximate solution (3.24). Three different regimes in the solution are separated by the two blue lines. While the second blue line, corresponding to X_c, is obtained by (3.25), three lime green lines the depict more accurate estimates of X_c, by using the condition that $\mathbf{u}_\star$ dropped below 90% of the linear solution. In the rightmost regime (III), with $X > X_c$, the solution is linear, $u_\star^{(III)}(X) \approx X/2$. Regime (II) is defined below X_c down to $X \approx V/\sigma$. There, a steep increase in the solution occurs, consistent with the approximate solutions $u_\star^{(II)}(X) \sim \exp\big(X^2/(2V)\big)$ (as part of (3.24)). Below $X \approx V/\sigma$, the solution decays exponentially with decreasing fitness X (defined as regime (I)).

Inverting this condition [Win08] leads to an approximation for X_c,

$$\begin{aligned} X_c \;&\approx\; \sqrt{2V}\sqrt{b + \log\big(N\sqrt{V}\big) - \sqrt{\Big(b + \log\big(N\sqrt{V}\big)\Big)^2 - 1/a \log\left(\frac{N\sqrt{V}}{\sqrt{2\pi}}\right)}} \\ &\sim\; \sqrt{2V\log\big(N\sqrt{V}\big)} + \mathcal{O}\left(\sqrt{V/\log\big(N\sqrt{V}\big)}\right) \;, \end{aligned} \tag{3.25}$$

with numerical constants $a = \frac{8}{3\pi}\frac{\pi-3}{4-\pi} \approx 0.140$ and $b = \frac{2}{\pi a} - \log\sqrt{2\pi} \approx 3.628$.

In regime (II), $\mathbf{u}_\star$ is small compared to its solution in (III), hence the quadratic term can be neglected. We have to solve $V\partial_X\mathbf{u}_\star \approx X\mathbf{u}_\star$, arriving at the solution $u_\star^{(II)}(X) \sim \exp\big(X^2/(2V)\big)$. Again, our solution (3.24) predicts the shape for this fitness range: in the denominator, the value of Erfi is small compared to the integration constant N, the solution is dominated by the nominator, $\sim \exp\big(X^2/(2V)\big)$.

Characterizing regime (I) requires more thought than for (II) and (III). Values of $\mathbf{u}_\star$ are tiny in this regime, and it has been set to zero *e.g* in [GRB+12]. Usually this zero-solution suffices, as regime (I) is not essential in the adaptation process (see next section).

Nevertheless, we want to improve on this (trivial) asymptotic solution. The mutational term, which we have ignored (or absorbed into the derivative of the comoving frame with a first order Kramers-Moyal expansion) contributes substantially in this regime. In appendix C.1 we present calculations for the *exponential* mutation kernel using an integral transform. This transformation reduces the original integro-differential equation (3.21) to a simple ODE. The solution predicts the cross-over from regime (I) to regime (II) at $X_c^{(I)\to(II)} = v/\sigma$ (for the exponential mutation kernel). Moreover, the main contribution to $\mathbf{u}_\star$ in regime (I) is an exponential increase: $u_\star^{(I)}(X) = \exp\big(X/\sigma\big) f(X)$, where $f(X)$ is a slowly varying function in X. This exponential tail is also found numerically, see Figure 3.1. In [Fis13], a more careful analysis based on the log-slope $\boldsymbol{\gamma} = \partial_X \log \mathbf{u}_\star$ and $u_\star(X) \approx X_c \exp\big(-\int_X^{X_c} dy\ \gamma(y)\big)$, extracts also the next to leading order, with exponents $X/\sigma \log\big(-X/\mu\big)$ for the staircase model and $XD^{-1/3}\left(2\log\big(-XD^{-1/3}\big)\right)^{1/2}$ for the diffusion mutation kernel.

Figure 3.1 depicts the numerical solution to (3.21) for all three mutation kernels. We also show the first order Kramers-Moyal solution (3.24), establishing it as a good approximation for regimes (II) and (III). However, the problem with (3.24) becomes apparent: Erfi is an odd function, hence at $X = -X_c$ it exhibits a singularity: then $N = -2\sqrt{\pi/(2V)}\mathrm{Erfi}\big(-X_c/\sqrt{2V}\big)$. Therefore (and for reasons stated next section, 3.3), we have to discard the approximation (3.23) in regime (I)[6]. The advantage of (3.24) is that it (approximately) solves regimes (II) and (III) simultaneously. Previously, the approximate solutions only allowed a piece-wise characterization of $\mathbf{u}_\star$[7].

3.3. Interpretation of the tuned constraint

A crucial open question is now if we can assign a relevant interpretation to $\mathbf{u}_\star$. We have shown that using this particular constraint renders the dynamics of constrained BRWs *linear* and decouples the moment hierarchy. However, this comes with the cost of having to solve an additional non-linear PDE (or ODE). From the last section, we also know the shape of the function $\mathbf{u}_\star$ for different fitness regimes. Its meaning in the context of tuned models is yet elusive.

We will show in this section that $u_\star(X)$ is the *fixation probability* of an individual with (relative) fitness X. Therefore, the constraint $\langle \mathbf{u}_\star | \mathbf{c}_t \rangle = 1$, (3.14), is the (obvious) statement that exactly *one* individual will fixate: it will be the ancestor of the whole population at some time point in the future. Having a single ancestor for *all* individuals in the current (finite) population is an absorbing state: in stochastic models that we are treating here, this absorbing state will be reached.

With this knowledge, we can assign different interpretations to the three regimes (shown in Figure 3.1): In regime (III), an individual has such an extraordinary high fitness that it just has to overcome genetic drift[8]. Classical theories [Gil10] predict that the probability for an allele to overcome genetic drift is proportional to its fitness advantage X in accordance with the linear regime we obtained. In regime (II), the effect of clonal competition is strong. Many individuals of same or similar fitness exist, and just by chance one of them is lucky enough to be picked as ancestor for the future population. They interfere with each others fixation, as we consider only asexual reproduction. Therefore, the fixation probability (for a single individual) drops significantly below X_c. Furthermore, we see why regime (I) is not as important as (II) and (III) in the dynamics of adaptation. It is very unlikely that an individual with such a low fitness background receives enough beneficial mutations to outgrow all others in the large populations we consider here.

---

[6] The solution $\mathbf{u}_\star$ should not increase with decreasing fitness. Its interpretation in the context of our model demands that it is a monotonic increasing function in X, see next section.

[7] Note added in proof: the same solution has been published in [GD14].

Before going into technical details, we present a short heuristic argument for $\mathbf{u}_\star$ being a fixation probability. Recall the equations of motion of $\overline{\mathbf{c}_t}$ and $\mathbf{u}^\star_t$,

$$\begin{aligned} \partial_t \overline{\mathbf{c}_t} &= \big(\mathcal{L} - 2\mathbf{u}^\star_t\big)\overline{\mathbf{c}_t} \,, && \text{(3.26a)} \\ -\partial_t \mathbf{u}^\star_t &= \big(\mathcal{L}^\dagger - 2\mathbf{u}^\star_t\big)\mathbf{u}^\star_t \,. && \text{(3.26b)} \end{aligned}$$

The first equation (3.26a) is the forward-in-time evolution of the population density. Sloppily, we can state its meaning as “where individuals in the population will go to”. Adjungating a differential operator reverses its time-direction. Hence, we see that the second equation (3.26b) is a backward-in-time dynamics with the same *expectation value* as (3.26a) (albeit time-reversed). Similar to before, we can state that it describes “where individuals in the population come from”. In a steady state, this statement on the origin of individuals is equivalent to a statement on their fixation when reversing time (again to forward dynamics). Several, more rigorous, derivations follow below.

3.3.1. Dynamics of sub-populations

The first method to assign the meaning of a fixation probability to $\mathbf{u}_\star$ is to investigate the dynamics of sub-populations within the whole population. In this section we use the Green's operator of the (mean) forward dynamics to accomplish this task. We define a (labelled) sub-population $\mathbf{w}_t$, consisting of one or more individuals and spread over one or more fitness classes, $\mathbf{w}_t \leq \mathbf{c}_t$. See Figure 3.2 for an example where the sub-population $\mathbf{w}_t$ is the orange part of the overall population density. In the stochastic dynamics the only two long time behaviors are extinction of the sub-population or a complete take over, such that all individuals in a later population are descendants of $\mathbf{w}_t$. To derive the dynamics of $\mathbf{w}_t$, we use the constrained generator (3.15) again. The sub-population $\mathbf{w}_t$ itself follows the forward dynamics of BRWs, but the constraint acts *globally* on the whole population:

$$\mathbf{w}_{t+\epsilon} = \frac{\mathbf{w}_t + \epsilon\mathcal{L}\mathbf{w}_t + \sqrt{\epsilon}\mathcal{D}\mathbf{w}_t}{\langle \mathbf{u}_t + \epsilon\partial_t\mathbf{u}_t | \mathbf{c}_t + \epsilon\mathcal{L}\mathbf{c}_t + \sqrt{\epsilon}\mathcal{D}\mathbf{c}_t\rangle} \,. \tag{3.27}$$

---

[8] The mean population density $\overline{\mathbf{c}_t}$ (solution to (3.17)) passes the threshold $\overline{c_t(X_c)} \approx \mathcal{O}\big(1/\sigma\big)$ at X_c: a fitness class of width σ would contain roughly one individual. Individuals with fitnesses beyond this value are unlikely to be present, although it can happen due to stochasticity.

Here, we introduced additional notation that will enhance readability in later sections: the (non-linear) genetic drift operator $\mathcal{D}\mathbf{w}_t \equiv \sqrt{2w_t(x)}\xi_t(x)$ denotes stochastic terms. The white noise (field) $\boldsymbol{\xi}_t$ is different from the white noise $\boldsymbol{\eta}_t$ before, insofar as it is the stochasticity of *only* the sub-population $\mathbf{w}_t$.

After expanding (3.27) to orders $\mathcal{O}\big(\epsilon\big)$ and $\mathcal{O}\big(\sqrt{\epsilon}\big)$, we arrive at

$$\begin{aligned}\mathbf{w}_{t+\epsilon} - \mathbf{w}_t \;&=\; \epsilon \left[\mathcal{L}\mathbf{w}_t - \mathcal{D}\mathbf{w}_t\big\langle \mathbf{u}_t|\mathcal{D}\mathbf{c}_t\big\rangle - \mathbf{w}_t\big\langle \partial_t\mathbf{u}_t|\mathbf{c}_t\big\rangle - \mathbf{w}_t\big\langle \mathbf{u}_t|\mathcal{L}\mathbf{c}_t\big\rangle + \mathbf{w}_t\big\langle \mathbf{u}_t|\mathcal{D}\mathbf{c}_t\big\rangle^2\right] \\ &\quad\; \sqrt{\epsilon}\left[\mathcal{D}\mathbf{w}_t - \mathbf{w}_t\left\langle \mathbf{u}_\star|\mathcal{D}\mathbf{c}_t\right\rangle\right] + \mathcal{O}\left(\epsilon^{3/2}\right) \;. \qquad (3.28)\end{aligned}$$

Upon averaging over the noise, the last three terms for the order $\mathcal{O}\big(\epsilon\big)$ in the expansion (3.28) yield the coupling to higher moments $\overline{\mathbf{w}_t\mathbf{c}_t}$. In particular, we arrive at $\big\langle(\partial_t + \mathcal{L}^\dagger - 2\mathbf{u}_t)\mathbf{u}_t|\overline{\mathbf{w}_t\mathbf{c}_t}\big\rangle$. This term vanishes when inserting $\mathbf{u}_t^\star$ to close the moment hierarchy.

In the average, the second term in (3.28) combines the noise terms of the sub-population and its counterpart, the "non"-sub-population. There, we compute

$$\begin{aligned}\overline{\mathcal{D}\mathbf{w}_t\big\langle \mathbf{u}_t|\mathcal{D}\mathbf{c}_t\big\rangle} \;&=\; \overline{\sqrt{2w_t(x)}\xi_t(x)\int dy\; u_t(y)\sqrt{2c_t(y)}\eta_t(y)} \\ &=\; \overline{\sqrt{2w_t(x)}\xi_t(x)\int dy\; u_t(y)\sqrt{2w_t(y)}\xi_t(y)} \\ &=\; 2u_t(x)\overline{w_t(x)} \\ &=\; 2\mathbf{u}_t\overline{\mathbf{w}_t} \;. \qquad (3.29)\end{aligned}$$

From the first to second line, we replaced the noise term of the complete density $\sqrt{\mathbf{c}_t}\boldsymbol{\eta}_t$ with $\sqrt{\mathbf{w}_t}\boldsymbol{\xi}_t$ in the integral: the noise of individuals in $\mathbf{c}_t$ in excess of $\mathbf{w}_t$ cancels to zero in the average. Correlations in the white noise occur *only* for indistinguishable individuals – and we can distinguish between the sub-population $\mathbf{w}_t$ and the "non"-sub-population $(\mathbf{c}_t - \mathbf{w}_t)$. To evaluate the integral in the second line we used the property of white noise, $\overline{\xi_t(x)\xi_s(y)} = \delta_{ts}\delta(x-y)$, after which the integration disappears.

Thus, after averaging over the noise and after inserting $\mathbf{u}_t^\star$ to close the moment hierarchy, we have identical operators acting on both *mean* densities, $\overline{\mathbf{c}_t}$ and $\overline{\mathbf{w}_t}$. From now on we continue with these mean quantities. We introduce the same comoving frame as in section 3.2.1 ($X = x - vt$, $\partial_t = -v\partial_X$) and arrive at

$$0 \;=\; \big(\mathcal{L} - 2\mathbf{u}_\star + v\partial_X\big)\overline{\mathbf{c}_\star} \;, \qquad (3.30a)$$

$$0 \;=\; \big(\mathcal{L} - 2\mathbf{u}_\star + v\partial_X\big)\overline{\mathbf{w}_\star} \;. \qquad (3.30b)$$

We define $\overline{\mathbf{c}_\star}$ as the mean population density in the comoving frame. Subsequently we employ a spectral decomposition of the operators in (3.30a) and (3.30b).The eigenvalue to $\overline{\mathbf{c}_\star}$, $\lambda = 0$, is the largest eigenvalue of the spectrum $\sigma(\lambda)$, all other eigenvalues are negative, $\lambda \leq 0$[9]. Over time, the average of the population density $\overline{\mathbf{c}_t}$ will decay to $\overline{\mathbf{c}_\star}$ with time scales given by the eigenvalues of the operator in (3.30a).

---

[9] This fact, $\sigma(\lambda) \leq 0$, arises from forcing the adaptation speed to the value of the comoving frame v and denying the population $\overline{\mathbf{c}_t}$ to reach its absorbing states, extinction or infinite growth. In the non-comoving frame, we *do not* have this condition on the spectrum.

In the comoving frame, the time-evolution from time t to time τ of the population densities can be stated by using the Green's operator $\mathcal{G}$, $\overline{\mathbf{w}_\tau} = \mathcal{G}\overline{\mathbf{w}_t}$. Together with the spectral decomposition of $\mathcal{G}$ we have

$$\mathcal{G}\,\overline{\mathbf{w}_t} \equiv \int d\sigma(\lambda)\; e^{\lambda(\tau-t)} \big\langle \boldsymbol{\zeta}^{(L,v)}_{\lambda,t} | \overline{\mathbf{w}_t} \big\rangle \boldsymbol{\zeta}^{(R,v)}_{\lambda,\tau} \,, \tag{3.31}$$

where $d\sigma(\lambda)$ is the measure on eigenvalues λ of the dynamics (3.30a) and (3.30b). $\boldsymbol{\zeta}^{(L,v)}_{\lambda,t}$ and $\boldsymbol{\zeta}^{(R,v)}_{\lambda,\tau}$ are the corresponding left (L) and right (R) eigenfunctions to eigenvalue λ. The other indices denote the comoving frame at speed v and times t or τ, respectively. On long times, $\tau - t \to \infty$, only the eigenfunction to $\lambda = 0$ survives, as all others decay exponentially. The Green's operator in (3.31) then reduces to

$$\mathcal{G}\,\overline{\mathbf{w}_t} \overset{\tau-t\to\infty}{\longrightarrow} \big\langle \boldsymbol{\zeta}^{(L,v)}_{0,t} | \overline{\mathbf{w}_t} \big\rangle \boldsymbol{\zeta}^{(R,v)}_{0,\tau} \,. \tag{3.32}$$

In order to evaluate this expression, we need explicit solutions for the eigenvectors to $\lambda = 0$. However, we have them already: by investigating (3.30a) we see that $\overline{\mathbf{c}_\star}$ is the right eigenvector to eigenvalue zero, $\boldsymbol{\zeta}^{(R,v)}_{0,\tau} = \overline{\mathbf{c}_\star}$. To determine left eigenvectors, operators have to be adjungated and the resulting equation solved. Hence, we also have $\boldsymbol{\zeta}^{(L,v)}_{0,t} = \mathbf{u}_\star$, as $\mathbf{u}_\star$ is the solution to eigenvalue zero to the equation governed by the adjungated operator, see (3.21). Using these two solutions, the large time limit in $\mathcal{G}$ can be written as

$$\mathcal{G}\,\overline{\mathbf{w}_t} \overset{\tau-t\to\infty}{\longrightarrow} \big\langle \mathbf{u}_\star | \overline{\mathbf{w}_t} \big\rangle \overline{\mathbf{c}_\star} \,. \tag{3.33}$$

At the start of this section we mentioned that in a single realization of the dynamics of the sub-population $\mathbf{w}_t$[10], the descendants of $\mathbf{w}_t$ either go extinct or become ancestors to the complete population. Denoting the probability for this second scenario as the fixation probability F of the sub-population $\mathbf{w}_t$, the long time limit can be written as

$$\mathbf{w}_t \overset{\tau-t\to\infty}{\longrightarrow} F\,\overline{\mathbf{c}_\star} + (1-F)\,\mathbf{0} \,, \tag{3.34}$$

where the right hand side is a superposition of probabilities of the respective states. Comparing the last relation (3.34) with the long time limit of the Green's operator, (3.33), we can identify

$$F = \big\langle \mathbf{u}_\star | \overline{\mathbf{w}_t} \big\rangle \,. \tag{3.35}$$

Therefore, $\mathbf{u}_\star$ is interpreted as the probability of fixation for an individual of fitness X. Establishing this statement has been the central aim of this section. It can be checked by inserting a single individual with fitness X_0 into the population, $w_t(X) = \delta(X - X_0)$. This leads to $u_\star(X_0)$ as is its fixation probability. If $\mathbf{w}_t$ consists of more than one individual, then each of those individuals fixates independent from all others. The overall fixation probability of the sub-population is the sum of the particular fixation probabilities of each individual within the sub-population.

---

[10] Here without overbar as we consider a single realization of the process!

The argument involving the Green's operator has already been stated in a comparable form in the supplementary material of [Hal11b].

As a last remark we derive the stochastic time evolution of the fixation probability F_t. In particular, we want to obtain the expectation value $\mathbb{E}\big[F_{t+\epsilon}\big]$ of the fixation probability F_t after a (stochastic) time step ϵ. In order to do so, we multiply (3.27) with $\mathbf{u}_\star$ and integrate over x. In the resulting equation the deterministic terms of order $\mathcal{O}\big(\epsilon\big)$ vanish and we arrive at

$$F_{t+\epsilon} = F_t + \sqrt{\epsilon}\Big[-F_t \left\langle \mathbf{u}_\star | \mathcal{D}\mathbf{v}_t \right\rangle + (1-F_t) \left\langle \mathbf{u}_\star | \mathcal{D}\mathbf{w}_t \right\rangle\Big] + \mathcal{O}\left(\epsilon^{3/2}\right) \ . \tag{3.36}$$

As the white noise terms $\mathcal{D}\mathbf{w}_t$ and $\mathcal{D}\mathbf{v}_t$ both have zero mean (in each infinitesimal fitness class), we see that $\mathbb{E}\big[F_{t+\epsilon}\big] = \mathbb{E}\big[F_t\big]$. This allows the interpretation of $F_t = \langle \mathbf{u}_\star | \mathbf{w}_t \rangle$ as the (current, at time t) expected probability of fixation for the sub-population $\mathbf{w}_t$.

3.3.2. Fixation probabilities in unconstrained BRWs

The probability for fixation of an allele in a population has been a central question in population genetics since the inception of the field in the early 20th century. In this section, we derive extinction and fixation probabilities in *unconstrained* BRWs. The derivation is based on standard branching process theory [All03, HJV05].

First, the focus is on a single fitness class, and we derive the generating function for all moments of its occupancy. The extinction probability follows directly from this calculation. In turn, we arrive at an equation that is similar to the dynamics of our constraint $\mathbf{u}_\star$. Generalizing these concepts to the complete population, we see again that the backward-in-time dynamics of the complete generating functional resembles (in part) the equation of motion for our constraint, $\mathbf{u}_\star$. At the end of this section, we will treat the differences between the calculations in this section on *unconstrained* BRWs and our tuned models, which are a special case of a *constrained* BRWs.

Generating function for moments of the occupancy in a fitness class

At first, we follow [NH13] to derive the generating function for the moments of the occupancy in a single fitness class. To this end, we define $\mathbb{P}\big[n,t|X,t_0\big]$ as the probability distribution of having occupancy n at time t after the fitness class has been seeded with a single individual at fitness X above the mean at time t_0. Thus, the distribution $\mathbb{P}\big[n,t|X,t_0\big]$ encodes all information on possible occupancies of the fitness class after this seeding event. Standard application of branching process methods leads to the dynamics

$$\begin{aligned}
\mathbb{P}\big[n,t|X+v\epsilon,t_0-\epsilon\big] \quad = \quad & \big(1-\epsilon(2+X-\mu)\big)\mathbb{P}\big[n,t|X,t_0\big] \\
& +\epsilon(1+X)\sum_{n'}\mathbb{P}\big[n',t|X,t_0\big]\mathbb{P}\big[n-n',t|X,t_0\big] \\
& +\epsilon\delta_{n,0} \\
& +\epsilon\mu\int ds\ m(s)\mathbb{P}\big[n,t|X+s,t_0\big] \ .
\end{aligned} \tag{3.37}$$

In order to account for the difference in mean fitness at t_0 and t, we assumed that its difference is given by $\sim v\big(t-t_0\big)$ and $\sim v\epsilon$ for tiny time-steps. The *direct* time evolution from initial time $t_0-\epsilon$ to final time t is identical to the time evolution from t_0 to t, but taking the events between $t_0-\epsilon$ and t_0 into account. The time difference ϵ is chosen small enough, such that (at maximum) one event occurs during ϵ. Then the time evolution from time t_0 sets in, encoded in $\mathbb{P}\big[n,t|x,t_0\big]$. The additional events occur at probabilities proportional to ϵ multiplied with the respective rates for birth, death and mutation (second to fourth lines). The first line in the derivation above corresponds to no additional event happening during ϵ.

The moment generating function is defined as

$$\begin{aligned} \psi_z\big(X,T=t-t_0\big) &= \sum_n e^{-zn}\mathbb{P}\big[n,t|X,t_0\big] \\ &= \overline{\exp\big(-zn\big)}\ , \end{aligned} \tag{3.38}$$

with the overbar as average over all possible configurations, as we have done already in the rest of this thesis. We arrive at its dynamics by inserting the definition in (3.37),

$$\begin{aligned} \partial_T\psi_z(X,T)-v\partial_X\psi_z(X,T) &= -\big(2+X\big)\psi_z(X,T) \\ &\quad +\big(1+X\big)\psi_z^2(X,T) \\ &\quad +1 \\ &\quad +\mu\int ds\ m(s)\psi_z(X+s,T)\ . \end{aligned} \tag{3.39}$$

As the name “moment generating function” suggests, we can derive all moments by

$$\partial_z^k\psi(X,T)\big|_{z=0}=(-1)^k\overline{n^k}\ . \tag{3.40}$$

More interestingly for us, we can take the limit

$$\lim_{z\to\infty}\psi_z(X,T)=\mathbb{P}\big[0,T|X,0\big]\ , \tag{3.41}$$

and arrive at the probability of having no individual left after time T. Hence, $z\to\infty$ yields exactly the extinction probability of the branching process. The complement to the extinction probability is the survival probability, $1-\mathbb{P}\big[0,T|X,0\big]$. To generalize, we define $\Phi_z(X,T)=1-\psi_z(X,T)$, and continue with the dynamics of $\Phi_z(X,T)$ after inserting the definition into (3.42):

$$\begin{aligned} \partial_T\Phi_z(X,T)-v\partial_X\Phi_z(X,T) &= X\ \Phi_z(X,T)-\big(1+X\big)\Phi_z^2(X,T) \\ &\quad +\mu\int ds\ m(s)\big(\Phi_z(X+s,T)-\Phi_z(X,T)\big)\ . \end{aligned} \tag{3.42}$$

As we have done before, we will approximate the coefficient of the quadratic term, $1+X\approx 1$. In the range of fitness we are interested in, the selection coefficients are tiny compared to one, $X\ll 1$. However, this term is ultimately responsible that $\Phi_z(X,T)$

does not grow without bound for huge X. Including $\big(1+X\big)$ leads to $\lim\limits_{X,z\to\infty}\Phi_z(X,T)=1$. Note that we had $u_\star(X)\approx X/2$ previously, which diverges for $X\to\infty$.

Taking the limit $z\to\infty$ in (3.42), we have an equation similar to the dynamics of $\mathbf{u}_\star$, *cf.* (3.18). The interpretation of $\lim\limits_{z\to\infty}\Phi_z(X,t)$ is the survival probability of the fitness class. In the long run, offspring from a fitness class can only survive, if their lineage fixates. Hence, we can assign the notion of a *fixation probability* to this limit. Noting that (3.42) itself is independent of z, and defining $u_T(X)=\lim\limits_{z\to\infty}\Phi_z(X,T)$, we have

$$\big(\partial_T - v\partial_X\big)\mathbf{u}_T = \big(\mathcal{L}^\dagger - \mathbf{u}_T\big)\mathbf{u}_T \ , \tag{3.43}$$

where we used the adjungated Liouvillean $\mathcal{L}^\dagger$ to denote the deterministic part of the selection and the mutation dynamics. Therefore, the dynamics of the fixation probability follows a time-evolution that is similar to the dynamics of our constraint $\mathbf{u}_\star$, *cf.* (3.18). Note, however, that the coefficient of the quadratic term is 1 instead of the factor 2 we have in tuned models.

In the following, we present a approximative solution to (3.42). It is in general worse than the approximation (3.24) above, as we only consider the linear regime (III), but now a *time-dependence* is included. For large fitness X, $\Phi_z(X,T)$ is also large, the main balance in (3.42) is between the time derivative and the two terms in the first line on the right hand side. Mutations and the term including the adaptation speed can be neglected. We arrive at an equation that is similar to logistic growth. Its solution is given by

$$\Phi_z(X,T) = \frac{\exp\big(XT\big)X}{\exp\big(XT\big) - \exp\big(Xf_z(X)\big)} \ , \tag{3.44}$$

where $f_z(X)$ is a (time-)integration constant, that has to be determined by boundary conditions. From the branching process above, we postulate to start with one individual, hence $\Phi_z(X,0)=1-e^{-z}$. The investigation of the time-evolution from such constant (in fitness) initial conditions has been a major aim in [Fis13]. Using our initial conditions in (3.44) yields

$$\Phi_z(X,T) = \frac{\exp\big(XT\big)X}{\exp\big(XT\big) - 1 + X\big(1-\exp(-z)\big)^{-1}} \ . \tag{3.45}$$

As a first check, we can evaluate (3.45) to obtain the fixation probability,

$$\lim_{z\to\infty}\Phi_z(X,T) = \frac{\exp\big(XT\big)X}{\exp\big(XT\big)-1+X} \xrightarrow{T\to\infty} X \ . \tag{3.46}$$

This is, reassuringly, the expected linear regime of the fixation probability $\mathbf{u}_\star$ (although in the formalism of tuned models we had an additional factor of 1/2).

Generating function in field theoretic approach

So far, we only treated the moments of occupancies in a single fitness class. A generalization to the population density $\mathbf{c}_t$ has been made in [Fis13, GD14]. The generating function resembles partition functionals in statistical mechanics. We have

$$Z_t\left[\mathbf{u}_t, \mathbf{c}_t\right] = \overline{\exp\Big(-\big\langle \mathbf{u}_t \big| \mathbf{c}_t \big\rangle\Big)} \; . \tag{3.47}$$

Here, the function $\mathbf{u}_t$ is at first only an auxiliary function that plays a similar role as z in the derivation above. From this generating function we obtain all moments of the population density with a functional derivative with respect to this auxiliary function:

$$\left.\frac{\delta Z_t\left[\mathbf{u}_t, \mathbf{c}_t\right]}{\delta \mathbf{u}_t}\right|_{\mathbf{u}_t=\mathbf{0}} = -\overline{\mathbf{c}_t} \; , \tag{3.48}$$

and, more general

$$\left.\frac{\delta^k Z_t\left[\mathbf{u}_t, \mathbf{c}_t\right]}{\delta \mathbf{u}_t \cdots \delta \mathbf{u}_t}\right|_{\mathbf{u}_t=\mathbf{0}} = (-1)^k \; \overline{\mathbf{C}_t^{(k)}} \; . \tag{3.49}$$

Details on such a formalism can be found in any advanced textbook on statistical mechanics from a field theoretic perspective, see *e.g.* [BDFN92] for a well written treatise.

We also can derive the time evolution of this generating functional, by using the dynamics of BRWs in (3.1) and the time evolution $\mathbf{u}_{t+\epsilon} \approx \mathbf{u}_t + \epsilon\partial_t\mathbf{u}_t$, neglecting orders $\mathcal{O}\big(\epsilon^{3/2}\big)$:

$$\begin{aligned} Z_{t+\epsilon} &= \overline{\exp\big(-\left\langle \mathbf{u}_{t+\epsilon} \middle| \mathbf{c}_{t+\epsilon} \right\rangle\big)} \qquad (3.50)\\ &\approx \overline{Z_t\left(1 + \sqrt{\epsilon}\big\langle \mathbf{u}_t \big| \sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t \big\rangle + \epsilon\left(\big\langle \partial_t\mathbf{u}_t \big| \mathbf{c}_t \big\rangle + \big\langle \mathbf{u}_t \big| \mathcal{L}\mathbf{c}_t \big\rangle - \frac{1}{2}\big\langle \mathbf{u}_t \big| \sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t \big\rangle^2\right)\right)} \; . \end{aligned}$$

The stochastic term $\sim \sqrt{\epsilon}\boldsymbol{\eta}_t$ is a non-anticipating function [Gar09] with respect to the density. Hence, in the average denoted by the overbar, it will cancel out. In turn, we arrive at the following expression for the time difference

$$\begin{aligned} Z_{t+\epsilon} - Z_t &= \overline{\epsilon Z_t\big\langle \mathbf{u}_t \big| \big(-\partial_t + \mathcal{L} - \mathbf{u}_t\big)\mathbf{c}_t \big\rangle + \mathcal{O}\big(\epsilon^{3/2}\big)} \\ &= \epsilon\big\langle \mathbf{u}_t \big| \big(-\partial_t + \mathcal{L} - \mathbf{u}_t\big)\overline{\mathbf{c}_t \big\rangle \; Z_t} + \overline{\mathcal{O}\big(\epsilon^{3/2}\big)} \qquad (3.51a)\\ &= \epsilon\big\langle \big(\partial_t + \mathcal{L}^\dagger - \mathbf{u}_t\big)\mathbf{u}_t \big| \overline{\mathbf{c}_t \big\rangle \; Z_t} + \overline{\mathcal{O}\big(\epsilon^{3/2}\big)} \; . \qquad (3.51b) \end{aligned}$$

The solutions are stationary, iff we have $\partial_t Z_t = 0$. In the average implicit in our partition functional, the forward dynamics of population density couples intricately to the partition functional itself, (3.51a). However, treating the backward dynamics allows to solve deterministic equations. The condition for stationarity is also fulfilled if this backward dynamics, (3.51b), is zero. The analysis done in [Fis13] is (partially) based on this insight. The dynamics to investigate is the the (already known) backward flow of the fixation probability,

$$-\,\partial_t\mathbf{u}_t = \big(\mathcal{L}^\dagger - \mathbf{u}_t\big)\mathbf{u}_t \; . \tag{3.52}$$

Difference in fixation probability to tuned models

Here we want to give an explanation of the factor 2 that arose in the context of tuned models: in the calculations in this section we used the dynamics of BRWs, or more specific, of *unconstrained* BRWs. The time evolution of such systems also includes the two cases we ruled out before: extinction of the population and infinite growth. On the other hand, the derivation for tuned models used *constrained* BRWs. There, the dynamics never reached one of those states, as the population density is rescaled each (infinitesimal) time step. Thus *counter-events* are triggered upon each birth, death and mutation event: the population always has to comply with $\langle \mathbf{u}_t | \mathbf{c}_t \rangle = 1$. These counter-events increase the stochasticity in the model, but as the constraint only acts globally on the whole population, this additional stochasticity is smeared out. Each fitness class receives only minute adjustment in the (fractional) number of individuals. Nevertheless, the large amount of these tiny correction adds up to twice the variance in reproduction.

This argument can be seen much easier in the constrained model with $\mathbf{u}_t \equiv 1/N$, the fixed population size model, which we used in simulations presented in chapter 2. For the sake of the argument, we restrict ourselves to only integer occupancies of fitness classes. Mutations do not trigger any counter-events, as they do not change the number of individuals[11]. Each birth event has to be followed by the death of another individual. There are N possible choices (as we have the uniform weighting from the constant $\mathbf{u}_t$), so each individual has probability $1/N$ to be killed due to the birth of another one. On average, there will be $\mathcal{O}(N)$ such counter-events per generation: the (original) death rate is 1, but with counter-events also occurring at rate $\mathcal{O}(N)\frac{1}{N} \approx 1$, we arrive at an effective death rate of ≈ 2. The same argument also holds for an increased effective birth rate. Upon each death event, another individual has to be born to comply with the constraint. Hence, the effective birth rate is $\approx \left(2 + x - \overline{x_t}\right)$.

Due to this increased effective birth and death rates, we would have a noise term (genetic drift) of $\sim \sqrt{4\mathbf{c}_t}\boldsymbol{\eta}_t$ in (3.1) or the prefactor $\left(\frac{N}{N} + \frac{N}{N} + 2 + x - \overline{x_{t_0}}\right)$ in the first line and $\left(\frac{N}{N} + 1 + x - \overline{x_{t_0}}\right)$ in the second line of the derivation for the generating function (3.37) above. Both would lead to the factor two in the final equation of the fixation probability with such adjusted coefficients. Therefore, we attribute the 2 in tuned models, (3.18), to this increased stochasticity.

The factor 2 is also reminiscent of the scaling of the time with $2/N$ in the Moran-model compared to the Wright-Fisher-model in classical population genetics: on average, there are N events before one individuals is chosen again to reproduce. The factor 2 in the Moran-model might similarly arise from the observation that an event never occurs alone: at first an individual is chosen to reproduce, then another individual has to give way for the newborn and die.

---

[11] Note that this assumption is violated in tuned models: different fitness classes are weighted with $u_\star(x)$ in the constraint, hence even a mutation does change $\langle \mathbf{u}_t | \mathbf{c}_t \rangle = 1$ and counter-events are triggered.

3.3.3. Simulation of fixation and extinction events

In order to validate our interpretation of $\mathbf{u}_\star$ as fixation probability for individuals with fitness X, we also conduct stochastic simulations with labelled sub-populations. The simulation code is similar to the code explained in chapter 2. However, the constraint is changed from the uniform weighting in a fixed population size model, *i.e.* $u_t(x) = 1/N$, to our tuned model $\mathbf{u}_\star$ in a comoving frame[12]. After generating different starting conditions (different stochastic snapshots of population densities) we reset time, $t = 0$, and label a sub-population $\mathbf{w}_0$ in the nose of the wave, such that

$$\big\langle \mathbf{u}_\star | \mathbf{w}_0 \big\rangle = 1/2 \; , \tag{3.53}$$

limited to the most fit individuals in the population $\mathbf{c}_0$. The fitness class with the lowest fitness in the sub-population usually consists also of unlabeled individuals, see Figure 3.2. From this configuration we run the stochastic time evolution of the population, until the sub-population either reaches fixation or goes extinct (in the extinction case, the *unlabeled* sub-population fixates). In accordance with the interpretation of $\mathbf{u}_\star$ as fixation probability, we expect half of the simulations to end in fixation of the sub-population, and half of the simulations in extinction. For definiteness, we abort simulations when one of the thresholds $\big\langle \mathbf{u}_\star | \mathbf{w}_t \big\rangle = 1 - 10^{-4}$ or $\big\langle \mathbf{u}_\star | \mathbf{w}_t \big\rangle = 10^{-4}$ is exceeded and count the simulation run as fixation event or extinction event, respectively.

After having amassed such simulation evidence, we can use Bayesian inference to check if our assumption of the interpretation of $\mathbf{u}_\star$ as fixation probability is consistent [OF94]. The distribution of the fixation probability F of the sub-population, given the (simulation-) data is computed as (using Bayes' theorem)

$$\mathbb{P}\big[\, F \,|\, \mathrm{data} \,\big] = \mathbb{P}\big[\, \mathrm{data} \,|\, F \,\big]\, \mathbb{P}\big[\, F \,\big] \; . \tag{3.54}$$

Here, the likelihood $\mathbb{P}\big[\, \mathrm{data} \,|\, F \,\big]$ of observing either an extinction or fixation event is a simple binomial distribution: when having N trials with n fixation events, the likelihood is $\mathbb{P}\big[\, \mathrm{data} = (N,n) | \, F \,\big] \sim F^n (1-F)^{N-n}$. Furthermore, we assume a flat prior $\mathbb{P}\big[F\big]$ for the fixation probability F, ignoring any knowledge about its value at the beginning. Such a flat prior can also be cast as a Beta distribution (incidentally a conjugate prior [OF94]), which in its general form is given by

$$\mathrm{Beta}\big(Y; \alpha, \beta\big) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} Y^{\alpha-1} (1-Y)^{\beta-1} \; . \tag{3.55}$$

Using the two hyperparameters $\alpha = 1$ and $\beta = 1$ we arrive at the uniform (flat) distribution on $\big[0;1\big]$. Thus, the posterior distribution for the fixation probability F of the sub-population can also be written as Beta-distribution:

$$\mathbb{P}\big[\, F \,|\, \mathrm{data} = (N,n) \,\big] \sim F^{n+\alpha-1} (1-F)^{N-n+\beta-1} \; , \quad \alpha = \beta = 1 \; . \tag{3.56}$$

From this posterior (3.56) we can evaluate the 95% confidence interval, and check if our assumption $F = 1/2$ is within its range. Increasing the value of $\alpha = \beta > 1$ would

---

[12] The simulation code is explained in the next chapter, in particular section 4.1.

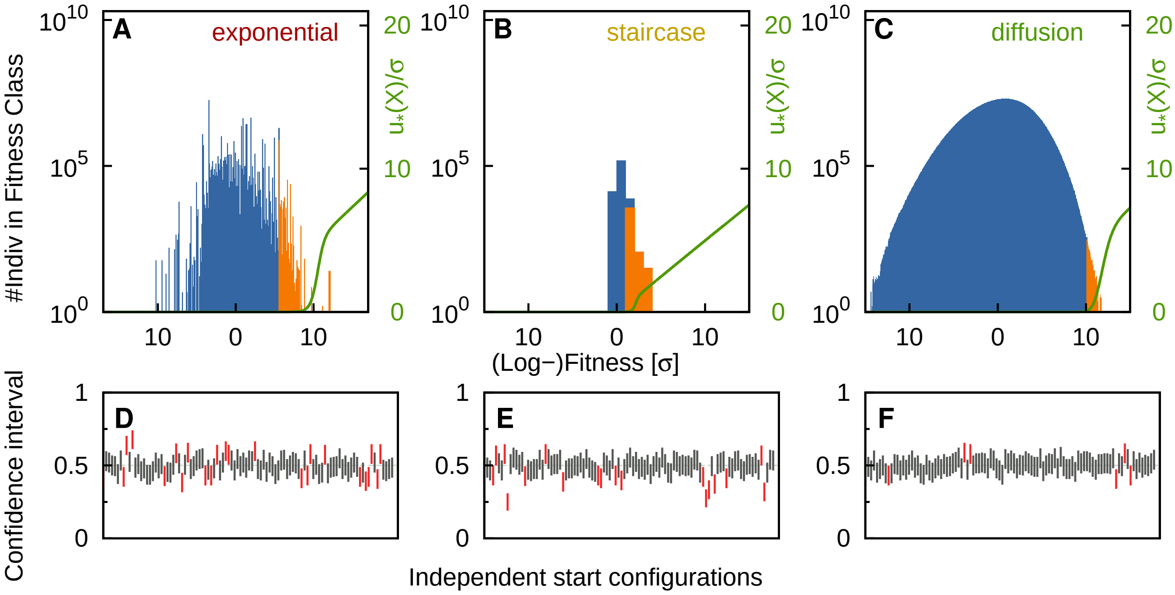

Figure 3.2.: **Measurements of fixation and extinction in stochastic simulations.** (A) - (C) We label a sub-population (orange) such that its overall expected fixation probability is 0.5. The green line is the (local) fixation probability $\mathbf{u}_\star$ (in non-logarithmic scale). (D) - (E) We count the fixation and extinction events from 100 independent starting conditions. The posterior distribution of the fixation probability is Beta-distributed (3.56), the two parameters are related to counts of fixation and extinction events (see main text for details). Here, we show the 95% interval of this distribution. Dark lines depict confidence intervals which include the predicted value, red ones do not include the predicted value, given that our predicted fixation probability is 0.5.

increase the certainness of our initial hypothesis, $F = 1/2$, that we put into the model, narrowing the distribution (3.56).

Figure 3.2 depicts the 95% confidence interval for 100 simulation runs started with different initial conditions $\mathbf{c}_0$, each iterated until 200 events occurred (either fixation or extinction). Most of the intervals contain the value we would have expected. Deviations are attributed to the discreteness in our simulations: we set the value of $\mathbf{u}_\star$ in a fitness class x as $\hat{u}_\star(x) = \frac{1}{\sigma}\int_{x-\sigma/2}^{x+\sigma/2} dy\; u_\star(y)$ (with fitness classes of size σ), which most likely will introduce lattice errors. Moreover, these deviations are most pronounced in the exponential mutation kernel that exhibits the most stochastic dynamics.

3.4. Interpretation of tuned models

The last sections were devoted to establish the interpretation of $\mathbf{u}_\star$ as fixation probability. With this knowledge we can review the methodology of using tuned models in the context of constrained BRWs. Specifically, we want to compare the two different choices for the constraint: first, we used a fixed population size in the simulations of chapter 2. Then, we used $\mathbf{u}_\star$, the tuned constraint introduced in the current chapter.

In the fixed population size model, the adaptation speed v is fluctuating. The analytic predictions cited in chapter 2, in particular (2.9), (2.10a) and (2.11), are valid for its mean value:

$$\overline{v} = \overline{v}(N, \mu, \sigma) \tag{3.57}$$

The fixation probability in these fixed-N models is complicated to calculate. It depends intricately on the current occupancies of different fitness classes, $\mathbf{c}_t$[13]. Other quantities with this constraint are also hard to obtain: in the *fully stochastic* modeling, the moment hierarchy does not close, as can be seen by inserting $\mathbf{u}_t = 1/N$ in the dynamics of the mean population density $\overline{\mathbf{c}_t}$, (3.17). However, the fixed-N model is conceptually simple, is in accordance with many previous theoretical investigations, and is easy to implement in simulations (as in chapter 2).

In tuned models, *the fixation probability is kept constant*, but the population size is allowed to fluctuate. This is the defining feature of tuned models. In particular, the solution to the fixation probability is computed in a *comoving frame* with speed v. The population size is adjusted such that the fixation probability constraint is always fulfilled, $\big\langle \mathbf{u}_\star | \mathbf{c}_t \big\rangle = 1$. Therefore, this model can be seen as an ensemble, where the adaptation speed is constant: at least fixation probabilities move constantly to higher fitness and drag the population density along, ultimately with average speed v:

$$\overline{N} = \overline{N}(v, \mu, \sigma) \tag{3.58}$$

This choice of constraint has a decisive advantage: the moment hierarchy closes, *cf.* (3.19), and the model is amenable to analytic treatment. Even if equations cannot be solved in closed form, numerical methods can be employed. We present the numerical code in appendix B. Therefore, we do not have to resort to averages over extensively long time series in simulations to obtain values for observables in the fully stochastic model.

The adjustment in population size depends crucially on variations and fluctuations in the nose of the wave, where $\mathbf{u}_\star$ grows large (see Figure 3.1). Without this particular constraint, the effects of such fluctuations would only be observable when those individuals present at the stochastic edge of the current population survived drift and grow to the largest clone in the population at time τ_{sw} later [Fis13]. Hence, our tuned models speed up the feedback mechanism of fluctuations on the population size.

The dichotomy of these two setups, (3.57) and (3.58), can be seen in the context of different ensembles in statistical mechanics. Both constraints allow different states (or trajectories, for that matter) in the population density $\mathbf{c}_t$. The global observables, population size N and adaptation speed v, switch roles in the two different constraints. One is a parameter of the model, while the other is a fluctuating quantity, that we want to obtain and investigate. Hence, choosing either $\mathbf{u}_t = N^{-1}$ or $\mathbf{u}_\star$ is like choosing different ensembles in statistical mechanics.

---

[13] Note here, that the fixation probabilities in the last section 3.3.2 were for *unconstrained* BRWs, while the fixed population size constraint represents of course a *constrained* BRW.

The next section will treat the analytical (and numerical) solutions of stationary quantities in the tuned model. Only in the next chapter 5, we compare time traces of simulations to analyze fluctuation behavior.

3.5. The distribution of fixating clones

The next step in our analysis is to evaluate different distributions that help to shed light on the dynamics of adapting asexual populations. The distribution of fixating clones is one of those important measures. It is the distribution of clones, of which one individual will become the (single) ancestor to the future population[14]. In the asexual populations we consider here, this will ultimately happen at some time. A similar measure has also been used for instance in [HN08] to investigate the advance of expanding populations. Later in this section, we also inspect other crucial distributions, like the distribution of the size of fixating mutations in the adaptation process. However, here we will only treat distributions generated by the averaged population density $\overline{\mathbf{c}_\star}$. In chapter 5 we will investigate the time dependent behavior of these measures.

3.5.1. Stationary population density

So far, we focused our analysis on the fixation probability $\mathbf{u}_\star$, but ignored the dynamics of actual population density $\overline{\mathbf{c}_t}$. Briefly, we established the existence of a stationary solution $\overline{\mathbf{c}_\star}$ in the same comoving frame we used for the fixation probability (see beginning of section 3.3). Recall, that in tuned models the stationary solution $\overline{\mathbf{c}_\star}$ is governed by the linear differential equation

$$-v\partial_X \overline{\mathbf{c}_\star} = \big(X - 2\mathbf{u}_\star\big)\overline{\mathbf{c}_\star} + \mathcal{M}\overline{\mathbf{c}_\star}\ . \tag{3.59}$$

In this particular arrangement of terms, we can assign the interpretation of "effective selection" to the first expression in brackets on the right hand side. Our constraint, arising from the finite size of the population, reduces the original selection term. We can use the solutions obtained in section 3.2.1 for the shape of $\mathbf{u}_\star$ to see that a observable effect sets in roughly before X_c: for fitness $X < X_c$ the value of $u_\star(X)$ is small (if not tiny) compared to X. Only at X_c and above the solution converges to the linear approximation, $u_\star(X) \sim X/2$, and the effective growth due to "selection" is almost zero. In Figure 3.3 (B), we show the shape of $\big(X - 2\mathbf{u}_\star\big)$ for all three mutation schemes.

Without this reduced selection term, deterministic traveling wave models also exhibit a strange feature: (exponential) growth of even minute fraction of individuals with

---

[14] Note again that individuals in a clone are indistinguishable, their only property is their (identical) growth rate.

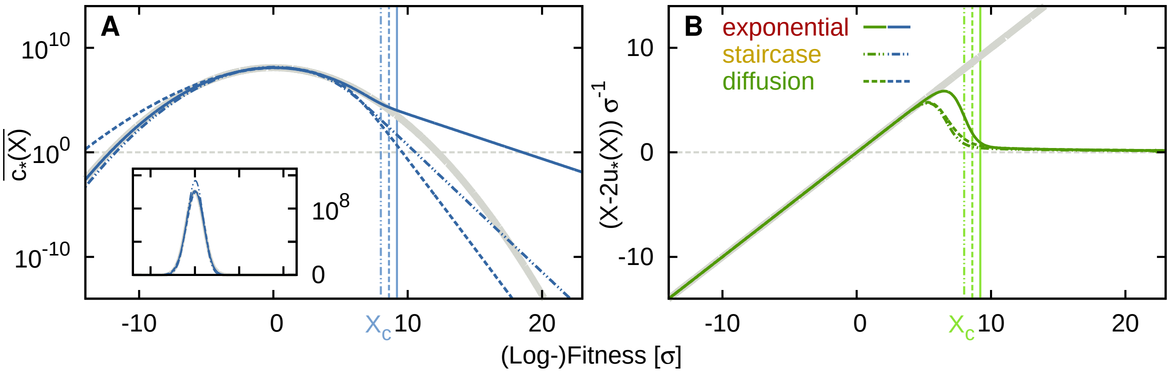

Figure 3.3.: **Stationary solution for the population density $\overline{\mathbf{c}_\star}$ and the shape of the function $\big(X - 2\mathbf{u}_\star\big)$.** All three mutation kernels are considered here: solid lines correspond to the exponential mutation scheme, dashed lines correspond to the diffusion mutation scheme, and dash-dotted lines correspond to the staircase mutation scheme. (A) Profiles shown exhibit the same adaptation speed, $v/\sigma^2 = 4$, and an identical mutational scale, $\sigma = 10^{-3}$. (In the diffusion mutation scheme we chose $vD^{-2/3} = 4$ and $D = 10^{-9}$.) Using the population size $\overline{N} \approx 6.30\ 10^5$ resulting from the diffusion model, we adjusted the mutation rates in the other two kernels such that they also agreed on this value. This lead to $\mu^{exp} \approx 7.97\ 10^{-4}$ and $\mu^{staircase} \approx 7.41\ 10^{-3}$. The gray line is a Gaussian density with the same integral (population size) and variance (adaptation speed) as the other profiles. Deviations from this generic solution can hardly be seen in non-logarithmic scale (see inset). However, on a logarithmic scale, we see that roughly at the stochastic edge X_c, the population density starts to decay exponentially. (B) The “effective” selection terms $\big(X - 2\mathbf{u}_\star\big)$ from the profiles in (A) are compared to the original selection, $\sim X$, shown as gray line. At roughly X_c the behavior changes qualitatively, as beyond this value the population does not grow deterministically anymore: the linear increase of fitness is offset by $\mathbf{u}_\star$. All solutions were obtained numerically with the code detailed in appendix B. Note that the code for the staircase mutation kernel is already quite unstable for such values.

extraordinarily high fitness leads to an acceleration of the population density moving to higher fitness, resulting finally in infinite adaptation speeds [DF07, Fis13]. This clearly nonsensical behavior is prohibited in stochastic models, but as we see here also in the mean population density of tuned models.

In addition, we present heuristics to solve (3.59), leading the the Gaussian we assumed for $\overline{\mathbf{c}_\star}$: we know that $X_c \sim \mathcal{O}\big(\sqrt{v}\big)$ from (3.25). Below that value, the term $\mathbf{u}_\star$ has only a small effect on the forward dynamics. With a small mutation rate, $\mu/\sigma \ll 1$, the major balance in (3.59) is between the derivation with respect to fitness (the comoving frame term) and the original selection term, $-v\partial_X \overline{\mathbf{c}_\star} \approx X\overline{\mathbf{c}_\star}$. This is solved by the proposed Gaussian profile, $\overline{\mathbf{c}_\star} \sim \exp\bigl(-X^2/(2v)\bigr)$. *A posteriori* we see that with this population variance the effect of the constraint only sets in when $\overline{\mathbf{c}_\star}$ has decayed sufficiently at X_c. Therefore, the Gaussian approximation appears valid.

We calculate our profiles with the code implementing the numerical methods detailed in appendix B.

3.5.2. Distribution of fixating clones

Clones, *i.e.* indistinguishable individuals comprising one fitness class, that will fixate in the population are influenced by two main factors: on the one hand, a high fitness of an individual increases its chances to fixate (as discussed at length in section 3.3). On the other hand, larger clones have higher fixation probability just by sheer numbers. Thus, the distribution of fixating clones arises from the trade-off between having enough individuals, but also being fit enough to outgrow all others. Thence, we define $\mathbf{g}_t \equiv g_t(x)$, the distribution of fixating clones as

$$\mathbf{g}_t = \mathbf{u}_\star \overline{\mathbf{c}_t} \ , \tag{3.60}$$

and correspondingly $\mathbf{g}_\star = \mathbf{u}_\star \overline{\mathbf{c}_\star}$ using the mean stationary population density. Thus, our constraint (3.14)

$$1 = \bigl\langle \mathbf{u}_\star | \mathbf{c}_t \bigr\rangle = \int dx \ u_\star(x) c_t(x) = \int dx \ g_t(x) \ , \tag{3.61}$$

is just the normalization condition on the distribution of fixating clones.

From the analysis done so far, we expect $\mathbf{g}_\star$ to be peaked around or on a slightly less fitness value than the cross-over X_c to the linear regime. There, the Gaussian population density and the fixation probability increasing like an inverse Gaussian tend to create a (broad) region, from where fixating clones will emerge. In Figures 5.5, 5.6 and 5.7 we show the fixation probability $\mathbf{u}_\star$, the stationary mean population density $\mathbf{c}_\star$ and the resulting distribution of fixating clones $\mathbf{g}_\star$ for all three mutation kernels with an appropriate range of parameters v, μ and σ (resp. D).

3.5.3. Distribution of fixating mutations

A successful individual from the distribution of fixating clones in the previous section usually features several new (beneficial) mutations in on its genome before fixation (when the population follows a dynamics in the clonal competition regime). In our probabilistic setting we cannot distinguish different mutations leading finally to the fitness of the respective clone. One may ask, however, what the effect of a *single* mutation on the fitness of an (arbitrary chosen) individual in the population is? It will turn out that this resembles closely another crucial question: which distribution of mutational effects do we actually observe in the population? This distribution is (most likely) not the

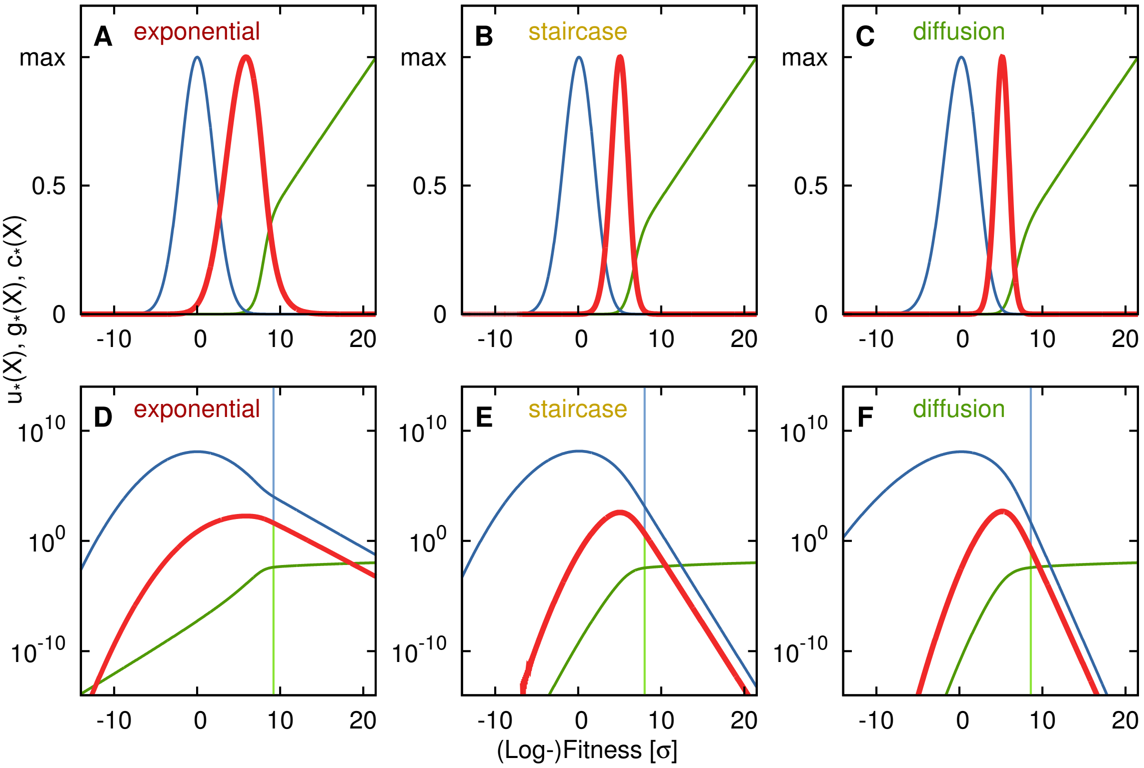

Figure 3.4.: **Distribution of fixating clones (red) for all three mutation kernels.** The blue lines depicts the mean stationary population density $\overline{\mathbf{c}_\star}$, while the fixation probability $\mathbf{u}_\star$ is shown in green. Vertical blue/green lines indicate X_c, the cross-over in the fixation probability to its linear solution, (3.25). Upper figures are scaled such that the maximum values of $\overline{\mathbf{c}_\star}$, $\mathbf{u}_\star$ and $\mathbf{g}_\star$ are identical in the chosen range. In all three mutation kernels, fixating clones arise most likely from a region just below X_c. (A) In the exponential mutation kernel fixating clones arise in a rather broad range of fitness below X_c. There, the inverse Gaussian solution to $\mathbf{u}_\star$, (3.24) and the approximate Gaussian population density $\overline{\mathbf{c}_\star}$ create a roughly constant region in $\mathbf{g}_\star$ (that is equivalent to regime (II) in section 3.2.1). (B) The staircase mutation kernel (C) The width of $\mathbf{g}_\star$ in the diffusion mutation kernel is essentially constant, $\sim D^{1/3}$, and independent of adaptation speed. The dynamical behavior of the the fixation process will be investigated in detail in chapter 5. Parameter values were $N = 6.3 \cdot 10^5$, $\sigma = 10^{-3}$, $\mu^{\text{exp}} = 4.2 \cdot 10^{-5}$, $\mu^{\text{staircase}} = 4$ and $D = 10^{-9}$.

raw DFE, but shaped by the actual dynamical process of adaptation. The analysis done in [GRB+12] is based on the idea of using this distribution of fixating mutations to determine a consistency relation for the adaptation speed. Briefly, we discuss their argument here.

First note, that the average probability of an mutation of size s to fixate is proportional to $\int dx\ \overline{c_t(x)}u_\star(x+s)$: after occurring at a fitness background x (weighted by the density $\overline{c_t(x)}$), its overall fitness is $x+s$. Then, the individual with its new mutation fixates proportional to $u_\star(x+s)$. A mutation of size s itself occurs proportional to $\mu m(s)$, with $\mathbf{m}$ the mutation kernel. Thus, the distribution of fixating mutations $\boldsymbol{\rho}_f$ is given by

$$\rho_f(s) \sim m(s)\int dx\ \overline{c_\star(x)}u_\star(x+s)\ . \tag{3.62}$$

This expression (3.62) can be used to define the adaptation speed: overall, the population advances towards higher fitness by fixating new mutations. Hence, the rate of fitness increase is the weighted average of fixating mutations [GRB$^+$12, Neh13]:

$$\begin{aligned} v &= \mu\int ds\ s\rho_f(s) \\ &\sim \mu\int ds\ sm(s)\int dx\ \overline{c_\star(x)}u_\star(x+s)\ . \end{aligned} \tag{3.63}$$

In [GRB$^+$12] this integral has been evaluated using a Gaussian population density $\overline{\mathbf{c}_\star}$ (with variance v) and approximations for the fixation probability $\mathbf{u}_\star$ (linear from X_c and above, inverse Gaussian below X_c). As a second relation for the two unknowns v and X_c they use an integrated version of (3.59). In turn, they obtain a numerical solutions for these values, which coincide well with the stochastic simulations results. From these equations they also derive the scaling already mentioned in section 2.2.2, in particular expression (2.10a), that captures the trend, but deviations are discernible for large population sizes.

3.6. Higher moments in tuned models

In section 3.2 we introduced tuned models, where a constraint $\mathbf{u}_\star$ was chosen such that the moment hierarchy closes at the first moment. Later, in section 3.3, we established the interpretation of $\mathbf{u}_\star$ as the profile of fixation probabilities in fitness space. Now we turn our attention to higher moments of the population density

$$\overline{\mathbf{C}_t^{(n)}} \equiv \overline{c_t(x_1)c_t(x_2)\cdots c_t(x_n)}\ . \tag{3.64}$$

Here, (3.64) defines a n-point correlation function with n variables x_1, x_2, …, x_n in fitness space, averaged over the noise. With knowledge about these correlation functions, we can analyze the behavior of fluctuations in adapting populations. Theoretically speaking, the whole spectrum of fluctuations can be estimated if all correlation functions are available [Kam92].

The structure of the derivation for tuned models allows to extend the method. Previously, we used the dynamics of the generator for *constrained* BRWs (3.15) to obtain two coupled equations for the fixation probability (3.18) and the population density (3.19). In order to solve for higher correlation functions, we choose a different constraint, denoted $\mathbf{u}_\star^{(n)}$, we can decouple the term including the $(n+1)^{\text{th}}$ moment $\overline{\mathbf{C}_t^{(n+1)}}$ from the dynamics of the n^{th} moment $\overline{\mathbf{C}_t^{(n)}}$. In turn we arrive at $n+1$ coupled equations: this set of equations for $\overline{\mathbf{c}_t} \equiv \overline{\mathbf{C}_t^{(1)}}, \overline{\mathbf{C}_t^{(2)}}, \ldots, \overline{\mathbf{C}_t^{(n)}}$ and $\mathbf{u}_\star^{(n)}$ is closed, however.

We start with the stochastic generator for the n-point correlation function $\mathbf{C}_{t+\epsilon}^{(n)}$ (and a general, not yet specified $\mathbf{u}_t$):

$$\begin{aligned}\mathbf{C}_{t+\epsilon}^{(n)} &\equiv \mathbf{c}_{t+\epsilon}\cdots\mathbf{c}_{t+\epsilon} \\ &= \frac{\mathbf{c}_t+\epsilon\boldsymbol{\mathcal{L}}\mathbf{c}_t+\sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t}{\left\langle \mathbf{u}_t+\epsilon\partial_t\mathbf{u}_t \middle| \mathbf{c}_t+\epsilon\boldsymbol{\mathcal{L}}\mathbf{c}_t+\sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\right\rangle}\cdots\frac{\mathbf{c}_t+\epsilon\boldsymbol{\mathcal{L}}\mathbf{c}_t+\sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t}{\left\langle \mathbf{u}_t+\epsilon\partial_t\mathbf{u}_t \middle| \mathbf{c}_t+\epsilon\boldsymbol{\mathcal{L}}\mathbf{c}_t+\sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t\right\rangle}\,.\end{aligned} \tag{3.65}$$

Similar to before, we can expand (3.65) to orders $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\sqrt{\epsilon})$. The readability is enhanced by using the (non-linear) operator $\mathcal{D}\mathbf{c}_t \equiv \sqrt{2c_t(x)}\eta_t(x)$ for genetic drift, as already introduced in the section 3.3.1:

$$\begin{aligned}\mathbf{C}_{t+\epsilon}^{(n)}-\mathbf{C}_t^{(n)} &= \epsilon\Big[\sum_{i=1}^{n}\mathcal{L}_i\mathbf{C}_t^{(n)}-n\sum_{i=1}^{n}\mathcal{D}_i\mathbf{C}_t^{(n)}\left\langle\mathbf{u}_t|\mathcal{D}\mathbf{c}_t\right\rangle+\sum_{i\neq j}\mathcal{D}_i\mathcal{D}_j\mathbf{C}_t^{(n)}\Big] \\ &\quad+\epsilon\Big[n\mathbf{C}_t^{(n)}\left\langle\partial_t\mathbf{u}_t|\mathbf{c}_t\right\rangle-n\mathbf{C}_t^{(n)}\left\langle\mathbf{u}_t|\mathcal{L}\mathbf{c}_t\right\rangle+\frac{n(n+1)}{2}\mathbf{C}_t^{(n)}\left\langle\mathbf{u}_t|\mathcal{D}\mathbf{c}_t\right\rangle^2\Big] \\ &\quad+\sqrt{\epsilon}\Big[\sum_{i=1}^{n}\Big(\mathcal{D}_i\mathbf{C}_t^{(n)}-\mathbf{C}_t^{(n)}\left\langle\mathbf{u}_t|\mathcal{D}\mathbf{c}_t\right\rangle\Big)\Big]+\mathcal{O}\big(\epsilon^{3/2}\big)\end{aligned} \tag{3.66}$$

Note that each term contains the same variables $x_1, \ldots, x_n$: in terms with additional $\mathbf{c}_t$, these population densities always occur in integrals $\left\langle\mathbf{u}_t|.\right\rangle$. Operators with indices act on the i^{th} factor in the product $\mathbf{C}_t^{(n)} \equiv c_t(x_1)c_t(x_2)\cdots c_t(x_i)\cdots c_t(x_n)$. As before, we average over the noise in the model, upon which the stochastic terms of order $\mathcal{O}(\sqrt{\epsilon})$ vanish. In the limit $\epsilon \to 0$ we arrive at

$$\begin{aligned}\partial_t\overline{\mathbf{C}_t^{(n)}} &= \sum_i\Big(\mathcal{L}-2n\mathbf{u}_t^{(n)}\Big)_i\overline{\mathbf{C}_t^{(n)}}+2\sum_{i\neq j}\delta(x_i-x_j)\overline{\mathbf{C}_t^{(n-1)}}_{\setminus x_j} \\ &\quad-n\left\langle\big(\partial_t+\mathcal{L}^\dagger-(n+1)\mathbf{u}_t\big)\mathbf{u}_t\middle|\overline{\mathbf{C}_t^{(n+1)}}\right\rangle\end{aligned} \tag{3.67}$$

When comparing this dynamics of $\overline{\mathbf{C}_t^{(n)}}$, (3.67), to the original dynamics of the mean population density $\overline{\mathbf{c}_t}$, (3.17), we see that the first term is related to the “standard” dynamics in each of the fitness dimensions. The second sum appears first for $n = 2$. It originates from the combined influence of drift at positions in different dimensions. In particular, we used here that the drift term calculates as

$$\begin{aligned}\overline{\mathcal{D}_i\mathcal{D}_j\mathbf{C}_t^{(n)}} &\equiv \overline{c_t(x_1)\cdots\sqrt{2c_t(x_i)}\eta_t(x_i)\ldots\sqrt{2c_t(x_j)}\eta_t(x_j)\ldots c_t(x_n)} \\ &= 2\delta(x_i-x_j)\overline{c_t(x_1)\ldots c_t(x_i)\ldots c_t(x_{j-1})c_t(x_{j+1})\ldots c_t(x_n)}\,,\end{aligned} \tag{3.68}$$

leading to the expression in (3.67). The coupling to the next order of correlation functions involves again the integral containing the dynamics of the constraint function $\mathbf{u}_t$. When inspecting (3.67), it becomes apparent that we can propose a specific constraint $\mathbf{u}_t$ rendering the last term in (3.67) zero. For the moment closure at the n^{th} level, the tuned constraint $\mathbf{u}_\star^{(n)}$ is governed by the (backward-in-time) dynamics:

$$-\partial_t \mathbf{u}_t^{\star(n)} = \left(\mathcal{L}^\dagger - (n+1)\mathbf{u}_t^{\star(n)}\right)\mathbf{u}_t^{\star(n)} . \tag{3.69}$$

Inserting this solution into the dynamics of the n-point correlation function $\overline{\mathbf{C}_t^{(n)}}$, we obtain a solvable system of equations (with all other correlation functions with $m < n$ given below).

Using the marginalization[15] $\langle \mathbf{u}_t^{\star(n)} | \mathbf{C}_t^{(n)} \rangle = \mathbf{C}_t^{(n-1)}$ and inserting the tuned constraint given by (3.69) into (3.67) we can state a single integro-differential equation for $\overline{\mathbf{C}_t^{(n)}}$,

$$\partial_t \overline{\mathbf{C}_t^{(n)}} = \sum_i \left(\mathcal{L} - 2n\,\mathbf{u}_\star^{(n)}\right)_i \overline{\mathbf{C}_t^{(n)}} + 2\sum_{i\neq j} \delta(x_i - x_j) \left\langle \mathbf{u}_\star^{(n)} \middle| \overline{\mathbf{C}_t^{(n)}} \right\rangle . \tag{3.70}$$

All other moments $1 \leq m < n$ are governed by (3.67), but with all n replaced by m. In turn, the term involving the coupling to higher moments *does not vanish* upon inserting the solution $\mathbf{u}_\star^{(n)}$. However, we can use the dynamics of our tuned constraint (3.69) to arrive at a stationary instead of a time-dependent term:

$$\left\langle \left(\partial_t + \mathcal{L}^\dagger - (m+1)\mathbf{u}_\star^{(n)}\right)\mathbf{u}_\star^{(n)} \middle| \overline{\mathbf{C}_t^{(m+1)}} \right\rangle = (n-m) \left\langle \mathbf{u}_\star^{(n)}\mathbf{u}_\star^{(n)} \middle| \overline{\mathbf{C}_t^{(m+1)}} \right\rangle .$$

Therefore, the dynamics of all other moments, $m < n$, is given by

$$\begin{aligned} \partial_t \overline{\mathbf{C}_t^{(m)}} &= \sum_i \left(\mathcal{L} - 2m\mathbf{u}_\star^{(n)}\right)_i \overline{\mathbf{C}_t^{(m)}} + 2\sum_{i\neq j} \delta(x_i - x_j) \overline{\mathbf{C}_t^{(m-1)}}_{\backslash x_i} \\ &\quad - m(n-m) \left\langle \mathbf{u}_\star^{(n)}\mathbf{u}_\star^{(n)} \middle| \overline{\mathbf{C}_t^{(m+1)}} \right\rangle . \end{aligned} \tag{3.71}$$

The equations (3.69), (3.70) and all $n-1$ equations (3.71) are a set of $n+1$ equations to solve for all correlation functions up to n variables. Usually, one does not solve (3.71) for lower correlation function, but use the marginalization mentioned above,

$$\overline{\mathbf{C}_t^{(m-1)}} = \left\langle \mathbf{u}_\star^{(n)} \middle| \overline{\mathbf{C}_t^{(m)}} \right\rangle , \tag{3.72}$$

and integrate out all variables one after the other until one arrives again at the mean population density $\overline{\mathbf{c}_t}$ ($m = 1$). The dynamics (3.71) of the m-point correlation function (with $m < n$) is only stated for completeness. However, one still ends up with $n+1$ equations for the correlation functions and the constraint: the $n-1$ intricate time evolutions in (3.71) are just replaced by (3.72).

---

[15] which holds with and without averaging over the noise,

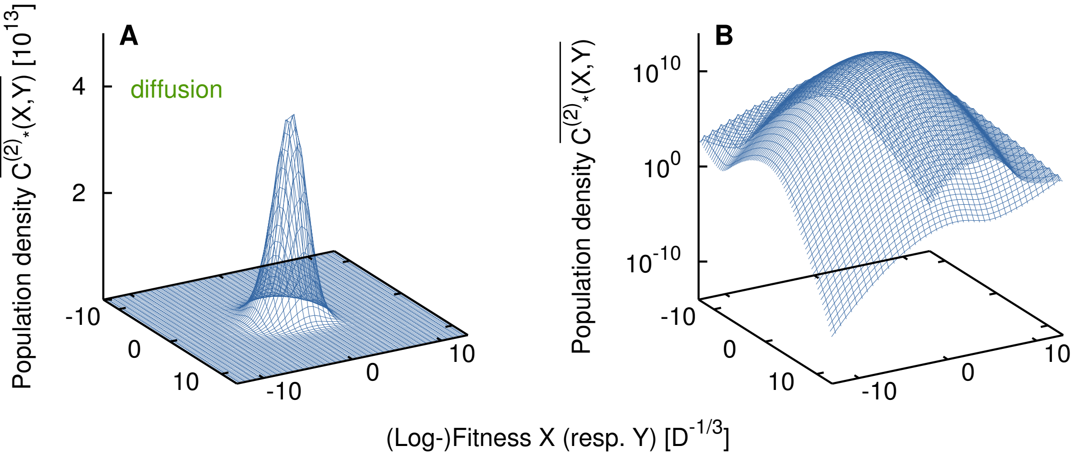

Figure 3.5.: **Stationary solution $\overline{C_\star^{(2)}(X,Y)}$ in the diffusion mutation scheme.** The solution is obtained by solving (3.74) numerically with parameters $vD^{-2/3} = 2.5$ and $D = 10^{-6}$.

The non-linear term in (3.69), $\sim \mathbf{u}_\star^{(n)}\mathbf{u}_\star^{(n)}$, has a coefficient $n+1$ instead of the coefficient 2 in the original equation when closing the moment hierarchy at the first moment, *cf.* (3.18). By the simple rescaling $\mathbf{u}_\star^{(n)} = \frac{n+1}{2}\mathbf{u}_\star$ (with $\mathbf{u}_\star = \mathbf{u}_\star^{(1)}$), both solutions can be transformed into each other. More general, we have the transformation

$$\big(n+1\big)\mathbf{u}_\star^{(n)} = \big(k+1\big)\mathbf{u}_\star^{(k)}\ , \quad n \geq 1,\ k \geq 1\ . \tag{3.73}$$

Hence, the different regimes in the dynamics of the stochastic population, explained at length in section 3.3, are still the same: if an individual has a extraordinary high fitness (beyond X_c) it just has to overcome genetic drift. Just below this cross-over, clonal competition still dominates the behavior of clones, while for negative fitnesses it is exponentially unlikely for an individual to fixate. However, when closing the moment hierarchy at higher moments, the number of ”counter-events“ to keep up with the *constrained* population is increased: each dimension in the n-point correlation function has to be considered.

The factor α in the first order Kramers-Moyal expansion, (3.23) and (3.24), should become clear as a generalization for the closure at higher moments. For $X > X_c$, the solution converges to $u_\star^{(n)}(X) \approx X/(n+1)$.

Obtaining a general analytical solution to (3.70) is a cumbersome, if not impossible, task. Already the stationary 2-point correlation function $\overline{C_\star^{(2)}(X,Y)}$ is governed by the (rather complicated) partial integro-differential equation

$$\begin{aligned}
0 \;=\;& v\partial_X \overline{C_\star^{(2)}(X,Y)} + v\partial_Y \overline{C_\star^{(2)}(X,Y)} \\
&+X\,\overline{C_\star^{(2)}(X,Y)} + Y\,\overline{C_\star^{(2)}(X,Y)} \\
&+\mu \int ds\; m(s)\big(\overline{C_\star^{(2)}(X-s,Y)} - \overline{C_\star^{(2)}(X,Y)}\big) \\
&+\mu \int ds\; m(s)\big(\overline{C_\star^{(2)}(X,Y-s)} - \overline{C_\star^{(2)}(X,Y)}\big) \\
&-4u_\star^{(2)}(X)\,\overline{C_\star^{(2)}(X,Y-s)} - 4u_\star^{(2)}(Y)\,\overline{C_\star^{(2)}(X,Y-s)} \\
&+2\delta(X-Y)\int dZ\; u_\star^{(2)}(Z)\overline{C_\star^{(2)}(X,Z)}\;. \qquad (3.74)
\end{aligned}$$

Often, one has to resort to numerical methods, for example those presented in appendix B.3. In Figure 3.5 we show an example of this 2-point correlation function with the diffusion mutation kernel computed numerically with our code.

Similar to tuned models with a moment closure already at the mean ($n = 1$), where we had $\mathbf{u}_\star \equiv \mathbf{u}_\star^{(1)}$ as fixation probability, we can interpret $\mathbf{u}_\star^{(n)}$ as fixation probability when closing the moment hierarchy at the n^{th} level. In appendix C.2 we present the derivation for this claim. It is based on similar considerations as those in section 3.3.1.

4. Simulations of tuned models

At this point, it is not yet clear that tuned models from chapter 3 exhibit a similar dynamical behavior than the simulations with fixed population size in chapter 2. So far, we know that tuned models produce *exactly* solvable equations for the mean population density $\overline{\mathbf{c}_t}$. In section 3.6 the method was extended to obtain all n-point-correlation functions $\overline{\mathbf{C}_t^{(n)}}$ of the population density. This extension allows to gain access to the exact fluctuation spectrum of traveling wave models. Analyzing the fluctuation behavior in simulations will be the main topic of this and the next chapter.

Tuned models replace the well-known constraint of a fixed population size with a special weighting of different fitness classes. In particular, the population size is adjusted such that the fixation probability is kept constant in a comoving frame with speed v. This forces the (mean) adaptation speed to coincide with v. However, the population size is not constant anymore in tuned models. In section 3.4 we argued that this difference, $\left(N, \overline{v}\right)$ vs. $\left(\overline{N}, v\right)$, can be interpreted in the light of different ensembles, like in statistical mechanics. It is plausible that tuned models are similar to the original fixed population size model, but the aim of this chapter is to quantify differences of these two. We focus on the fluctuating population size in this chapter. The next chapter is devoted to the fluctuations in the adaptation speed. These fluctuations in adaptation speed have a much more interesting interpretation for real populations, that will be explained in chapter 5.

First, we introduce a slightly modified stochastic simulation here, that exhibits the tuned constraint $\left\langle \mathbf{u}_\star \middle| \mathbf{c}_t \right\rangle = 1$. We will review all results presented first in chapter 2 with these stochastic tuned simulations, but now augmented with detailed numerical solutions for averaged simulation observables.

4.1. Stochastic simulation of tuned models

The simulation code is similar to the algorithm in section 2.1. Revisit Figure 2.1 for reference of the three sub-steps (A), (B) and (C) to advance the population a step ϵ in time. At first, deterministic changes in the occupancies $\mathbf{n}$ are applied due to selection and mutation[1], as in (2.2). Then, genetic drift is accounted for, *cf.* (2.3). Both sub-steps are similar in simulations of tuned models:

---

[1] Recall that in simulations we use occupancies $\mathbf{n} \equiv n_i$ of discrete fitness classes: each fitness class has a finite width $dx > 0$. The continuous description with the population density $\mathbf{c}_t$ in the last chapter is valid for infinitesimal dx. The ultimate reason for this difference is that the white noise is correlated over distances dx here, and not the Dirac-delta as in the continuous description.

$$\Delta^{(A)} n_i/\epsilon = X_i n_i + \mu\Big(\sum_j m_j n_{i-j} - n_i\Big) \ , \tag{4.1}$$

$$\Delta^{(B)} n_i/\sqrt{\epsilon} = \sqrt{2}\big(\text{Poisson}(n_i) - n_i\big) \ . \tag{4.2}$$

The last sub-step in the simulation algorithm is different, however: we rescale the occupancies by

$$\Delta^{(C)} n_i = \left(\left(\sum_j u_j n_j\right)^{-1} - 1\right) n_i \ , \tag{4.3}$$

instead of the uniform weighting done in (2.4).

In order to obtain the values for u_j in (4.3), we use a numerical solution to the fixation probability $\mathbf{u}_\star$. Recall from (3.21), that in a comoving frame the governing equation is

$$v\partial_X \mathbf{u}_\star = X\mathbf{u}_\star + \mathcal{M}^\dagger \mathbf{u}_\star - 2\mathbf{u}_\star^2 \ . \tag{4.4}$$

The numerical code to solve (4.4) is based on a multi-dimensional Newton-Raphson iteration scheme and detailed in appendix B.1. Usually we choose a finer lattice spacing dx in the numerics than in stochastic simulations[2]. Parameters of the numerical solution are adaptation speed v (fixed to the speed of the comoving frame), and the mutation parameters μ and σ (respectively D in the diffusion mutation kernel).

In the stochastic simulation, we average over lattice points in the numerical solution $\mathbf{u}_\star$ to obtain the coefficients u_j in (4.3). We also account for the comoving frame: in each step (of duration ϵ) the profile $\mathbf{u}_\star$ is shifted by $v\epsilon$ to higher fitness. Hence, over time the weighting u_j associated with a fixed fitness class j declines, as the fixation probability $\mathbf{u}_\star$ is a monotonically increasing function. Fluctuations in the occupancy of this fitness class j are affecting the overall population size less and less. Small fluctuations in the nose of the wave, however, can lead to considerable changes in the population size, as these fitness classes are weighted with a much larger coefficient in (4.3).

As our scale, we set $\sigma = 10^{-3}$ for all simulations with the exponential and staircase mutation model. For the diffusion kernel the diffusion constant D is varied, but in a range to comply with the assumption $D \sim \mathcal{O}(\mu\sigma^2)$, such that "reasonable“ mutation rates μ result.

4.2. Fluctuations in population size

4.2.1. Trajectories of fitness increase and population size

At first we examine trajectories of observables in the stochastic simulation of tuned models. Figure 4.1 depicts the fitness increase over time. For each of the different

---

[2] In most cases that are presented in this thesis we use $dx/\sigma = 1/20$ in numerics, while we have $dx^{\text{staircase}}/\sigma = 1$ and $dx^{\text{exp}}/\sigma = 1/10$ in simulations.

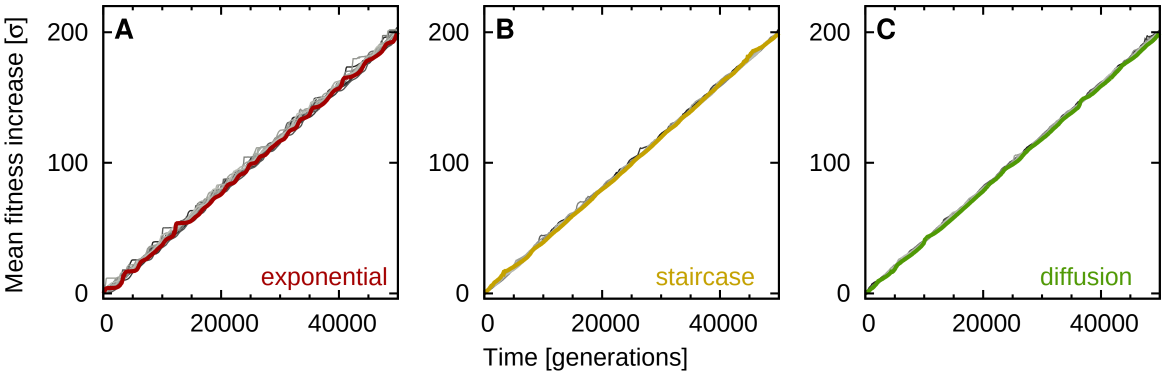

Figure 4.1.: **Trajectories of mean fitness increase in tuned models.** The forcing to a fixed adaptation speed v is clearly visible. Small undulations exists only when an extraordinary fit sub-populations emerges, increasing the mean fitness fast while growing to become the dominant sub-population. Then, fitness increase in the population grinds to a halt. The relatively "large", very fit sub-population poses a major weight in the constraint $\sum_j u_j n_j = 1$. Compared to the fitness increase in the fixed-N simulations, deviations from the *linear* fitness increase are small. Simulation parameters were $v = 4 \cdot 10^{-6}$, $\mu^{\text{exp}} = 4.1 \cdot 10^{-5}$, $\mu^{\text{staircase}} = 6.1 \cdot 10^{-4}$, $\sigma = 10^{-3}$, $D = 10^{-9}$. Different quantities from these 60 simulation runs (20 for each mutation kernel) are also shown in Figures 4.2 and 5.2 with identical colors.

mutation kernel we present 20 example trajectories. These trajectories are analogous to the fitness increase in fixed-N simulations in Figure 2.3. Parameters in the tuned simulations were adjusted, such that they are comparable to data points in the fixed population size simulations. Inspecting the data it becomes apparent that the population is forced to adapt with speed v, although small fluctuations are visible. The mean adaptation speed is v. Fluctuations that were present in adaptation speed (in fixed-N simulations) are now present in fluctuations of the population size (in tuned models). These trajectories of population size are shown in Figure 4.2.

Deviations from the linear increase in fitness occur, when a fitness class in the nose of the wave receives a large positive fluctuation in occupancy. Then, the non-uniform weighting rescales the overall population size down to small values. These drops in population size can be observed in Figure 4.2, where the time-dependent population size is shown for the same simulations as in Figure 4.1. Hence, fast increases in fitness trigger a considerable reduction in population size. We argue, however, that these (large) fluctuations in population size do not alter the the dynamics of traveling wave models too drastically. Here, one has to keep in mind that the population size enters observables, like the adaptation speed, only logarithmically. This fact has been discussed in section 2.2.2.

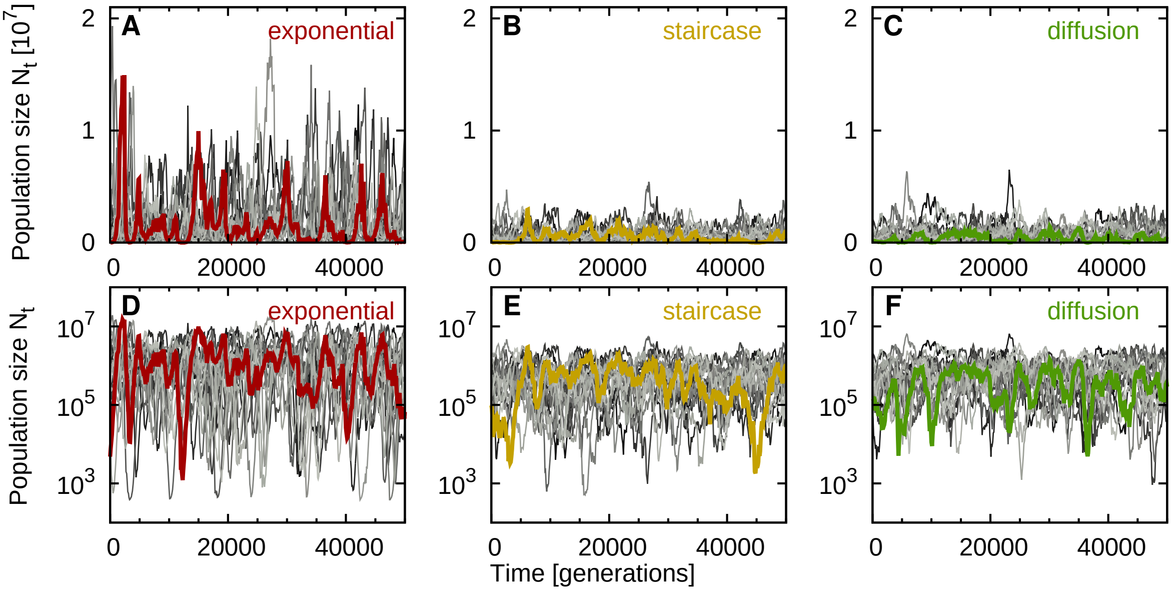

Figure 4.2.: **Trajectories of population size in tuned models.** In tuned models the adaptation speed is constant, but large fluctuations in population size occur. These excursions often span several orders of magnitude. Minima in population size correspond to a state with a single "large" sub-populations in the nose of the wave, which contribute a major weight in the constraint, (4.3). Fluctuations are most pronounced in the exponential mutation kernel, where mutations can increase fitness substantially. However, fluctuations are also present in the other two mutation kernels. Top and bottom rows show the trajectory in population size in normal and logarithmic scale, respectively. Simulation parameters were $v = 4 \cdot 10^{-6}$, $\mu^{\text{exp}} = 4.1 \cdot 10^{-5}$, $\mu^{\text{staircase}} = 6.1 \cdot 10^{-4}$, $\sigma = 10^{-3}$, $D = 10^{-9}$. Different quantities from these 60 simulation runs (20 for each mutation kernel) are also shown in Figures 4.1 and 5.2 with identical colors.

The tuned constraint just emphasizes and anticipates the expected outcome of such a fluctuation: individuals with this extraordinarily high fitness will most likely become ancestors to the future population. All other fitness classes are doomed to extinction and individuals in these fitness classes do not contribute to the future dynamics much. Therefore, even such large drops in the absolute number of individuals are only minor changes in the expected outcome of the dynamics.

However, we do not see a slower increase in mean fitness than forced by the comoving frame speed v. The undulation are only fast increases in fitness and a subsequent stasis, until the comoving frame catches up. This is not surprising in tuned models. When a very fit sub-population is absent, the constant decrease in the coefficients u_j scales the population size to large values. This also increases the occupancies in the very fit fitness

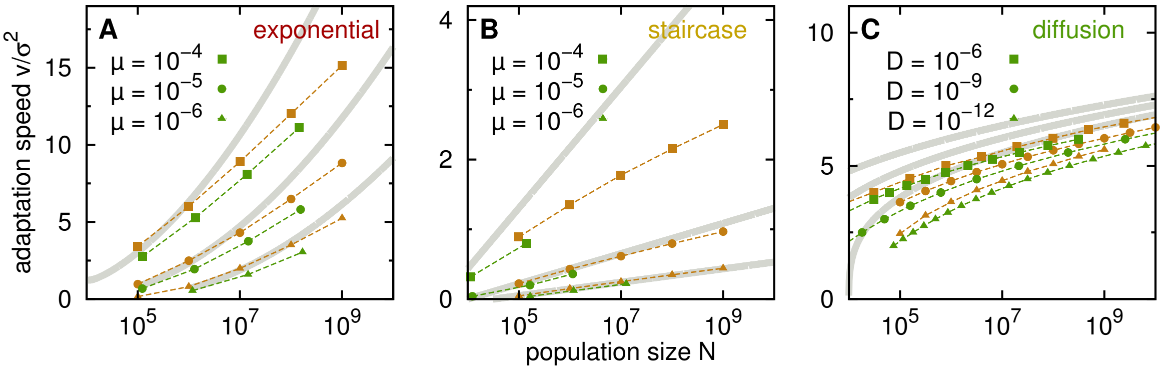

Figure 4.3.: **Mean population size dependence on adaptation speed in tuned models (green) compared to results obtained in fixed population size models (brown).** Results on the mean adaptation speed $\overline{v}$ in fixed-N simulations were shown in Figure 2.4. Gray lines are analytic predictions of section 2.2.2 (in particular (2.9), (2.10a) and (2.11)). Overall, values for adaptation speed in both, fixed-N and tuned, models have similar magnitude, but tuned models exhibit always a slower adaptation speed. This difference can be explained by the non-constant population size in tuned models (see main text for detailed arguments). The numerical code for the staircase mutation kernel is quite unstable, reliable results could only be obtained for small adaptation speeds.

classes. In turn, they reach a dynamical regime where the exponential growth due to selection is a stronger force than the random changes in occupancy due to genetic drift. A new clone emerges, that increases mean fitness substantially.

In Figure 4.3 we summarize the results on the population size and compare them to the results from chapter 2. The first observation is that results on both models are similar, but not identical. However, results are close in the N-v-plane. In tuned models, we measure a larger mean population size $\overline{N}$ than in the fixed-N simulations, when adaptation speed v and the mutation parameters are identical. In the next sections, we will quantify this observed difference.

4.2.2. Distribution of population sizes

In Figure 4.4 we show the distribution of population sizes recorded in tuned models for all three mutation kernels. For $N > \overline{N}$, a clear exponential decay visible. For large adaptation speeds v, the decay is dominated by only a single parameter, the mean population size $\overline{N}$:

$$\mathbb{P}\big[N|v \text{ large}\big] \sim \exp\big(-N/\overline{N}\big) \ . \tag{4.5}$$

For smaller speeds in the exponential and staircase mutation kernel, the decay parameter increases up to $\approx 4\overline{N}$, but the tail remains exponential. In the diffusion kernel, the exponential decay with the single parameter $\overline{N}$ in (4.5) is valid for all adaptation speeds, even small ones.

The tail for small population sizes, $N < \overline{N}$, is different in the three kernels. In the exponential mutation kernel, a clear power-law with exponents $-1 < \alpha < 0$ is indicated, where α depends on the model parameters v, μ and σ. For large mean population sizes $\overline{N}$, this algebraic decay is visible over 6 orders of magnitude (see Figure 4.4 D). Such small population sizes correspond to the rare emergence of extraordinarily fit clones in the nose of the wave that contribute a major weight in the constraint. For smaller $\overline{N}$, the power-law breaks down already at smaller relative population sizes, as seen in Figure 4.4. However, the minimum of recorded population sizes is $N \approx \mathcal{O}\big(10^2\big)$, a value that holds for all conducted simulations (that have identical simulation length). For a single choice of parameters, this lower limit can also be guessed from the trajectories of population sizes shown in Figure 4.2.

In the staircase and diffusion mutation kernel, the whole distribution of population sizes appears to be exponential. The exponential tail appears to be valid even for $N < \overline{N}$. The power-law seen in the exponential mutation kernel is not observed. The lower limit on observed population sizes is again $\mathcal{O}\big(10^2\big)$ in the staircase mutation kernel for all (checked) parameters. In the diffusion mutation kernel, the observed lower limit depends on the choice of adaptation speed.

Generating function predicts an exponential tail

Some insight into this exponential distribution can also be gained with analytic methods. In section 3.3.2, we introduced the generating function for the occupancy n of a single fitness class. Recall that we defined

$$\Phi_z(X,T) = 1 - \sum_n e^{-zn}\mathbb{P}\big[n,T|X,0\big] \; , \tag{4.6}$$

where $\mathbb{P}\big[n,T|X,0\big]$ is the probability of having n individuals at time T in the fitness class X ahead of the mean, after it was seeded with a single individual at time 0. Neglecting the mutation term that couples different fitness classes, we derived

$$\Phi_z(X,T) \approx \frac{\exp\big(XT\big)X}{\exp\big(XT\big) - 1 + X\big(1-\exp(-z)\big)^{-1}} \; . \tag{4.7}$$

This expression has been inverted in [DF07]. The general solution is a geometrical distribution in n, that can be approximated by an exponential for times $T \gg X^{-1}$:

$$\mathbb{P}\big[n,T|X,0\big] \sim \exp\big(-\beta_T\, n\big) \; . \tag{4.8}$$

Note that the coefficient β_T converges to infinity for $T \to \infty$, such that the probability for finite n decays to zero. This fact has been mentioned already several times in this

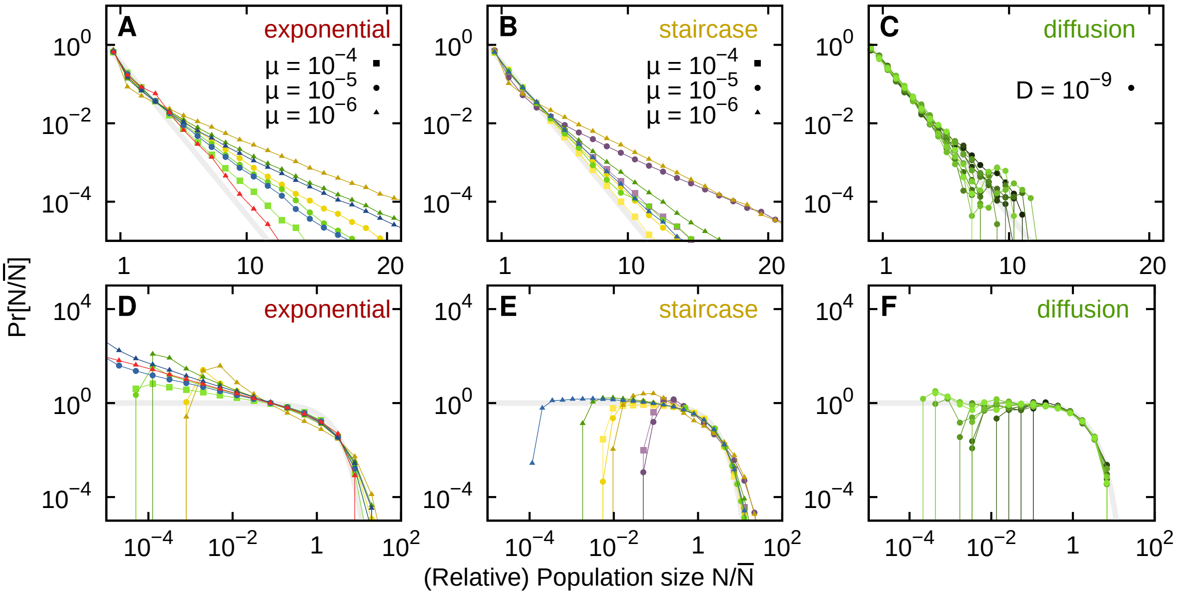

Figure 4.4.: **Population size distribution in tuned models with different mutation schemes.** Although different parameters lead to population sizes that are several orders of magnitude apart (see Figure 4.3), the shape of the distributions looks similar when the horizontal axis is scaled by the mean population size $\overline{N}$. An exponential scaling $\mathbb{P}\left[N/\overline{N}\right] = \exp\left(-N/\overline{N}\right)$ is shown as gray line. Colors in the exponential and staircase mutation kernel indicate the order of magnitude in the population size, the correct mean values are shown in Figure 4.3: red $\mathcal{O}\left(10^8\right)$, blue $\mathcal{O}\left(10^7\right)$, green $\mathcal{O}\left(10^6\right)$, yellow $\mathcal{O}\left(10^5\right)$, purple $\mathcal{O}\left(10^4\right)$. In the diffusion mutation kernel, the color gradient from light to dark green corresponds to different adaptation speeds: light is faster, dark is slower. (A) - (C) In the upper panels a semi-logarithmic plot of the population size distribution reveals an exponential tail. In the exponential and staircase mutation kernel the tail of the distribution of population sizes converges to the exponential distribution with the parameter $\overline{N}$ (gray line) for increasing population sizes. In the diffusion mutation kernel the exponential scaling is observed, irrespective of the value for adaptation speed. Different diffusion constants also lead to the exact same exponential tail (data not shown). (D) - (F) In a double-logarithmic plot, the left tail of the distribution for $N < \overline{N}$ is emphasized. Values smaller than the mean occur, when a sub-population ahead of the mean emerges, such that the tuned constraint decreases population size considerably: these sub-populations contribute a major weight in (4.3). Hence, the left tail is dominated by effects of our constraint.

thesis: the only two long-time behaviors are extinction or infinite growth in *unconstrained* BRWs. Obviously, in unconstrained BRWs the probability for any finite occupancy n converges to zero.

In [DF07] this analysis is extended. The authors couple the occupancy of such a exceptionally fit and stochastically growing sub-population to an second *exponentially growing* sub-population. From this second sub-population, mutations are generated in a Poisson process with rate μ, that enter the sub-population of interest, described by the generating function above. In turn, they arrive at an algebraic distribution of n for the highest fitness class that is occupied: $\mathbb{P}\big[n\big] \sim n^{\alpha}$, with exponents $-2 < \alpha < -1$.

Although the general solution (4.8) to a branching process suggests an exponential decay, it is not immediately clear how to interpret it in the light of our tuned models. The size distribution in the nose of the wave will influence the overall population size only later. This time delay is given by the time when occupancies in these fitness classes have grown large enough such that individuals within them contribute a major weight in the constraint. For fixed-N simulations this happens roughly at a time, when the wave moved half its width to higher fitness[Fis13]. However, in our tuned model the overall population size is very sensitive to changes in the occupancy in the nose of the wave. Most likely, values smaller than the mean $\overline{N}$ are crucially influenced by our constraint: small population sizes only occur when relatively large, very fit sub-populations emerge. In such situations the overall population size is given almost entirely by the occupancies of these fitness classes. Hence, we expect the left tail ($N < \overline{N}$) of the distribution $\mathbb{P}\big[N\big]$ to be heavily influenced by $\mathbf{u}_{\star}$. The algebraic tail in the exponential mutation kernel might be a consequence.

Another relevant result was obtained in [Fis13]. There, the traveling wave model was conditioned on *non-extinction*. Then, the generating function Ψ_z of the whole population size N is given by

$$\Psi_z = \sum_N e^{-zN}\mathbb{P}\big[N\big] \approx \frac{1}{1+z\overline{N}} \; . \tag{4.9}$$

This generating function yields an exponential distribution of populations sizes (see equation (114) in [Fis13]).

A more detailed investigation of this exponential distribution of population sizes in tuned models is still missing, and will be pursued in future work.

4.2.3. Adaptation speed dependence on population size

After having obtained the distribution of population sizes (at least in simulations), we can use this knowledge to give an explanation for the perceived difference in adaptation speed, when comparing fixed-N and tuned simulations (see In Figure 4.3). For identical adaptation speeds, the population size differs by a factor of roughly 2. However, in tuned models we computed the mean population size as *algebraic mean* over an extensive time series. Using this fact, we present a calculation explaining this difference in the following.

In section 2.2.2 we briefly demonstrated results that derived the adaptation speed with various analytical arguments. The overarching conclusion was that population size enters the expression for adaptation speed *logarithmically.* However, in many of the cited publications population size N was treated as a fixed parameter, ignoring possible fluctuations over time. Our tuned model is constructed that fluctuations in adaptation speed are suppressed, but large deviations from the mean population size occur. This "change of ensembles" allows to quantify the influence of fluctuations in population size. In the following, we will argue that the adaptation speed depends on $\overline{\log N}$ instead of $\log \overline{N}$. To this end, we use the exponential tail of the population size distribution. Assuming large adaptation speeds, we can use the distribution with the single parameter, $\mathbb{P}\left[N\right] \sim \exp\left(-N/\overline{N}\right)$. This yields

$$\begin{aligned}
\overline{\log N} &= \int_0^\infty dN \, \left(\log N\right)\mathbb{P}\left[N\right] \\
&\approx \int_0^\infty dN/\overline{N} \, \left(\log N\right)\exp\left(-N/\overline{N}\right) \\
&= \log \overline{N} + \int_0^1 dk \, \log\log\frac{1}{k} \\
&= \log \overline{N} - \gamma \, . \qquad (4.10)
\end{aligned}$$

The definite integral in the third line is evaluated as the Euler-Mascheroni constant $\gamma \approx 0.5772156649$. When using the exponential distribution in (4.10) we ignored any further structure of the distribution $\mathbb{P}\left[N\right]$ for small N. Writing the numerical constant inside the logarithm yields

$$\overline{\log N} \approx \log\left(\frac{\overline{N}}{1.78}\right) \, . \qquad (4.11)$$

Hence, commuting averaging and logarithm introduces an additional factor of $\mathcal{O}\left(2\right)$.

We also conducted simulations to confirm the observation that this factor of ≈ 2 arises from the dichotomy between $\overline{\log N}$ and $\log \overline{N}$. This is done by two different setups, with results shown in Figure 4.5. Both methods rely on the fact that we have recorded the trajectory of population sizes in tuned models. First, we can compute the logarithm of the single values on the trajectories, then take the average, *i.e.* we compute $\overline{\log N}$. When inspecting the resulting $\left(\overline{\log N}, v\right)$ in the N-v-plane, the corresponding data point is much closer to the fixed-N models than to the tuned models: compare the green pentagons in Figure 4.5 to the brown line corresponding to results in the fixed-N simulations. Secondly, we use the code of "fixed population size" simulations, but impose a fluctuating populations size. Here, the trajectories of population sizes N_t recorded in tuned models are used. In these simulations, adaptation speed is *not* fixed as in the simulations of tuned models. We only measure the effect of fluctuating population sizes on adaptation speed. The resulting data point, $\left(\overline{N_t}, \overline{v}\right)$ is shown as red triangles in Figure 4.5. Here, $\overline{v}$ resembles closely the value of v that has been used in tuned models to produce the solution for $\mathbf{u}_\star$ and in turn the time trace in stochastic simulations. As with tuned models, we use the algebraic mean $\overline{N}$ to specify population size.

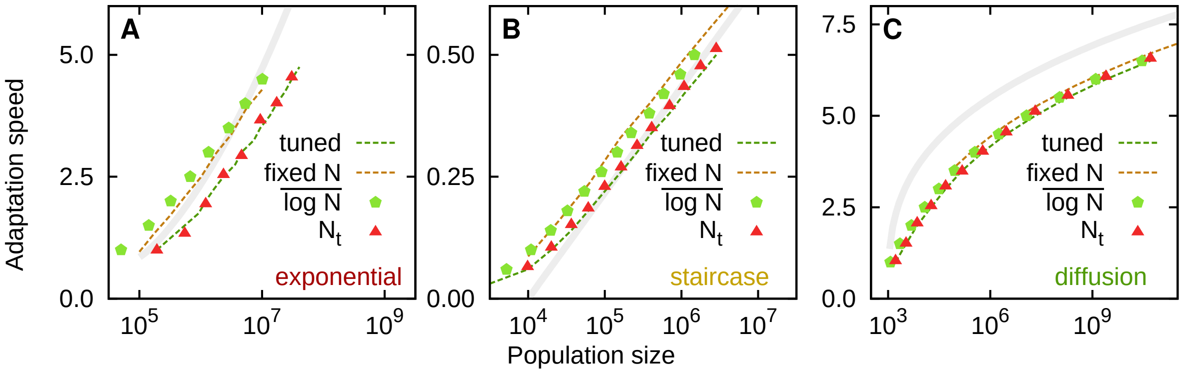

Figure 4.5.: **Mean adaptation speed depends on $\overline{\log N}$ instead of $\log \overline{N}$.** The brown dashed line depicts the mean adaptation speed $\overline{v}$ in fixed-N models. The green dashed line depicts the mean population size $\overline{N}$ in tuned models with a comoving frame of speed v. Consult also Figure 4.3 for a wider range of parameters. Here, we also draw $\overline{\log N}$ in the tuned model as green pentagons. For all three mutation kernels, the resulting coordinates in the $N - v$ plane are much closer to the results of the fixed population size model (brown line). Red triangles show simulations where the population size was adjusted to follow a trajectory N_t recorded in tuned models (see main text for simulation details). Even without imposing the particular constraint $\mathbf{u}_\star$ of tuned models, the mean adaptation speed $\overline{v}$ in this model with fluctuating population size is much closer to tuned models than to fixed population size models.

Therefore, we conjecture that the governing *dynamical quantity* in traveling wave models is the logarithm of the population size $\left(\log N\right)_t$, and not the population size N_t itself (or $\left(\log N_t\right)$ for that matter). This manifests itself in the fact that in simulations with fluctuating population sizes $\overline{\log N}$ instead of $\log \overline{N}$ reproduces a similar behavior to fixed-N simulations. When the population size is constant, the mathematical operations of averaging and taking the logarithm commute: both quantities are identical, $\overline{\log N} = \log \overline{N}$. Therefore, our investigation revealed a genuine feature of fluctuations in the fully stochastic traveling wave model.

This concludes our analysis of fluctuating population sizes. The next chapter deals with fluctuating adaptation speed.

5. Oscillations in traveling wave models of adaptation

5.1. Introduction

The last few chapters have been rather technical treatises on tuned models. There, we established traveling wave models, and in particular tuned models, as a minimal model of adaptation in highly diverse, microbial populations. These organisms have been the focus evolutionary experiments recently, as detailed in section 1.2. This chapter tries to deal with an biologically relevant problem: How does the diversity in such populations change over time?

The random emergences of novel, beneficial mutations in a real population do not cause fitness to increase completely linear. For a *single* beneficial mutation the dynamics for a sweep is well understood [Gil10]. If the novel mutation manages to overcome genetic drift, the fitness increase has a sigmoidal shape in time. This shape leads to a temporary increase of adaptation speed, but after the sweep adaptation speed decays again.

In the clonal competition scenario, mutations enter the population frequently, indicated by the parameter regime $N\mu \gg 1$. The dynamics of single sweep events is smeared out. Different clones in the population usually accumulate several beneficial mutations before one of them takes over the population. Such *cohorts of mutations* have been found by sequencing the DNA of adapting microbial populations at multiple times during an evolution experiment [LRH$^+$13]. Similar to single mutations, the sweep of a lineage carrying a cohort of mutations increases adaptation speed temporarily.

After a sweep of a clone with either a single mutation or a cohort of mutations, diversity in a population is low. If the experiment is run long enough, however, diversity will increase again: novel mutations enter the population frequently for $N\mu > 1$. Ultimately, the process repeats itself, leading to another sweep event. One of the aims in this chapter is to quantify these dynamics: What are the timescales involved in such fluctuations in diversity? We will try to answer this by using traveling wave models where the adaptation process entered its “steady state”.

Fluctuations in diversity can also be observed experimentally. However, the timescales of this process are usually a few hundred if not a few thousand generations. Experiments with such a long duration are scarce. The Lenksi populations are one of the few examples that span enough generations to contain multiple sweep events. In section 5.4 we will try to interpret the experimental observations with the results obtained in this chapter.

It should be noted here, that the contents of this chapter are not yet a complete investigation of the fluctuating behavior of the diversity in a population. We believe, however, it can act as starting point for future investigations.

5.2. Time-dependent dynamics of BRWs

A solution to the time-dependent equation of motion for the population density in general BRWs is not possible due to the non-closing moment hierarchy. In our tuned model the equation of motion for the mean is a “simple” PDE, (3.19). However, a analytic solution is not yet found. Even a general numerical solution is currently beyond our reach.

In order to investigate time-dependent behavior, we rely on time traces in stochastic simulations. Several methods to investigate possible regularities or irregularities in time-dependent behavior exists. For instance, computing the power spectrum from the trajectory of the observable is an often used method. However, we employ autocorrelation functions to analyze time-dependence. This allows us to clearly distinguish two crucial time scales in the adaptation process of populations in the clonal competition regime, as we will see below.

5.2.1. Autocorrelation functions

Consider a dynamical observable a_t. An autocorrelation function (ACF) A_T of a_t is defined as

$$\begin{aligned} A_T &= c\,\overline{\delta a_{t+T}\delta a_t} \\ &= c\,\overline{\left(a_{t+T}-\overline{a_t}\right)\left(a_t-\overline{a_t}\right)} \\ &= c\left(\overline{a_{t+T}a_t}-\overline{a_t}^2\right) . \end{aligned} \tag{5.1}$$

The overbar denotes averaging over the whole time series or the whole ensemble. The dependence on a single time point t drops out in the average.

The ACF is a measure of the correlation of two values in the time series, that are a time difference T apart. Usually, one chooses $c = \left(\mathrm{Var}\left[a_t\right]\right)^{-1}$ such that $A_0 = 1$. With this choice, the ACF is bounded by $-1 < A_T < 1$. Positive values of A_T correspond to time differences T, when a_{t+T} and a_t tend to lie on the same side of the mean value, i.e. δa_{t+T} and δa_t are both negative or both positive. At time differences T when A_T is negative, the values of a_{t+T} and a_t tend to be on opposite sides of the mean. Averaged over the whole ensemble (or time series), characteristic time scales of the process can be found in the shape of A_T.

The simplest examples are stationary Markov processes a_t, which do not have a memory of their past. In these cases A_T decays exponentially, $\sim \exp\left(-T/\tau\right)$. Fluctuations to either side of the mean decorrelate on the timescale τ. A decorrelation time scale will be present for all stochastic processes, although this explicit exponential decay does not necessarily occur. An arbitrary evolving process could exhibit more complex behavior, ranging from long power-law tails to oscillations in one or multiple frequencies.

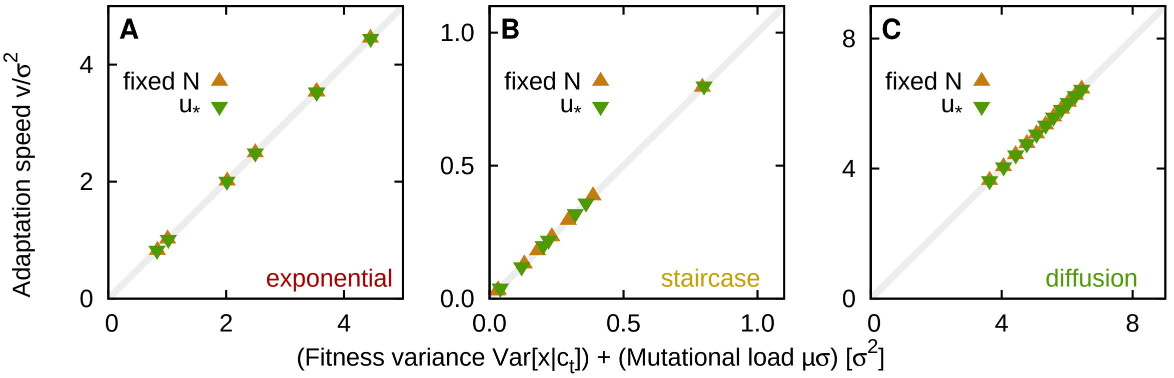

Figure 5.1.: **"Fisher's fundamental theorem" measured in simulations.** The "theorem" identifies adaptation speed v with the sum of the variance in fitness of the population and an additional mutational bias $\mu\sigma$, *cf.* (5.2) and (3.3). Both sides agree to a very good extent for all three mutation kernels. Moreover, we could not discern a difference of simulations with our tuned constraint and simulations with the fixed-N constraint. The relation between fitness variance and adaptation speed holds with both constraints. The mutational parameters for the exponential and staircase mutation kernel are $\sigma = 10^{-3}$ and $\mu = 10^{-4}, 10^{-5}, 10^{-6}$. In the diffusion mutation kernel the mutational bias is zero and $D = 10^{-9}$.

5.2.2. "Fisher's fundamental theorem" revisited

"Fisher's fundamental theorem" has been described in section 3.1. It relates adaptation speed v with the fitness variance of the population density. In particular, it can be stated as

$$v \approx \mathrm{Var}\big[x|\mathbf{c}_t\big] + \mu\sigma \;. \tag{5.2}$$

The last term is the mutational load $\mu\sigma$, that is assumed to be deterministic in our model.

In the following we compute the ACF of fitness variance to evaluate the time-dependence of the diversity in the modeled population. We defined the ACF with deviations from the mean value, see (5.1). Both, adaptation speed and fitness variance, will have identical deviations from their respective mean values: they only differ by a deterministic additive constant, as seen in "Fisher's fundamental theorem" above, (5.2). Thus, the ACF of adaptation speed and fitness variance are identical. We show the mean values of both sides of relation (5.2) in Figure 5.1. They agree to a very good extent.

The ACF of fitness variance can also be used in tuned models, even though adaptation speed is kept fixed. Computing the ACF of adaptation speed would lead to nonsensical results in tuned models. There, fluctuations in adaptation speed are translated into fluctuations in population size. The original derivation of (5.2) in section 3.1 included this fact: “Fisher’s fundamental theorem” has been obtained by multiplying the equation of motion for the population density with x/N_t and a subsequent integration over fitness. The ACF of *fitness variance*, however, does not suffer the problem of the choice of the constraint we are employing to keep the population size finite.

Another problem with the ACF of *adaptation speed* arises in the computation of the instantaneous adaptation speed. First, the mean fitness of the population density has to be calculated at consecutive time points. With the difference in mean fitness $\Delta\overline{x_t}$, the adaptation speed is obtained by $v = \Delta\overline{x_t}/\Delta t$. This discreteness in time Δt increases the noise in the computation of the ACF.

Therefore, we will use the ACF of fitness variance to determine timescales from simulation data. Correlations in the adaptation speed exhibit the same timescales by construction.

5.2.3. Timescales in clonal competition scenarios

Before presenting results on ACF of fitness variance, we present trajectories of the fitness variance in stochastic simulations in Figure 5.2. Data from the same simulations was shown in previous chapters, in particular in Figures 2.3, 4.1 and 4.2.

Positive fluctuations to large fitness variance values are similar with both constraints (fixed-N and tuned). Thus, broad, extended states of the population density are comparable in both models. The distribution of all recorded fitness variances values shows an exponential tail for values larger than the mean (data not shown). However, negative fluctuations to small values in fitness variance are different. Here, tuned models permit very small values in the trajectory. These states have been described already in the last chapter: a small variance of the population density corresponds to the emergence of extraordinarily fit fitness classes. If those grow to an appreciable number of individuals, the fitness class contributes a major weight in the constraint. The whole population consists almost entirely of individuals in this particular fitness class. In the distribution of fitness variances this behavior induces a power-law tail including very small values, as already seen in the distribution of population sizes. In contrast, with the fixed-N constraint, the distribution falls off quickly for values less than the mean.

Figure 5.3 shows the ACF of fitness variance in simulations. Independent of the choice of the mutation kernel, oscillations are clearly discernible. Oscillations appear for parameters in the clonal competition regime, $N\mu \gg 1$. They vanish, when parameters are reduced to values $N\mu < 1$. The latter, small parameter values correspond to the periodic selection regime.

When investigating the shape of the ACF, we can distinguish two different timescales: the first is the decay time (or decorrelation time) τ_{dec}. The stochastic process underlying the traveling wave loses its memory on its past shape due to fluctuations within a time of order $\mathcal{O}(\tau_{dec})$. In clonal competition scenarios, a second timescale τ_{osc} emerges. It describes the average timescale, on which the population density is “breathing”: starting from a maximum, the fitness variance passes its minimum and reaches its maximum again during a time τ_{osc}. In order to extract relevant time scales from the autocorrelation

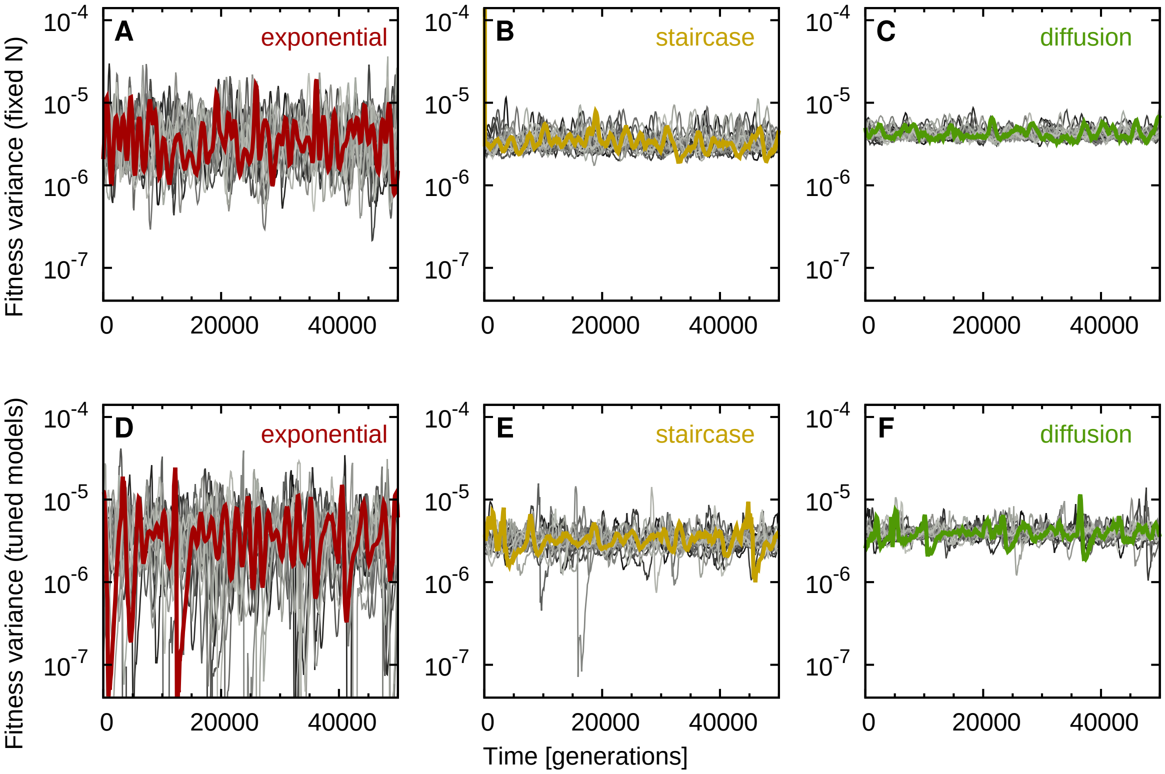

Figure 5.2.: **Trajectories of fitness variance for all three mutation kernels.** Here, the dynamics of simulations with fixed-N constraint (upper panels) is compared to the dynamics of simulations with the tuned constraint (lower panels). Positive fluctuations to larger values in fitness variance appear similar with both constraints. However, with the tuned constraint negative fluctuations are much more pronounced. These occur when the population of individuals concentrated in only a single, extraordinarily fit fitness class, that contributes a major fraction in the constraint, $\sum_j u_i n_j = 1$. See main text for more details. The amplitudes of fluctuations in different mutation kernels reflect the expectations: the large jumps that are possible in the exponential mutation kernel lead to large extensions or contractions of the population density. The staircase mutation kernel exhibits intermediate amplitudes. The diffusion kernel exhibits the smallest amplitudes. This behavior can also be discerned from Figures 2.3 and 4.1, where the increase in fitness is shown for the very same simulations: the fitness variance is related to the first derivative of these trajectories, *i.e.* the adaptation speed, via “Fisher’s fundamental theorem” (*cf.* section 5.2.2).

function, we use the fit

$$A_T^{(FV)} = c\,\exp\big(-T/\tau_{dec}\big)\,\cos\big(2\pi(T+\phi)/\tau_{osc}\big)\ . \tag{5.3}$$

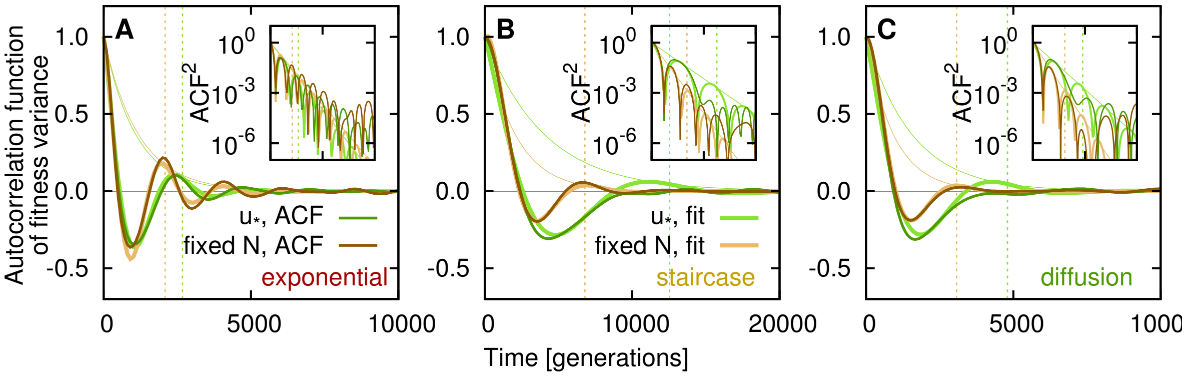

Figure 5.3.: **Autocorrelation function of fitness variance for all three mutation kernels.** Parameters were chosen such that the oscillations are pronounced. In the fixed-N simulations this resulted in $N^{\text{staircase}} = 10^{6.5}$, $N^{\text{exp}} = 10^8$, $\mu^{\text{staircase}} = \mu^{\text{exp}} = 10^{-6}$, $\sigma = 10^{-3}$, $N^{\text{diffusion}} = 10^{10}$, $D = 10^{-9}$. In the simulations with a tuned constraint we used the adaptation speed v obtained by averaging over time series in the fixed-N simulations. The autocorrelation function (ACF) of fitness variance is computed from stochastic simulation data. Bold lines in lighter colors indicate the matching numerical fit given by (5.3). Vertical lines indicate the oscillation timescale τ_{osc}. Insets show the square of the ACF on a logarithmic scale.

The normalization constant is $c = \left(\cos\left(2\pi\phi/\tau_{osc}\right)\right)^{-1}$ to ensure $A_0^{(FV)} = 1$. The offset ϕ is used to capture the first minimum in the ACF better. Subsequently, this results in a better agreement of the oscillation timescale τ_{osc}. The fit with only a *single* parameter for each timescales, τ_{dec} and τ_{osc}, shows a good agreement to simulation data, *cf.* Figure 5.3.

We tested if the hypothesis of using only a single oscillation time scale can hold. To this end, we implemented a similar fit, but with 11 oscillation timescales instead of a single one. This is implemented by the target function

$$A_T^{(FV)} = \exp\left(-T/\tau_{dec}\right)\left(\sum_{i=0}^{10} \alpha_i\, c_i \cos\left(\frac{2\pi(T+\phi)}{\frac{i+5}{10}\,\tau_{osc}}\right)\right)\,, \tag{5.4}$$

with the constraint $\sum_i \alpha_i = 1$. The coefficients $c_i = \left(\cos\left(\frac{2\pi\phi}{(i+5)/10\,\tau_{osc}}\right)\right)^{-1}$ ensure the value 1 for each cosine term at $T = 0$ separately. The function (5.4) has 14 parameters. These are the known ones from above, τ_{dec}, τ_{osc} and ϕ. The 11 coefficients α_i are a measure on the discretized distribution of oscillation timescales. The distribution ranges from 0.5 to 1.5 times the value of the simultaneously fitted τ_{osc}. The mean oscillation time might deviate from this "center" τ_{osc}.

When using (5.4) as target function to fit simulation data, the resulting measure α_i is sharply peaked around the value of the previous fit in (5.3). The mean value deviates only slightly from τ_{osc} in (5.3). Hence, our initial hypothesis can be assumed. We will only use this single timescale τ_{osc} to indicate the oscillation timescale for further analysis[1].

Interpretation of timescales

We have two distinct timescales in our model of the adaptation process, τ_{dec} and τ_{osc}. Usually, they are *not* equal, although they are within one order of magnitude of each other (for the parameters we checked). When choosing simulation parameters that the population exhibits a dynamics in the *periodic selection* regime, the anti-correlation peak in the ACF, *i.e.* the first negative minimum, vanishes. The population enters this regime when the mutational supply drops below one mutation per generation, $N\mu < 1$. For a large mutational supply $N\mu > 1$, the ACF starts to show a oscillations with a timescale τ_{osc}. Then, the population is in the *clonal competition* regime.

In the staircase mutation kernel, the first scenario with periodic selection is simple to describe: when the mutational supply $N\mu$ is low, then only a single fitness class is populated. The time until the next fitness class is established does not produce a clear signal in the ACF. The distribution of waiting times for this establishment is broad. In turn, the fitness variance is *not* anti-correlated with itself at a later time. The ACF just decays to zero, but not with an exponential tail.

Simulation results for all mutation kernels suggest, that the decay time τ_{dec} depends only on the combined parameter $N\mu$. The actual values of the population size N and mutation rate μ do not matter, as long as their product is identical.

The dependence of the oscillation times τ_{osc} on the parameters in the simulation is more complicated. As mentioned above, when the mutational supply increases above the threshold $N\mu > 1$, the population density starts "breathing": it oscillates back an forth between states of small and large variance. In the clonal competition regime of the dynamics several lineages simultaneously try to accumulate beneficial mutations, independent from other lineages in the same population. When a single lineage accumulated enough beneficial mutation to sweep through the population, the variance will be low after this sweep event. However, it usually takes several beneficial mutations to trigger such a sweep.

Mathematically speaking, this "breathing" behavior of traveling waves model is expected when two conditions are met. First, a constraint has to restrict population sizes. Secondly, the whole population has to be modeled stochastically, *i.e.* the occupancy of *all* fitness classes is allowed to fluctuate. Then, the "vertical" fluctuations from genetic drift will inevitably create the large "horizontal" fluctuations in fitness variance (and adaptation speed in turn).

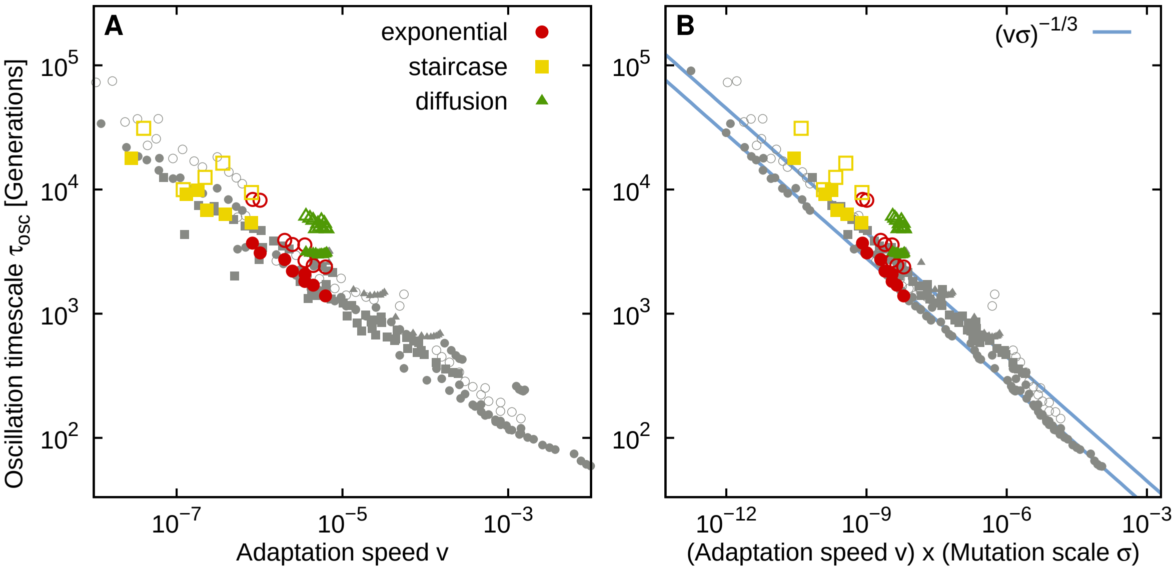

Figure 5.4.: **Scaling of the oscillation time τ_{osc} with adaptation speed v.** Measurement of oscillation times was carried out for a wide range of parameters. Circles correspond to results in the exponential mutation kernel, squares represent the staircase mutation kernel, and triangles indicate the diffusion mutation kernel. Empty symbols are results from tuned models, while data indicated with full symbols is obtained with a fixed-N constraint. Simulation data in colored points has also been shown in Figures 5.3, D.1, D.2, D.3 with additional details. Data in gray is presented only here. A few observations are imminent. Tuned models usually result in a larger oscillation time τ_{osc} for identical adaptation speed. The oscillation time in the exponential and staircase mutation kernel scales with $\tau_{osc} \sim (v\sigma)^{-1/3}$ (blue lines). Oscillations in the diffusion mutation kernel seem almost independent of adaptation speed, the main contribution is $\tau_{osc} \sim D^{-1/3}$. Simulation parameters are $N = 10^4 \dots 10^8$, $\mu = 10^{-6} \dots 10^{-3}$, $\sigma = 10^{-4} \dots 10^{-2}$, $D = 10^{-9} \dots 10^{-7}$ in various combinations. Colored points correspond to $\sigma = 10^{-3}$ or $D = 10^{-9}$.

5.2.4. Scaling of oscillation times

Although a theoretical argument to explain the oscillation timescale is still lacking, one simulation result seems almost universal. When plotting the oscillation time τ_{osc} against the adaptation speed v, a clear trend is visible (see Figure 5.4). With increasing adaptation speed, the oscillation time decreases. This relation can be quantified further:

---

[1] Computing the Fourier transform of the trajectory in fitness variance produces also a very narrow peak, with the center close to $2\pi/\tau_{osc}$.

using the combined parameter $v\sigma$, the oscillation time τ_{osc} is given by the simple scaling

$$\tau_{osc} \sim \left(v\sigma\right)^{-1/3} . \tag{5.5}$$

The prefactor is of order $\mathcal{O}\left(1\right)$ and depends on the mutation kernel. A quantity of the dimension $v\sigma$, referred to as *fitness flux*, has also been identified as crucial parameter for adaptation in changing environments [ML10].

For the exponential mutation kernel, the prefactor in the scaling (5.5) is ≈ 2.8. Deviations from this scaling occur for extreme values of any of the parameters N, μ and σ. However, the overall scaling seems robust and holds over 8 orders of magnitude. The behavior of the oscillations in the staircase mutation kernel is similar. Here, the prefactor in (5.5) is estimated as ≈ 4.5.

The scaling in the diffusion mutation kernel is different. There, the main contribution to the oscillation time τ_{osc} is dominated by the diffusion constant, $\tau_{osc} \sim D^{1/3}$. Different adaptation speeds affect the oscillation time only slightly.

5.3. Fitness measurements in tuned models

We can use tuned models to obtain exact profiles for the mean stationary population density $\overline{\mathbf{c}_\star}$ and the fixation probability $\mathbf{u}_\star$. Additionally, the choice of the mutational kernel $\mathbf{m}$ is a crucial model parameter. From these three profiles in fitness space, several other dynamically relevant distributions can be derived. In section 3.5 we introduced for instance the *distribution of fixating clones* $\mathbf{g}_\star = \mathbf{u}_\star\overline{\mathbf{c}_\star}$ and the *distribution of fixating mutations* $\boldsymbol{\rho}_f$. Due to the importance of $\mathbf{g}_\star$ in our tuned models we present this distribution together with $\mathbf{u}_\star$ and $\overline{\mathbf{c}_\star}$ for all three mutation kernels for a variety of parameter sets in Figures 5.5, 5.6 and 5.7.

The time scales in the ACF of fitness variance can be converted into fitness scales. $v\tau_{osc}$ and $v\tau_{dec}$ indicate the mean fitness increase of the population density during one oscillation period or during the decay of the ACF, respectively. These two values can be compared to mean values and variances of the distributions mentioned above.

However, the results on this comparison are not yet conclusive. Preliminary results are presented in appendix D. The fitness increase during these two timescales is of the same order of magnitude as all characteristic fitness scales in the profiles, but a clear trend cannot be distinguished. In general, timescales are hard to extract from the ACF. In Figure 5.4 we presented oscillation times τ_{osc} for a huge set of parameters. The scaling of oscillation times with exponent of $-\frac{1}{3}$ is visible for the large parameter set presented in the figure, but the colored data points, that indicate a single choice of σ, scatter around this scaling result.

We conjecture, however, that the distribution of fixating clones $\mathbf{g}_\star$ might play are role in determining the correct fitness scale during τ_{osc}. All other distributions mentioned later in appendix D only consider the effect of a *single* mutation. However, the oscillations in fitness variance occur usually due to the sweep of a whole *cohort of mutations*. Our conjecture is also based on the observation, that in the exponential mutation kernel, the oscillation times τ_{osc} tend to cluster in a similar fashion as the profiles for $\mathbf{g}_\star$ in

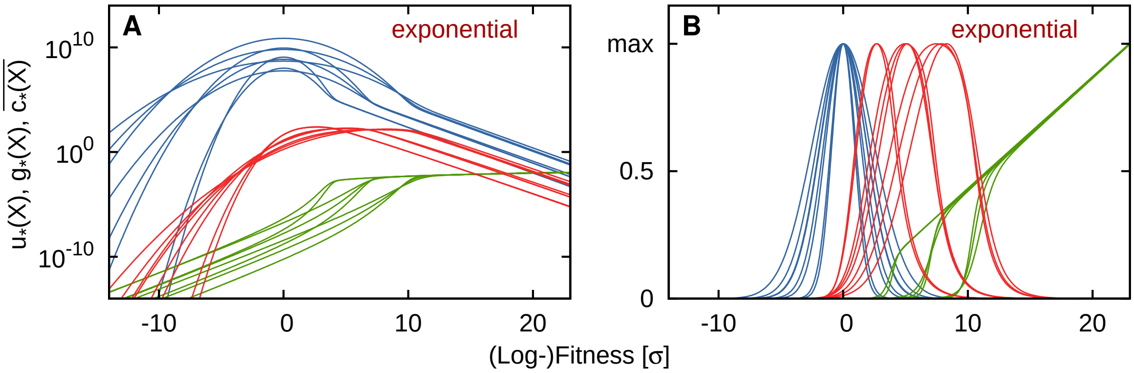

Figure 5.5.: **Population density $\overline{\mathbf{c}_\star}$ (blue), fixation probability $\mathbf{u}_\star$ (green) and distribution of fixating clones $\mathbf{g}_\star$ (red) for the exponential mutation kernel.** The value for adaptation speed $\overline{v}$ was obtained as measurement in fixed-N simulations. Original simulation parameters are $\left(\log_{10} N, -\log_{10}\mu\right) \in \left\{(5,4),(5,4),(6,4),(6,5),(6,6),(7,5),(7,6),(8,6)\right\}$ and $\sigma = 10^{-3}$. In the numerics for the tuned models the parameter population size is replaced by adaptation speed. Results cluster into several groups. Profiles with a similar stochastic edge X_c in the fixation profiles $\mathbf{u}_\star$ agree in the combined parameter $\overline{N}\mu$. The mean stationary population density $\overline{\mathbf{c}_\star}$ has a roughly Gaussian shape up until to the stochastic edge. For fitness values larger than X_c, $\overline{\mathbf{c}_\star}$ decays exponentially. The distribution of fixating clones $\mathbf{g}_\star$ is rather broad. It exhibits the same clustering with identical $\overline{N}\mu$.

Figure 5.5. When the combined parameter $N\mu$ is similar, also the oscillation times are comparable. Furthermore, the oscillation times in the diffusion mutation kernel are almost independent of the adaptation speed, similar as the width of the distribution of fixating clones, that is almost constant for the increasing adaptation speed (as in Figure 5.7).

Investigating details of this oscillation process is still left for future work.

5.4. Experimental evidence

The scaling of the oscillation times τ_{osc} found in the exponential and staircase mutation kernel can also be used to deduce an interesting result from experimental data. In section 1.2.2 we described the *long-time evolution experiment* in the Lenski lab, where 12 strains of *E. coli* have been grown with daily transfers since 1988. At the time of writing this thesis, roughly 60000 generation have passed. Data on the initial relative fitness increase of the twelve strains is freely available on the project's webpage [Len15]. In a recent publication [WRL13b] the fitness increase over the first 50000 generation has been provided. The raw data accompanying this latter publication is available in the

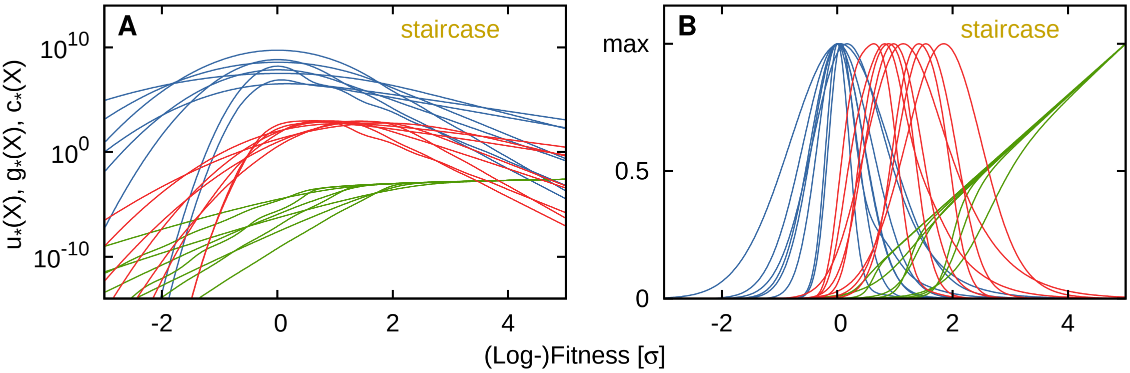

Figure 5.6.: **Population density $\overline{c_\star}$ (blue), fixation probability $u_\star$ (green) and distribution of fixating clones $g_\star$ (red) for the staircase mutation kernel.** For small adaptation speeds and/or population sizes, the numerical solution for the fixation probability oscillates slightly with a period of σ. This fact was reported in [Fis13]. Overall, the behavior of solutions in the staircase mutation kernel is comparable to the exponential mutation kernel. Simulation parameters were chosen such that the population size ranges from $\overline{N} \approx 5 \cdot 10^3 \ldots 5 \cdot 10^6$ and the mutation rate $\mu = 10^{-6} \ldots 10^{-4}$. The mutation scale was set to $\sigma = 10^{-3}$. The numerical code for this mutation kernel is unstable for large population sizes and/or adaptation speeds due to the singular mutation kernel. Therefore, we cannot present similar large population sizes and/or adaptation speeds as in the exponential mutation kernel.

Dryad Digital Repository [WRL13a].

The increase in relative fitness over time has already been shown in Figure 1.1. We present this data again in Figure 5.8, but now with additional information. We extract estimates for the adaptation speed and the oscillation time from these experimental data sets.

In order to obtain the (average) adaptation speed, we use a single linear fit through the increase in fitness over the whole time span. We define adaptation speed as the slope of the resulting fit. This most likely introduces errors, but we are only interested in a rough estimate of adaptation speed. Clearly, as indicated by the data, adaptation slows down over time. This is also reproduced by the two values we obtain. For the first 10000 generations, adaptation speed is estimated as $v = 4 \cdot 10^{-5}$, while for 50000 generations we retrieve $v = 10^{-5}$.

We also compute the ACF of adaptation speed to estimate the oscillation time. This is done by first calculating the current speed of adaptation between two consecutive data points as difference in relative fitness divided by the difference in time. The insets in Figure 5.8 show the rather noisy trajectory of the current adaptation speed. Note

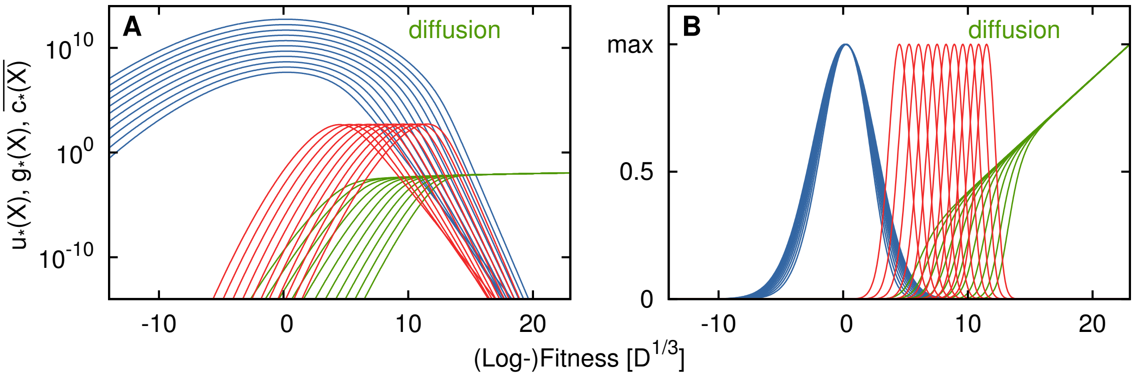

Figure 5.7.: **Population density $\overline{\mathbf{c}_\star}$ (blue), fixation probability $\mathbf{u}_\star$ (green) and distribution of fixating clones $\mathbf{g}_\star$ (red) for the diffusion mutation kernel.** All solutions were obtained with the same diffusion constant $D = 10^{-9}$. Adaptation speed is first measured in stochastic simulations with a fixed population size ranging from 10^5 up to 10^{10}. Increasing the population size N (with factors of $10^{0.5}$) increases the stochastic edge X_c in fixation probabilities almost linearly. In turn, the distribution of fixating clones shifts to higher higher fitness. However, the width of $\mathbf{g}_\star$ is almost unaffected, the range of fitness where ancestors for the future population emerge is roughly constant.

that at the beginning of each of the data sets, the time differences are smaller and more data points are available. In order to be able to calculate the ACF, we take the smallest common multiple of all time differences. This value is 500 generations in the first data set (10000 generations) and 2000 generations for the second dataset (50000 generations). Data on intermediate points is discarded. The ACF has been calculated for all twelve (respectively nine in the longer dataset) time series on fitness. The resulting curves have been averaged.

The ACF of real data shows a negative minimum at the first non-zero time point. After that, the value of the ACF fluctuates around zero. We argue, that this first anti-correlated time point is significant and represents the oscillations investigated in this chapter. In appendix E we present the ACF of adaptation speed for other time intervals than the values mentioned above (500 and 2000 generations). Although still quite noisy, these additional results strengthen our hypothesis of an oscillating adaptation speed in the Lenski experiment. To investigate these effects further, additional data would be needed. However, measuring the relative fitness (compared to a wild-type strain) in experiments with enough replicates is usually a cumbersome task.

We assume the oscillation time τ_{osc} is indicated by twice the time to the minimum in the first data point. Then, adaptation speed went through one oscillation period with passing its minimum and reaching its maximum again. As with the estimate for the

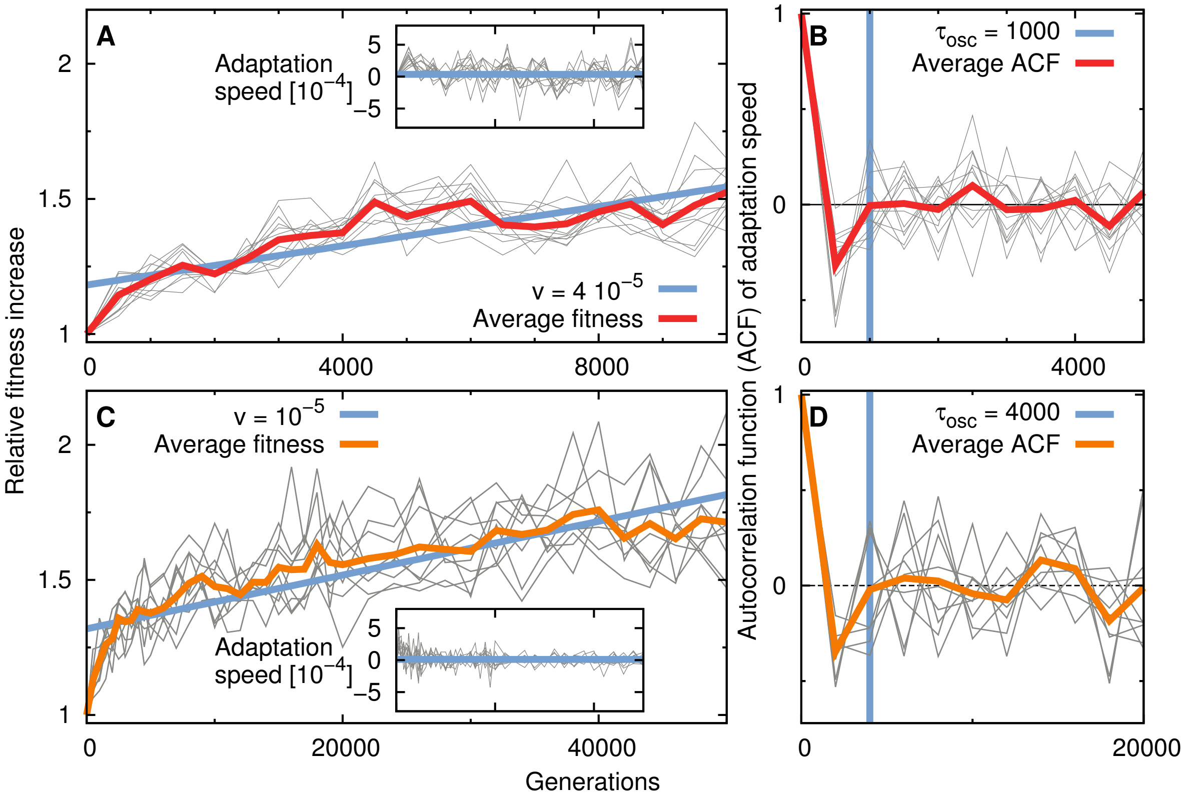

Figure 5.8.: **Experimental results on the fitness increase in the Lenski experiment during the first 10000 generations (upper panels) and first 50000 generations (bottom panels).** The data on the relative fitness for the first 10000 generations is available on the project's webpage [Len15]. Data for 50000 generations have been published in [WRL13b, WRL13a]. Time traces in (A) and (C) have already been shown in Figure 1.1. In insets we present the numerically evaluated adaptation speed. The time of the insets covers the complete dataset. From trajectories of adaptation speed we compute its autocorrelation function (ACF): (B) We use the adaptation speed every 500 generations and compute the ACF from these values. An anti-correlation at the first time point at $T = 500$ is visible. Hence, we estimate the oscillation time as $\tau_{osc} \approx 1000$. (D) We compute the ACF using only values 2000 generations apart and discard the rest. Similar to before, the value at $T = 2000$ generations is anti-correlated. We estimate the oscillation time as $\tau_{osc} \approx 4000$ in this longer time-series. Additional details can be found in the main text.

adaptation speed above, we do not expect this value to be exact. However, with the current situation on available data, we cannot obtain better estimates.

However, we can combine our result on the scaling, $\tau_{osc} \sim \left(v\sigma\right)^{-1/3}$, and the values for

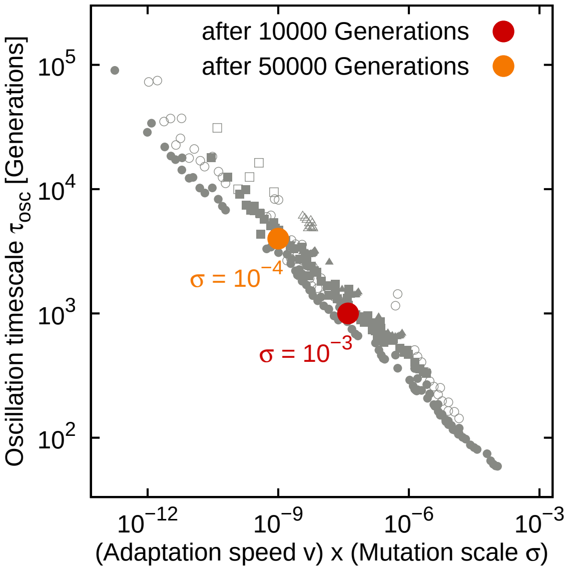

Figure 5.9.: **Scaling of oscillation time τ_{osc} with adaptation speed v times mutational scale σ with experimental data added.** Data from simulations in gray is identical with Figure 5.4. Figure 5.8 indicates the adaptation speed and the oscillation time in the Lenski experiment. The mutational scale for the experimental data is chosen such that they fit onto the scaling $\tau_{osc} \sim \left(v\sigma\right)^{-1/3}$. For the first 10000 generations [Len15] the mutational scale is estimated as $\sigma^{(10000)} \approx 10^{-3}$. For the larger dataset of 50000 generations [WRL13b] the mutational scale is estimated as $\sigma^{(50000)} \approx 10^{-4}$.

the quantities adaptation speed and oscillation time in experimental data. This allows us to give an estimate of the mutational scale σ. Inverting the scaling relation (5.5) yields

$$\sigma \approx \frac{c^3}{\tau_{osc}^3 v} \,. \tag{5.6}$$

The value of the constant c is assumed to be roughly 5 in accordance with the scaling obtained in the exponential and staircase mutation kernels. Notice that the large uncertainty in the oscillation timescale is amplified with the exponent -3 in (5.6).

For the experimental data for the first 10000 generations we use the values $v = 4 \cdot 10^{-5}$ and $\tau_{osc} = 1000$. This results in a mutational scale of

$$\sigma^{(10000)} \approx 10^{-3} \,. \tag{5.7}$$

For the larger dataset on the first 50000 generations, we use the values $v = 10^{-5}$ and $\tau_{osc} = 4000$ and arrive at

$$\sigma^{(50000)} \approx 10^{-4} \,. \tag{5.8}$$

In Figure 5.9 we show the results on the mutation scale in the experiment together with all previous simulations on the scaling of oscillation time τ_{osc} with adaptation speed v.

The difference of the two values in (5.7) and (5.8) is not surprising. The slow down of adaptation has been mentioned already above. A smaller mutation scale might be a simple explanation for this fact: mutations with large effects are incorporated early in the adaptation process, over time only mutations with a smaller beneficial effect remain available. However, other effects also can slow down adaptation. Epistatic interaction between mutations can reduce the actual effect.

However, we still regard (5.7) and (5.8) as an interesting result. Usually, the underlying mutation scale of adaptation is obscured in the dynamical process. All observed mutations are conditioned on being present in the genomes of enough individuals such that they actually become observable. Especially in the clonal competition regime, many mutations are "wasted“ on lineages that go extinct. Here, we are now able to calculate the average size of a novel mutation. However, currently these values are at best rough estimates.

6. Applications of tuned models: oasis model

Part of this thesis was published already as [GH13] in a special issue entitled "Statistical Mechanics and the Dynamics of Evolution":

GEYRHOFER L, AND HALLATSCHEK O. Stochastic delocalization of finite populations. *Journal of Statistical Mechanics: Theory and Experiment*, **2013**(01):p. P01007. 2013

The topic of this chapter is not directly related to other results in this thesis, although the mathematics behind the tuned models is quite similar. In that respect, this chapter can be seen as a single, self-contained part.

Only parts of the original text have been written by myself. Here we only briefly summarize our solution. Explicit mathematical derivations and further details can be found in the original publication. Figures reproduced here have been slightly modified to match others in this thesis. The abstract and all captions for the figures have been taken directly from the publication.

Abstract

Heterogeneities in environmental conditions often induce corresponding heterogeneities in the distribution of species. In the extreme case of a localized patch of increased growth rates, reproducing populations can become strongly concentrated at the patch despite the entropic tendency for population to distribute evenly. The localization of populations of replicating bacteria, viruses or autocatalytic chemicals arises in the various contexts, such as ecology, evolution, medicine or chemistry. Several deterministic mathematical models have been used to characterize the conditions under which localized states can form, and how they break down due to convective driving forces. It has been repeatedly found that populations remain localized unless the bias exceeds a critical threshold value, and that close to the transition the population is characterized by a diverging length scale. These results, however, have been obtained upon ignoring number fluctuations ("genetic drift"), which are inevitable given the discreteness of the replicating entities. Here, we study the localization/delocalization of a finite population in the presence of genetic drift. The population is modeled by a linear chain of sub-populations, or demes, which exchange

migrants at a constant rate. Individuals in one particular deme, called “oasis”, receive a growth rate benefit, and the total population is regulated to have constant size N. In this ecological setting, we find that any finite population delocalizes on sufficiently long time scales. Depending on parameters, however, populations may remain localized for a very long time. The typical waiting time to delocalization increases exponentially with both population size and distance to the critical wind speed of the deterministic approximation. The effective death rate serves as an order parameter in our model. As long as the population remains localized, there must be an overall positive death rate in order to balance the production of new individuals at the growth rate. However, as the population approaches the delocalization transition, the death rate decays to zero. We find that genetic drift has a similar effect, driving the death rate below the deterministic expectation. We augment these simulation results by a mathematical analysis that treats the reproduction and migration of individuals as branching random walks subject to global constraints. For a particular constraint, different from a fixed population size constraint, this model yields a solvable first moment equation. We find that this solvable model approximates very well the fixed population size model for large populations, but starts to deviate as population sizes are small. Nevertheless, the qualitative behavior of the fixed population size model is reproduced very well. In particular, the analytical approach allows us to map out a phase diagram of the order parameter as a function of the two driving parameters, inverse population size and wind speed. Our results may be used to extend the analysis of delocalization transitions to different settings, such as the viral quasi-species scenario.

6.1. Motivation

In previous chapters, we focused on evolutionary applications of BRWs. Here, we turn our attention to ecological problems. In the main interpretation of the model, differences in growth rates of individuals are not due to accumulating (beneficial) mutations, but rather individuals are subject to external heterogeneous environmental conditions. Moreover, movement in space is not a mutation process, but locomotion in an extended habitat.

Heterogeneities in environment are ubiquitous. We focus on environments exhibiting a narrow region of enhanced growth, that does *not* deplete on time scales relevant for reproduction of the population. Microbial experiments can be adjusted to match such conditions. In [NPLK00, LMTO+04] the authors create this growth spot by shielding a colony with a small platelet from UV light illuminating the rest of an agar plate. In the shade of the platelet microbes can grow, whereas outside they are killed by hazardous UV light.

However, also natural examples of such non-depleting growth regions exist. An oasis in a desert is the prime example: localized water supply leads to a confined, stable ecosystem. It allows organisms to grow within a bounded region, but (usually) death awaits them if they move too far away for too long. Another important example are hydrothermal vents in deep sea. It has been speculated that life itself might have originated near such localized nutrient sources [Wäc90]. Nowadays whole ecosystems form in the vicinity of these underwater structures [Tun91, TMM98]. However, ocean currents constantly try to drive organisms away.

Although the main analysis is focused on these ecological settings, the model can be applied in an evolutionary context. In Eigen's *quasi-species theory* the space coordinate denotes different genotypes [Eig71, Wil05]. In a special version of quasi-species theory, a single "master sequence" exists in genotype space. This master sequence has a higher reproduction rate than all other sequences. Mutations change the genotype, usually away from the singular master sequence. If the mutation rate exceeds a certain threshold, the excess growth rate of the master sequence cannot localize the population on it anymore. This transition to a delocalized population in sequence space has been termed *mutational meltdown* in population genetics literature [LBBG93, HRWB02, WGP03] and *error catastrophe* in quasi-species theory [Eig71, EMS89, Wil05].

The structure of this chapter is as follows. After this introduction, we start explaining stochastic simulations of the model with a fixed population size constraint. Then we switch to analytic predictions, first stating relevant results from a deterministic approach [DNS00]. After that, we show that tuned models, which were the focus of the rest of this thesis, can be applied in this varied context. We compare simulation results with a fixed population size to the exact analytic results obtained from tuned models.

6.2. Simulations

6.2.1. Simulation model

The oasis model uses a simulation code that is similar to the code for traveling waves in the previous chapters (*cf.* section 2.1). Deviations from these instructions exist, therefore we restate an independent version here.

Sub-populations are arranged on a linear chain of demes with distance $d = 1$ to each other. Individuals in one of the demes receive an additional boost in growth rate (the 'oasis'). A time step of duration ϵ consists of two sub-steps. In the first sub-step, the occupancy n_i of deme i is changed due to biased migration, aswell as the deterministic part of reproduction:

$$\Delta^{(A1)} n_i/\epsilon = D\big(n_{i+1} - 2n_i + n_{i-1}\big) + v/2\big(n_{i+1} - n_{i-1}\big) + \delta_{i0} a\, n_i \ . \tag{6.1}$$

The deme $i = 0$ indicates the oasis. Stochastic changes due to the randomness in the reproduction process are applied immediately after,

$$\Delta^{(A2)} n_i/\sqrt{\epsilon} = \sqrt{2}\big(\mathrm{Poisson}(n_i) - n_i\big) \ . \tag{6.2}$$

Poisson(n_i) a Poisson-distributed random variable with parameter n_i. This particular choice of noise ensures that (i) occupancies n_i do not drop below zero, (ii) the expectation value of the random term is zero and (iii) the variance within a deme amounts to $2\epsilon n_i$. All three conditions match the dynamics one would expect from BRWs. If ϵ is chosen small enough, higher moments of the noise term can be neglected.

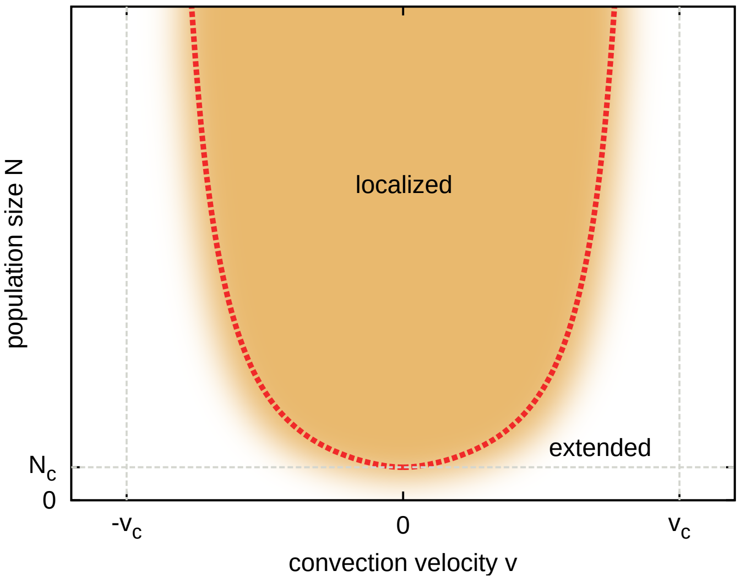

Figure 6.1.: **Hypothesized phase diagram for finite populations localized at an "oasis".** Deterministic calculations of a population localized at a favorable growth spot ("oasis") predict a sharp delocalization transition: When the convection speed exceeds a critical value v_c, populations escape from the oasis. Our stochastic simulations of populations of fixed size suggest that this picture has to be modified to account for number fluctuations. i) Populations always delocalize on sufficiently long time scales, due to number fluctuations. The *escape time*, however, depends exponentially on population size. Thus, delocalization is effectively unobservable in sufficiently large populations. ii) In the presence of convection, the escape times are strongly reduced. In an exactly solvable model with fluctuations but finite populations, a sharp transition line can be derived (red dashed line). This line also describes qualitatively the crossover from small to very large escape times (compared to the coalescence time) in fixed population size models.

After these changes in n_i, a second sub-step adjusts the occupancies n_i with a global death rate r to comply with the fixed population size constraint,

$$\Delta^{(B)} n_i / \epsilon = -r n_i \ , \tag{6.3}$$

where r is calculated as $r = 1 - N / \sum_i n_i$. N is the population size.

Periodic boundaries are imposed, such that after a delocalization event the population will be recaptured by the oasis. The dimension of the simulation box is chosen large enough that the population does not interfere with itself. We define the population to be localized if the occupancy at the oasis exceeds the threshold $n_0 > 1$. This condition is used when presenting averaged simulation data of localized populations in the following sections.

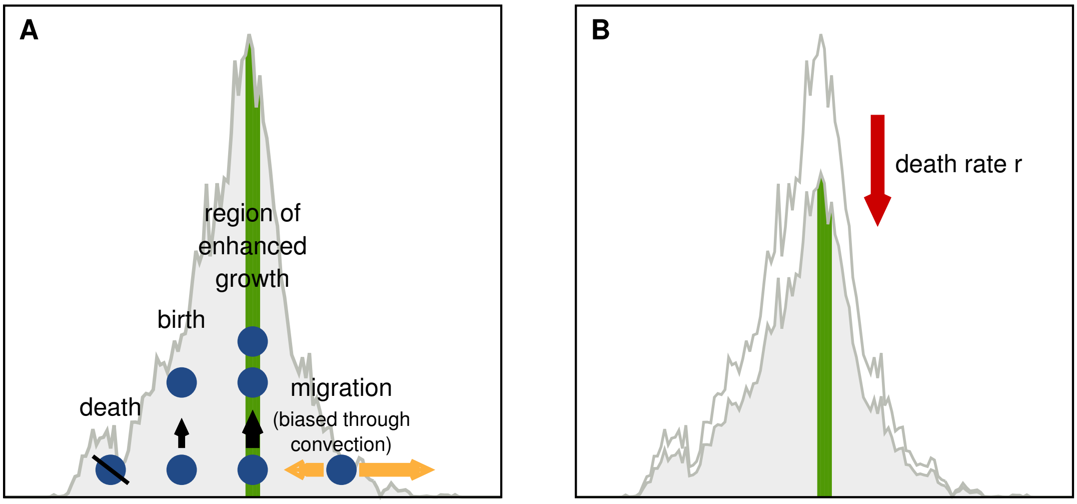

Figure 6.2.: **Stochastic dynamics of the "oasis" model.** A single time-step in our simulations consists of two sub-steps. (A) Individuals migrate between neighboring lattice sites, or demes, in general with a bias in one direction to account for convection. Number fluctuations are incorporated to account for the population turn over from one generation to the next. The sub-population located at the oasis enjoys an increased reproduction rate a. This first sub-step effectively describes a branching random walk with a branching rate that depends on the location. (B) In the second sub-step, the (global) population is regulated to comply with the constraint of fixed population size (see main text for details).

A graphical sketch of the simulation instructions is drawn in Figure 6.2.

6.2.2. Simulation results

Snapshots of the population with and without convection are shown in Figure 6.3. Localized populations are usually much narrower than the average profile, that accounts for excursions of the population.

Death rates r are recorded and averaged for localized populations. In Figure 6.4 we present values for various population sizes and convection speeds.

A delocalization of the population will always happen on sufficiently long time scales. The distribution of escape times T for the population can be captured by

$$\mathbb{P}[T] \sim \exp(-T/\tau) . \tag{6.4}$$

Deviations from this exponential distribution at small escape times are caused by fluc-

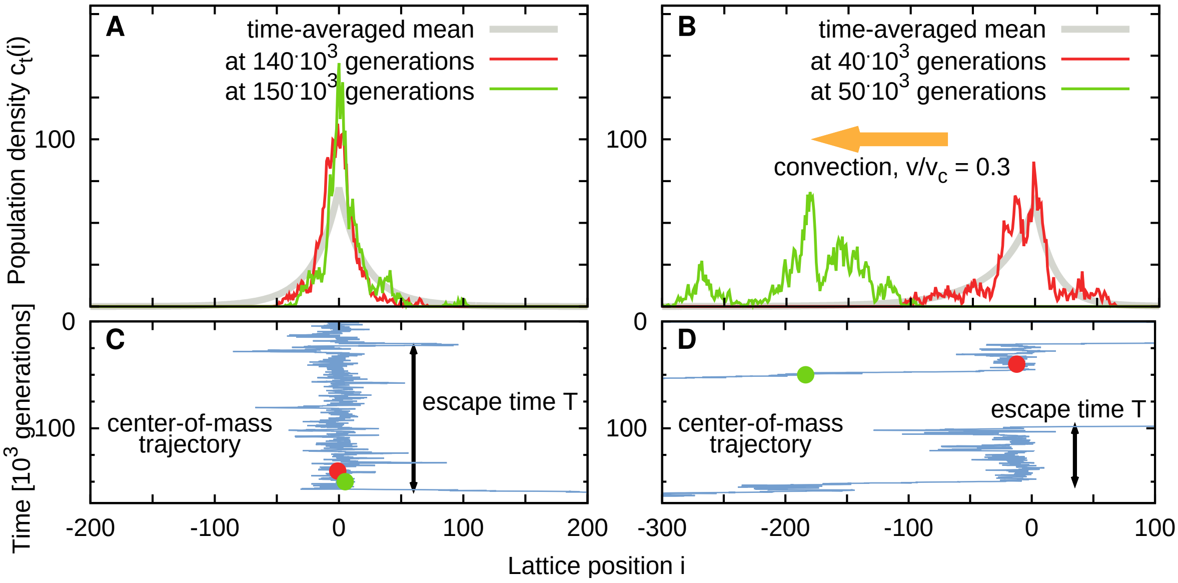

Figure 6.3.: **Snapshots of simulations with fixed population size.** Figure (a) depicts snapshots for our simulation model, in which individuals carry out an unbiased branching random walk subject to the global constraint of fixed population size (see Figure 6.2 for a description of the time step). All lattice sites (labeled by an index i) are equivalent, except for deme $i = 0$ where the growth rate is increased by $a = 0.1$. The resulting density profiles are typically quite narrowly centered around the oasis (*cf.* green and red lines). The average profile (gray line) is much broader than typical profiles, due to large excursions of the center-of-mass. These excursions can be seen in the center-of-mass trajectory at the bottom part of the figure. Notice that the population occasionally escapes from the oasis, due to a rare large fluctuation. Such delocalizations are recorded when the occupancy at the the oasis drops below a critical value, $c_0 < 1$. In the convection-less case, the population returns quickly after a random excursion (e.g. the first delocalization event in (A)). With convection, (B), the population usually traverses the (periodic) simulation box after a delocalization event until it reaches the oasis again. The escape time between two delocalizations is indicated by a black arrow. The green and red bullets indicate the spatio-temporal position of the center-of-mass corresponding to the red and green snapshots in the upper part of the figure. The convection speed in (B) is set to 30% of the critical convection speed $v_c = ad$, at which delocalization is observed in the deterministic limit [DNS00]. Notice that convection leads to more frequent escapes from the oasis. Simulation parameters were $a = 0.1$, $N = 3000$, $\epsilon = 0.2$.

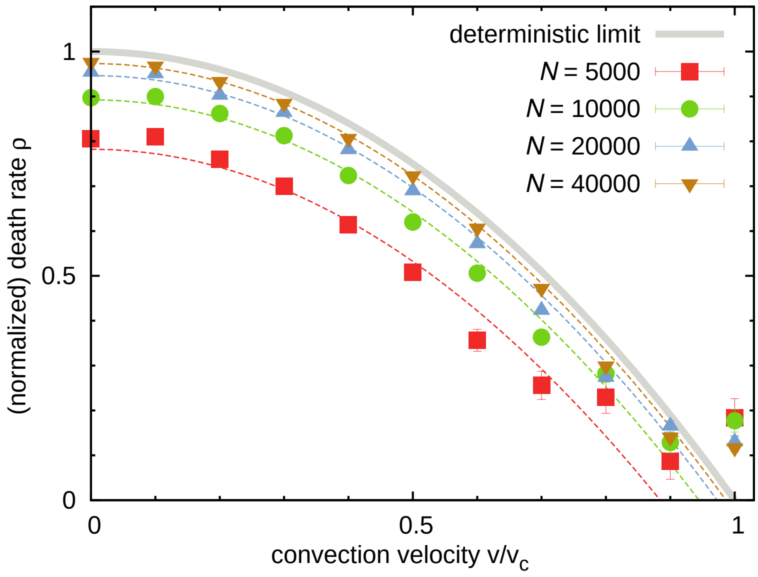

Figure 6.4.: **Mean death rates for varying population size and convection speed.** In our simulations, the death rate r is the factor with which the population has to be scaled down to keep its global size constant (*cf.* section 6.2). Here, we display the scaled death rate $\rho = \overline{r} / \left(a^2 d^2 / 4D\right)$ normalized by its deterministic no-convection limit $a^2 d^2 / 4D$. According to the deterministic limit, the death rate should decay with increasing convection speed according to $\rho = 1 - (v/v_c)^2$ (solid gray line) up to the critical convection speed $v_c = ad$ [DNS00]. For finite populations the death rates are shifted to lower values roughly by the constant $2/N$, as predicted by the analytical approach explained in section 6.3. As a consequence, smaller convection speeds are required for the death rates to cross 0, which signals the delocalization of the population. Our time average only includes states in which there is at least one individual at the oasis, $c_0 > 1$. For convection speeds larger or equal than the deterministic critical velocity, this is just a tiny minority of all states. Hence, the points for large convection speeds (close to v_c) are slightly biased. Solid lines are obtained by our theoretical approach explained in section 6.3, in particular by the relation (6.19).

tuating occupancies n_i. As mentioned above, we only define profiles with $n_0 > 1$ as localized. In the transition period of a leaving or an arriving population, this localization condition can be met and not met in fast succession. In turn, small times are significantly overrepresented in the distribution of escape times. Simulation evidence suggests that the characteristic escape time τ is approximated by the fit

$$\tau \sim \exp\bigl(N/\eta\bigr) \ , \tag{6.5}$$

with $\eta \approx \beta(1 - v/v_c)$ and the numerical constant $\beta \approx 560$. Figure 6.5 summarizes these

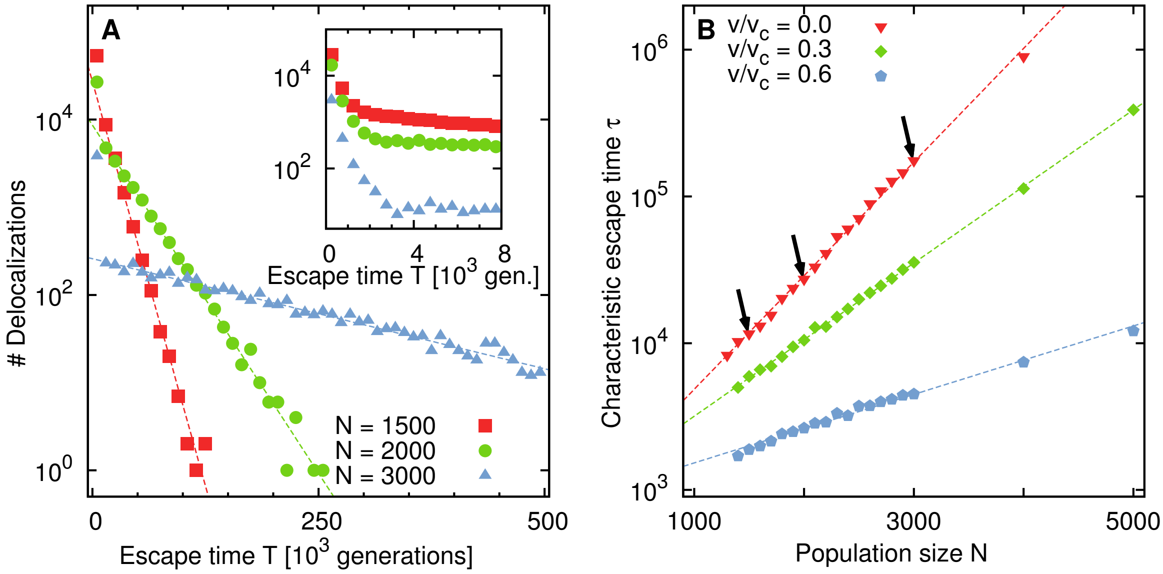

Figure 6.5.: **Distribution of escape times T measured in simulations of fixed population sizes.** (A) Distribution of escape times T for different population sizes in the absence of convection. In all cases, the distribution is close to exponential. Deviations at early times (inset) are attributed to shape fluctuations, see main text. From the exponential tail of the distribution, a characteristic escape time τ can be extracted. This time strongly depends on the population size and convection velocity, as shown in figure (B). Solid lines in figure (B) represent the fit in (6.5). The three black arrows indicate the characteristic times τ of the distributions shown in (A). The simulation results were obtained using parameter values $a = 0.1$, $\epsilon = 0.2$.

findings. Hence, we conclude that if the population size is large enough, these escape events are practically unobservable.

More explanations and details can be found in our publication [GH13].

6.3. Analytical approach

6.3.1. Deterministic solution

Similar to our analytical approach in the last chapters, we switch to the continuous description in our analytic calculations, with $n_i = c_t(x_i)dx$. The model in the previous section can be captured in a continuous reaction-diffusion-advection equation,

$$\partial_t \mathbf{c}_t = D\partial_x^2 \mathbf{c}_t + v\partial_x \mathbf{c}_t + \mathbf{R}_t \ . \tag{6.6}$$

The reaction term $\mathbf{R}_t \equiv R_t(x)$ accounts for (stochastic) birth and death of new individuals. Moreover, $\mathbf{R}_t$ contains a non-linearity to keep population sizes finite in the stochastic case. If the population is large enough a deterministic approximation is valid, and we can use

$$\mathbf{R}_t = \big(\alpha\delta(x) - r\big)\mathbf{c}_t \;, \tag{6.7}$$

where individuals at the oasis ($x = 0$) enjoy the additional growth rate $\alpha = ad$. In (6.7) the death rate r is adjusted such that (6.6) admits a steady state.

The stationary solution to the deterministic case is given by the profile [DNS00]:

$$c_\star(x) \sim \exp\left(\frac{-\alpha|x| + vx}{2D}\right) \;. \tag{6.8}$$

The rescaled growth rate $\rho \equiv 4Dr\alpha^{-2}$ decays to zero with increasing convection speed,

$$\rho = 1 - \big(v/v_c\big)^2 \;. \tag{6.9}$$

Upon reaching the critical convection speed $v_c = \alpha$, the stationary solution ceases to exist, as easily verified by inserting this value into (6.8). The elaborate analysis in [DNS00] predicts extended (delocalized) states with complex eigenvalues of the Louivillean for such large convection speeds.

6.3.2. Tuned models

The match the oasis model to the analysis done in the rest of this thesis, we resort back to the original description of BRWs. Recall the dynamics of BRWs in its Langevin form, (3.1). Previously it included selection, mutation and genetic drift. We keep the same structure and state

$$\mathbf{c}_{t+\epsilon} - \mathbf{c}_t = \epsilon\big(\mathcal{S} + \mathcal{M}\big)\mathbf{c}_t + \sqrt{\epsilon}\sqrt{2\mathbf{c}_t}\boldsymbol{\eta}_t \;. \tag{6.10}$$

The former mutation operator $\mathcal{M}$ now represents movement in space. The operator $\mathcal{S}$ still constitutes the differential growth at different positions in space. Reproduction noise, *i.e.* genetic drift, is identical in both models. In the context of the oasis model, the explicit expressions for $\mathcal{M}$ and $\mathcal{S}$ are:

$$\begin{aligned}
\mathcal{M}\mathbf{c}_t &= \big(D\partial_x^2 + v\partial_x\big)\mathbf{c}_t \;, && (6.11\text{a})\\
\mathcal{S}\mathbf{c}_t &= \big(\alpha\delta(x) - r\big)\mathbf{c}_t \;. && (6.11\text{b})
\end{aligned}$$

From here, we can construct tuned models identical to section 3.2: first, the fixed population size constraint is relaxed to allow for weighting with a general function $\mathbf{u}_t$,

$$1 = \big\langle \mathbf{u}_t \big| \mathbf{c}_t \big\rangle = \int dx \; u_t(x) c_t(x) \;. \tag{6.12}$$

The population size fluctuates for any other choice than a constant weighting, $u_t(x) = 1/N$. The constraint (6.12) is coupled infinitesimally to the generator (6.10) and the series expansion in orders of the time step ϵ is carried out. In the next step we chose the constraint such that higher moments, *i.e.* $\overline{\mathbf{c}_t\mathbf{c}_t}$, decouple from the dynamics of the mean population density, $\overline{\mathbf{c}_t}$. We arrive at the two (already familiar) equations[1],

$$\begin{aligned} -\partial_t \mathbf{u}_t^\star &= \big(\mathcal{M}^\dagger + \mathcal{S}^\dagger - 2\mathbf{u}_t^\star\big)\mathbf{u}_t^\star \; , && (6.13a) \\ \partial_t \overline{\mathbf{c}_t} &= \big(\mathcal{M} + \mathcal{S} - 2\mathbf{u}_t^\star\big)\overline{\mathbf{c}_t} \; . && (6.13b) \end{aligned}$$

The dichotomy between the two setups of fixed population sizes in simulations versus the tuned constraint now manifests in the following: in the simulations of section 6.2, we had a fixed population size N and measure the mean death rate,

$$\overline{r} = \overline{r}(N) \; . \tag{6.14}$$

In tuned models, we set the death rate as parameter, and obtain a mean population size,

$$\overline{N} = \overline{N}(r) \; . \tag{6.15}$$

In order to obtain this relation for the population size, we have to investigate the *stationary* solutions for the population density and the weighting function. Explicitely, these two profiles are governed by the two differential equations,

$$\begin{aligned} 0 &= D\partial_x^2 \mathbf{u}_\star - v\partial_x \mathbf{u}_\star + \big(\alpha\delta(x) - r\big)\mathbf{u}_\star - 2\mathbf{u}_\star^2 \; , && (6.16a) \\ 0 &= D\partial_x^2 \overline{\mathbf{c}_\star} + v\partial_x \overline{\mathbf{c}_\star} + \big(\alpha\delta(x) - r\big)\overline{\mathbf{c}_\star} - 2\mathbf{u}_\star\overline{\mathbf{c}_\star} \; . && (6.16b) \end{aligned}$$

The constraint, $1 = \int dx\; u_\star(x) c_\star(x)$, sets the scale of the mean population density. The equations depend on two control parameters, $\rho = 4Dr\alpha^{-2}$ and $\nu = v\alpha^{-1}$. Both parameters only admit values in the interval $\big[0;1\big]$.

By analogy with a mechanical problem, we can solve the convection-less case ($v = 0$) analytically. Details on the derivation can be found in [GH13]. The stationary weighting function $\mathbf{u}_\star$ is given by

$$u_\star(x) = \frac{3Dr}{4\alpha^2}\left[\sinh\left(\sqrt{\frac{Dr}{4\alpha^2}}\,|x| + \operatorname{artanh}\left(\sqrt{4Dr\alpha^{-2}}\right)\right)\right]^{-2} \; . \tag{6.17}$$

Due to the similarity of (6.16b) and (6.16a) for $v = 0$, both profiles are proportional to each other. In particular, we have $\overline{\mathbf{c}_\star} = \big\langle \mathbf{u}_\star | \mathbf{u}_\star \big\rangle^{-1} \mathbf{u}_\star$.The stationary mean population size $\overline{N_\star}$ is computed by integrating $\overline{\mathbf{c}_\star}$:

$$\overline{N_\star} = 2\left(\frac{\alpha^2}{8D} + \frac{\alpha}{4\sqrt{D}}\sqrt{r} - r\right)^{-1} \; . \tag{6.18}$$

Thus, (6.18) is now the desired relation between population size and death rate. With convection, $\nu > 0$, we employ a “small velocity approximation” in the calculations. Thus, the inversion of relation (6.18) to an expression for the death rate acquires a correction term for convection:

---

[1] Recall the notation we introduced in section 3.2. Profiles with a asterisk as lower index were stationary solutions in a comoving frame. Although we do not have this comoving frame here, we use $\mathbf{u}_\star$ and $\overline{\mathbf{c}_\star}$ for *stationary* solutions. As before, $\mathbf{u}_t^\star$ is *any* (even time-dependent) solution that renders the coupling to higher moments zero.

$$r \approx \left(\frac{\alpha}{4\sqrt{D}} + \sqrt{\left(\frac{3\alpha}{8\sqrt{D}}\right)^2 - \frac{2}{\overline{N_\star}}} \right)^2 - \frac{v^2}{4D} \,. \tag{6.19}$$

Notice that for $\overline{N_\star} \to \infty$ we recover the original (deterministic) relation (6.9), while for $v \to 0$ we arrive at the direct inverse of (6.18). Ultimately, relation (6.19) has been used for the (fuzzy) phase boundary in the phase diagram pictured in Figure 6.1, but also in Figure 6.4 to predict the death rates ρ depending on the population size and the convection speed.

6.4. Discussion

In this chapter we investigated a (de-)localization problem of finite populations. We employed branching process methods, similar to those in the rest of this thesis: Individuals reproduce and migrate in demes arranged along a linear chain, with individuals in one deme receiving a boost in growth rate (the "oasis"). The aim was to derive conditions upon which the population remains pinned at the oasis. Simulations of the model indicate that population will always delocalize when reproduction is stochastic. The tail of escape times of the population is captured by an exponential distribution, (6.4). The characteristic escape time depends itself exponentially on two parameters: convection speed v and population size N. This relation is indicated in (6.5). Due to this double exponential dependence, the boundary between localized and delocalized states appears sharp: the transition from short to long escape time is rapid. This extends a previous deterministic analysis [DNS00] predicting a delocalization occurring at a critical convection v_c.

An exact theoretical boundary has been obtained by at first relaxing the assumption of a constant population size. A special constraint can be tuned such that the dynamics of mean population density decouples from higher moments. These tuned models are a slight variation of the original model, but they create exact tractability. Contrary to the traveling wave models treated in the rest of this thesis, we were able to analytically compute and state a closed form of the solution. Thus, we can define the delocalization transition to be at those parameter values where our solutions cease to exist.

This chapter only constitutes a short summary of [GH13]. Additional explanations and more details can be found in the original publication.

7. Discussion

This thesis dealt with models of adaptation in microbial populations. Experimental evidence, summarized in section 1.2, indicates that populations of these microbial organisms are genetically highly diverse when evolved over many generations in laboratories: often, many beneficial mutations are present simultaneously and interfere on their way to fixation or extinction.

The ensuing dynamics of alleles within the population is highly influenced by these interfering beneficial mutations. Sometimes, albeit rarely, stable polymorphisms of multiple clones evolve to coexist, each with their own "ecological" niche [BBL08, LGPH+12, HD13, PHLG+14]. However, in asexual populations of microbes, usually only one of the beneficial mutations will outcompete the others and fixate. This dynamical regime of clonal competition has been investigated in recent years with *traveling wave models*.

7.1. Traveling waves are simplified models of adaptation

Traveling wave models of adaptation have been considered in an evolutionary context for almost two decades [TLK96, RWC03, CKL05, DF07, PK07, BRW08, RBW08, SG10, PSK10, Hal11b, GRB+12]. Many interesting analytical observations have been made while using them. However, compared to the real experimental situation, these models make considerable simplifying assumptions.

The main simplifying assumption is that the description of individuals is reduced to a single property: their growth rate. The observed diversity in experiments is implemented by a density in a one-dimensional fitness space, which remains only a caricature of the real population. On average the ensuing density is bell-shaped. When modeling adaptation, parameters are chosen such that the mean fitness of the population increases over time (on average). This soliton-like behavior results in the label of "traveling waves".

Often, and so did we, traveling wave models of adaptation focus on a dynamical regime, where the population is infinitely crawling up a *linear* fitness gradient. Our selection operator $\mathcal{S} \equiv \left(x - \overline{x_t}\right)$ in (3.1) and (3.15) indicates this linearity. This choice clearly helps understanding processes happening during adaptation, but real data indicate that fitness does not increase linearly on long timescales (see also Figure 1.1). This phenomenon, called *diminishing returns*, decelerates adaptation over time.

A recent publication [WRL13b] tried to fit the data in the long-time evolution experiment [Len15] of the first 50000 generations to different algebraic relations for fitness increase over time (see Figure 1.1B for an illustration of the data). In particular, the data was fitted to a relation of the form

$$\overline{x_t} \approx \left(\beta t + 1\right)^{\alpha} , \tag{7.1}$$

such that $\overline{x_0} = 1$. Their experimental data suggests $\alpha \approx \mathcal{O}\left(10^{-1}\right)$. This behavior, akin to a power-law, is consistent with an increase of fitness that would last for very long times. Even if the experiment would be carried on for much longer (several orders of magnitude longer), the expected relative fitness from (7.1) of the evolved cells would still remain within the range that is observed among other microbes nowadays [LC13].

A possible explanation for this slowdown could be epistasis, *i.e.* non-linear interactions between different mutations[1]. Epistasis was observed repeatedly in experiments [KDBP11, CCD+11, KDS+11, KRJD14, PHLG+14].

7.1.1. Differences and similarities of the mutation kernels

In this thesis we compared different models that incorporated variants of the mutation process. In experimental settings the mutation kernel is known as *distribution of fitness effects* [EWK07, SSF+13]. So far, only little is known about *all* the possible effects a mutation can have on the fitness of an organism. Several models have been proposed that reduced the mutational landscape of an organisms to a simple one-dimensional distribution, which we also tried to mimic in our traveling wave models by implementing different mutation kernels.

In general, the different mutation kernels lead to slightly different dynamical behavior. Mean quantities appeared quite unaffected. However, the fluctuation behavior changed more drastically. The amplitudes of the fluctuations were largest in the exponential mutation kernel and smallest in the diffusion mutation kernel.

The diffusion mutation kernel exhibited the smoothest dynamics. This kernel is thought to be mostly applicable to viral populations with enormous population sizes and large mutation rates [TLK96, CKL05, Hal11b, NH13].

The staircase mutation kernel is the simplest model for the mutation process, and has therefore gained much attention [DF07, DWF13, GD13]. Only beneficial mutations are possible, where all of them have the same effect. Many quantities are directly accessible in this setting. It has been argued that other mutation models can be mapped onto this mutation kernel, if the (positive) tail of the distribution decays fast enough [GRB+12, Fis13]. Its dynamics are slightly more erratic than the diffusion mutation kernel above.

The exponential mutation kernel is likely the most realistic of the three mutations kernel if one seeks to describe the emergence and competition of cohorts of mutations in microbial experiments.

Overall, all three mutation kernels are only simplifications of the mutation process in experimental populations. These mutation kernels are neither time-dependent nor fitness-dependent[2]. Their shape is constant during the whole adaptation process. Implementing a time-dependent or fitness-dependent mutation kernel might be a first step towards modeling of epistatic effects.

---

[1] In our formalism, this non-linearity would result in the non-additivity of the fitness of single mutations.

[2] *i.e.* the effect of a novel mutation did not depend on the fitness of the genetic background it occurred on.

7.1.2. Epistatic interactions might be implemented in higher dimensional fitness waves

Epistasis is by definition non-linear. When increasing the number of dimensions of the fitness landscape, epistatic effects between different mutations are much simpler to implement than in a single dimension, that requires additional (ad-hoc) constraints on the dynamics.

Incidentally, our extended tuned models (where the moment hierarchy is closed at a higher level) exactly exhibit the correct properties: theoretically, we could assume two (or more) different fitness dimensions instead of investigating the behavior of fluctuations in a single fitness dimension. Still, we are restricted in the types of interactions, and as seen in 3.6 we automatically acquire interaction terms between the fitness dimensions, but the method can act as starting point. However, this idea has not yet been pursued and it might be an important avenue for future directions.

7.2. Tuned models exhibit exactly solvable dynamics

One of the central topics in this thesis was the construction and use of *tuned models*. They were first proposed in [Hal11b]. A direct application of this formalism was published in [GH13] and it has been used to some extent in [GRB$^+$12, NH13].

In contrast to many previous implementations of traveling wave models, tuned models do not restrict population size to a constant value. Here, the fixation probability $\mathbf{u}_\star$ (in a comoving frame with speed v) is kept constant and the population size is adjusted to match this fact (see section 3.4).

7.2.1. Population size fluctuations

The assumption of a constant population size is not very realistic. Rarely in natural populations are birth and death events exactly synchronized. Moreover, consecutive populations do not have identical sizes in species with non-overlapping generations, *e.g.* annual plants. Nevertheless, the assumption of a constant population size was instrumental in deriving many important results in population genetics. Often, the reduction of a complicated demographic dynamics to a single number allows evolutionary scenarios to be treated with analytic expressions or powerful approximations.

Microbial populations also rarely have a constant size. In the experiments described in section 1.2 we introduced two basic experimental setups that both exhibit different population size dynamics. In liquid culture the population size grows exponentially over time until a dilution step drastically reduces the number of cells. The bottleneck size is usually several orders of magnitude smaller than the peak population size. During the experiment the process of exponential growth and dilution is repeated at regular intervals.

In the spatial experiments depicted in Figure 1.2 microbes grow centrifugally on agar. The total population size is proportional to the area of the colony and increases quadratically over time. Only a thin layer at the circumference finds the necessary conditions for growth: hence, the number of reproducing cells grows linearly over time until a fraction of the cells are transferred to the next step of the experiment.

In the light of these experimental considerations, a population genetic model with fluctuating population size is not surprising. Fluctuations of population size in tuned models were investigated in detail in chapter 4. However, we did not want to study a complicated demographic scenario: the fluctuating population size is just the by-product of our intention to construct an *exactly solvable* model.

However, having fluctuating populations size helped us to discover a peculiarity of traveling wave models: apparently, the population size enters the model not as the logarithm of the mean population size ($\log \overline{N}$), but as the mean of the logarithm of the population size ($\overline{\log N}$). Adaptation in these models depends on instantaneous properties of the population, not on features of a whole trajectory[3]. This observation was illustrated in detail in section 4.2.3.

In this thesis, we also emphasized the dichotomy between a constant population size and a constant adaptation speed[4]. We used a time-independent constraint by transforming the full PDE for the constraint (3.18) to an ODE in the comoving frame. However, any time-dependent constraint that solves (3.18) would be possible. The only condition is that it has to be a solution that closes the moment hierarchy (at the first level). For instance, one could impose a trajectory that mimics the large-scale population-size changes mentioned above: either exponential or linear growth in the number of (reproducing) cells. The time-evolution of the full non-linear equation, or at least its relaxation to the steady state, was also analyzed in [Fis13].

7.2.2. Tuned models require a square-root dependence on population size for genetic drift

At the beginning of chapter 3 we presented the time-evolution of unconstrained branching random walks (BRW). The ensuing dynamics of (3.1) did not have any limitation to its overall population size growth. *Every* such limitation – or constraint – will introduce non-linearities in the dynamics. We chose the general form of non-linearities such that we can describe the dynamics by the fraction displayed in (3.15): the dynamics is divided by the forward propagation of a (possibly time-dependent) weighted integral over the density. Then we proceeded by expanding the fraction into order of $\mathcal{O}\left(\sqrt{\epsilon}\right)$. Thus, when averaging over the noise to arrive at the moments of the population density, *i.e.* $\overline{\mathbf{c}_t}$, $\overline{\mathbf{c}_t\mathbf{c}_t}$, $\overline{\mathbf{c}_t\mathbf{c}_t\mathbf{c}_t}$, *etc.*, the dynamics couples to other moments. However, tuned models construct the non-linearity essentially in such a way that the coupling to other moments averages out.

A crucial feature of the original BRW dynamics (3.1) is that the amplitude of the noise term is proportional to the square root of the density $\sqrt{\mathbf{c}_t}$. Without this particular form of the noise, the construction of tuned models would fail. Any other exponent α in the amplitude would create a peculiar coupling to $\overline{\mathbf{c}_t^{2\alpha}}$ in the resulting dynamics of the mean population density. Thus, *only* for $\alpha = 1/2$ is the moment hierarchy is closeable.

---

[3] Note that our model is not completely memory-less. The decay of the autocorrelation function τ_{dec} gives an estimate of how long features in the population density persist.

[4] Although a more exact characterization would be that of a constant speed of the comoving frame with which fixation probabilities shift to higher fitness.

7.2.3. Tuned models are applicable to more general scenarios

As soon as the conditions on the noise amplitude is met (see above) one can construct tuned models. The derivation will be similar to the presented one in section 3.2. The actual deterministic dynamics does not influence the construction of equations. Any Louivillean $\mathcal{L}$ would be possible if the noise amplitude is given by $\sim \sqrt{\mathbf{c}_t}$.

An example for this fact is presented in chapter 6 (with its extended version in [GH13]). There, we used a slightly modified operator $\mathcal{L}$. In this modified context, we were even able to state exact solutions in *closed form*.

We expect tuned models to be a more accurate description at large population sizes [Hal11b]. Differences between the fixed population size model and the tuned model become emphasized for small population sizes. For such parameter values, the dynamics is depends sensitively on fluctuations. This behavior at small population sizes was observed in both models that were treated in this thesis.

In traveling wave models of adaptation, fluctuations in the nose of the wave are pronounced in tuned models, and therefore alter the dynamics considerably. However, population sizes where these effects become important correspond to parameter values that are different than those expected for microbial experiments. In the figures presented in this thesis we always placed ourselves in parameter regimes characteristic of experimental microbial populations, in which the dynamical behavior of the constraint did not dominate the dynamics.

In the oasis model, we found a minimum in population size when decreasing the death rate, *cf.* (6.18). This increasing population size for tiny death rates is a feature of tuned models, where the population was forced to stay at the oasis, although fluctuations would have delocalized it already.

Although these peculiarities at small population sizes exist, they disappear at large population sizes. Overall, tuned models still have a decisive advantage: they create exact tractability.

7.3. Fixation probabilities help to discern effects of clonal competition

The process of fixation of a novel mutation crucially depends on the highly stochastic dynamics prevalent when the mutation is still rare. At first, every mutation is only present in a single individual. Genetic drift can bring even a very beneficial mutation to extinction when it is present in such small numbers. In this thesis, we devoted section 3.3 to establish the interpretation of the weighting function $\mathbf{u}_\star$ as fixation probability.

Thus, the fixation probability plays a decisive role in our tuned models, but our results just extend and generalize a long series of mathematical treatments.

7.3.1. Classical results

One of the first results was derived by Haldane [Hal27]: he used branching process methods to show that the fixation probability $u(X)$ of a mutation with fitness advantage X is linear in the fitness advantage for small X. The coefficient in the linear relation depends on the actual reproduction process of the organisms [JG02].

By symmetry, a single, completely neutral mutation ($X = 0$) in an asexually reproducing population of size N has fixation probability $1/N$.

Both results have been generalized in the *diffusion approximation* introduced by Kimura [Kim62, Gil10]. Based on the expansion of a master equation, this approximation allows to derive quantities in large populations analytically, among them the fixation probability of a mutation with arbitrary selection coefficient X. Assuming that the allele is already present in n individuals in a population of size N, its fixation probability u is given by

$$u(X) = \frac{1 - e^{-nX}}{1 - e^{-NX}} \,. \tag{7.2}$$

For a novel mutation ($n = 1$) and for an intermediate selection coefficient ($1/N \ll X \ll 1$) we arrive at the previous result by Haldane, $u(X) \sim X$. Neutral mutations have $u(0) = \frac{n}{N}$. For deleterious mutations ($X < 0$), the fixation probability decays exponentially.

7.3.2. Clonal competition reduces the fixation probability in intermediate fitness regimes

The results of the previous section apply to the dynamical regime of *periodic selection*: mutations emerge one at a time in the population and then either sweep to fixation or go extinct long before the next mutation occurs. In clonal competition scenarios, the dynamics is more intricate. Fixation does not only depend on the selection coefficient of a single mutation, but also on its genetic background and on the general distribution of (relative) fitness within the population.

In this thesis, we derived the fixation probability in the clonal competition scenario. Tuned models are exceptionally valuable in this respect, as the fixation probability $\mathbf{u}_\star$ is constant. There, the occupancy of all fitness classes is rescaled in order to keep the fixation probability constant. This is in fact the *defining feature* of our tuned models (as established in section 3.4). Moreover, we also presented derivations for the fixation probability in unconstrained branching random walks that follow a similar dynamical equation, (3.43) and (3.51b). In tuned models, however, we found a factor 2 as the coefficient of the non-linear term (*cf.* (3.18)), that was not present in the calculations of unconstrained branching random walks. We attributed this difference to the increased stochasticity in tuned models, that arises due to the need of "counter-events" to keep the population size from reaching either infinity or zero (see section 3.3.2). In generic traveling wave models (including those with fixed population size) the fixation probability is not as easy to obtain as in tuned models (or in unconstrained BRWs), because the fixation process depends crucially on *all* occupancies of different fitness classes.

In section 3.2.1 we introduced three different regimes in the shape of the fixation probability $\mathbf{u}_\star$. From low to high fitness we labelled them regime (I), (II) and (III), as depicted in Figure 3.1. In the rightmost regime (III), the profile of fixation probabilities is linear, $u_\star^{(III)}(X) \approx X/2$. The classical results above correspond to this regime: if an individual has such an extraordinarily large fitness, it will fixate in the population as long as it overcomes genetic drift without having to compete with others of similar fitness for fixation (as such individuals are almost singular occurrences). Similarly, regime (I) also corresponds to the classical exponential decay in (7.2) for small and negative fitnesses.

Contrary to the large fitness and negative fitness limits in the paragraph above, regime (II) is a *distinct feature* of clonal competition models. This regime extends from roughly $X \approx 0$ up to just below X_c, above which the solution converges to its asymptotic limit, $X/2$. In this range, the fixation probability is considerably reduced compared to the classical result. From the fitness of individuals alone, one would expect a much larger fixation probability. Many individuals with similar fitness exist in this regime. We also cannot distinguish different cohorts of mutations: if they have the same fitness, they are identical in traveling waves models. Hence, regime (II) in the fixation probability is one of the few instances where we can *directly pinpoint* the effects of clonal competition, even though we cannot track the fates of individuals (or individual clones).

For tiny adaptation speeds $v \to 0$, regime (II) starts to vanish from the profiles $\mathbf{u}_\star$. The characteristic feature of the sharp shoulder at X_c becomes more and more insignificant with decreasing v. Then, the fixation probability exhibits a profile that looks similar in shape to (7.2). Only on logarithmic scales are deviations from (7.2) discernible. The tail for $X \lesssim 0$ depends significantly on the mutation kernel, whereas the population size is the single parameter for decay in (7.2). A comparison of tails in different mutation kernels is depicted in Figure 3.1, although Kimura's result (7.2) is not shown there.

So far, we presented the large fitness limit of $\mathbf{u}_\star$ as proportional to fitness itself, $u_\star^{(III)}(X) \approx X/2$. Clearly, the interpretation of fixation probability *cannot* hold for values of X close to or above 2. This unnatural limit for a probability arises from the fact that we neglected the dependence on fitness in the variance of the genetic drift term. In fact, the noise would have an amplitude $\sim \sqrt{(2+X)\mathbf{c}_t}$ in (3.1) and (3.15). In section 3.3.2 we first kept this term in the derivation (3.42) of the generating function, but dropped the dependence in the subsequent presentation of solutions. For the range of fitness, where the main mass of the population density $\overline{\mathbf{c}_\star}$ is located, this additional term in the noise amplitude. However, when we include the correct noise term in the numerical solution, we can safely neglect the fixation probability saturates asymptotically at 1 (data not shown). The first-order Kramers-Moyal expansion we employed in section 3.2.1 can also be generalized to the "correct" noise amplitude: the governing equation is then $V\partial_X \mathbf{u}_\star \approx X\mathbf{u}_\star - (2+X)\mathbf{u}_\star^2$, which has the solution

$$u_\star(X) = \frac{\exp\bigl(X^2/(2V)\bigr)}{\exp\bigl(X^2/(2V)\bigr) + N + 2\sqrt{\pi/(2V)}\mathrm{Erfi}\bigl(X/\sqrt{2V}\bigr)} \xrightarrow{X\to\infty} 1\,. \tag{7.3}$$

In (7.3), the first term in the denominator ensures the correct convergence properties and therefore the interpretation as fixation probability. However, for all practical purposes this correct convergence is not needed: any reasonable realistic parameter set in traveling wave models exhibits $X_c \ll 1$. Individuals with fitness larger than X_c are exceptionally rare.

7.3.3. Emergent neutrality as different view on similar phenomena

A different perspective on the effects seen in clonal competition scenarios has been considered already for a long time: *genetic linkage* denotes the effect that mutations that are close to each other on the DNA sequence are inherited together by offspring. Often, this distance has been measured in terms of how likely different alleles recombine during meiosis. Increasing the genomic distance also increases also the probability of recombination. The effect of genetic linkage on observable traits has first been reported in the early 20th century [BSP06].

Often, theoretical studies investigated and quantified genetic linkage for neutral or deleterious alleles near a beneficial mutation, such that the former *hitch-hikes* with the latter [MSH74, Bar98, BEKV13]. Similar concepts were examined as *genetic draft*: evolution exhibits another type of stochasticity depending on which genetic background a novel mutation occurs [Gil10].

Recently, this effect of genetic linkage was analyzed in the context of adaptive evolution [SSML11]. There, the authors examine the effect of beneficial mutations under the influence of highly beneficial driver mutations. While in [SSML11] the focus is on the fate of single mutations for a parameter regime $N\mu < 1$, we aimed at describing the ensuing dynamical behavior for $N\mu > 1$. Nevertheless, results are comparable. In their analysis they also compute a threshold similar to our X_c. Mutations with a fitness effect below this threshold most likely only hitch-hike with other beneficial mutations, that have a larger fitness effect. The already small effect of the former mutations appears even smaller, as their fixation or extinction depends crucially on the fitness of the genetic background they have. Effectively, such mutations seem almost neutral, hence the authors coined the term *emergent neutrality*.

7.3.4. Deleterious mutations

In this thesis we neglected a detailed investigation of deleterious mutations. This is partially based on the experimental observation that in microbial populations multiple beneficial mutations are present, so that some of them will sweep to fixation. Even more so, it can be argued that in the rich environments most of these microbes experience, loss of function of many “unneeded” parts of the metabolism can be seen as neutral mutations, if not beneficial[5]. Hence, a larger fraction of all possible mutations might actually be beneficial for these populations, compared to organisms not living in such exuberant nutrients.

---

[5] Even when using minimal media as in many experiments, these environments are at least sufficient for growth and stay that way over long times.

Still, deleterious mutations will occur in these microbial populations. Their impact on the overall dynamics of adaptation is under debate. If the mutation rate for deleterious mutations is not too large, one could focus on the lineage that does not receive any deleterious mutation (or only few mildly deleterious ones) to track the adaptation process.

Recently, a study investigated the ratio of beneficial mutations needed to halt Muller's ratchet[6] [GBJ+12]. The results showed, that even a minute fraction (a few percent) of all mutations being beneficial suffices the lead to a stable mean fitness of the population.

In the context of results presented in this thesis, there are basically two approaches to gauge the impact of deleterious mutations. First, we can investigate the forward dynamics and analyze the impact of deleterious mutations on the population density. Second, we could treat the backward dynamics and check the changes in the fixation probability with the inclusion of deleterious mutations[7]. In a crude first estimate, both approaches will indicate that few, mildly deleterious mutations will not have a considerable influence on the overall dynamics [DF07].

The first approach is using the forward dynamics of branching random walks. Recall that "Fisher's fundamental theorem" (3.3) arises from an integrated version of the dynamics (3.1) in traveling wave models[8]:

$$v \approx \mathrm{Var}\big[x|\mathbf{c}_t\big] + \big(\mu\sigma\big)\ . \tag{7.4}$$

Deleterious mutations have two opposite effects on adaptation speed: first, they obviously decrease the term for the mutational load $\big(\mu\sigma\big)$. Here, deleterious mutations decrease the mean mutational effect $\sigma = \int ds\ sm(s)$ or render it negative, if a surplus of deleterious mutations exists. On the other hand, the spread of the population in fitness space is increased: deleterious mutations broaden the density at smaller fitness. The larger variance in fitness allows very fit sub-populations to grow faster, as they have larger difference to the mean fitness in the population. This increased growth rate might offset the deceleration by the mutational load. Adaptation speed is unaffected by (mildly) deleterious mutations in a first approximation.

Moreover, we can investigate the effect of deleterious mutations on the profile of fixation probabilities $\mathbf{u}_\star$. Note that even when we include deleterious mutations in the mutation kernel $m(s)$, we can employ the first order Kramers-Moyal expansion (3.23). The shape of the solution in regime (III) and regime (II) will still be the same. With the broader variance, see (7.4) above, the Gaussian increase in regime (II) might be slower. The crossover X_c might also be shifted to larger fitness values. Still, those changes are most likely only minor. We expect the biggest changes to the fixation probability $\mathbf{u}_\star$ in regime (I): there, the profile will decrease faster for $X \leq 0$ when including deleterious mutations. Indeed, the behavior can also be seen in the diffusion mutation kernel, where deleterious mutations are included. In Figure 3.1 all three mutation kernels are compared for the same population size and mutational scale σ. The profile in the diffusion mutation kernel is the steepest.

---

[6] Recall that Muller's ratchet denotes the dynamical scenario where the highest fitness class in a population is continually lost due to genetic drift and beneficial mutations are too rare to increase fitness again. See also Figure 1.3.

[7] Note added in proof: [GD14] treats exactly this case.

[8] (7.4) is obtained by first multiplying the dynamics with x/N_t, then integrating over fitness x.

Finally, we can use the results of our oasis model in chapter 6 to interpret the addition of a negative tail to the mutation kernel. The main focus in the oasis model was to investigate the conditions for delocalization of the population that was "pinned" at a point-like nutrient source. There, we included a term for convection velocity that drives the population away from the nutrients. Previously, it was found that he population delocalizes when convection is increased beyond a threshold [DNS00]. In our analysis we concluded that another threshold exists: the population size has to be large enough, otherwise frequent stochastic delocalizations occur. Our results on this model were summarized in the phase diagram of Figure 6.1. Increase the fraction of deleterious mutations in the mutation kernel can be seen in a similar context: when deleterious mutations (equivalent to convection in the oasis model) become too frequent, the variance in the population will also increase. If the population is too extended, it might stochastically delocalize if the single occupancies of fitness classes are too small. Therefore, we expect a maximal fraction of deleterious mutations until when the population stays localized. However, this fraction depends (most likely) on population size, in accordance with the results of [GBJ+12].

7.4. Fluctuations in the adaptation process

This thesis treated fluctuations in traveling wave models with various approaches: First, an extension to tuned models allowed for the construction of all correlation functions $\mathbf{C}_t^{(n)}$. The derivation was presented in section 3.6. Secondly, we used extensive stochastic simulations to investigate details of specific observables. Chapter 4 dealt with fluctuations in population size. Chapter 5 treated oscillations in the adaptation speed. There, we used "Fisher's fundamental theorem" to show that the autocorrelation function of adaptation speed and autocorrelation function of the fitness variance are identical. In turn, we were investigating the time dynamics of the fitness variance, that is much more amenable to analytical (and numerical) methods than adaptation speed.

7.4.1. Fitness variance oscillates with well defined frequency

A simple argument can explain why we would see large scale fluctuations of fitness variance in experiments: in the clonal competition regime a cohort of extraordinarily beneficial mutations will (most likely) arise. If it had survived genetic drift, it will sweep through the population, erasing much of the diversity. During the sweep the mean fitness increases substantially: the sweeping clone was fitter than the rest of the population. This fast increase in mean fitness corresponds to a large adaptation speed. After the sweep diversity will be low. Mutations will enter the population, and build up diversity again. During this time of build-up, adaptation will proceed at a slower pace. Different cohorts of mutations will be created; each at first accounting only for a tiny increase in mean fitness until one of them receives again a "final" mutation such that again a sweep is triggered. Overall, such a scenario will create one oscillation period in adaptation speed.

In stochastic traveling wave models with finite size, fitness variance will fluctuate. This can be deduced directly with mathematical arguments: in a model where we have a constraint of the form $\int d\chi(x)\; c_t(x) = \text{const}$ (with $d\chi(x)$ some measure on fitness space[9]) all fluctuations in vertical direction, *i.e.* the fluctuating occupancies due to genetic drift, will be transformed to fluctuations in horizontal direction, *i.e.* in the width of the density $\mathbf{c}_t$. Thus, fluctuations in fitness variance are *not* surprising. In contrast, the regularity of these fluctuations is surprising, as we showed in chapter 5. When analyzing time-traces of simulations that are long enough we observed oscillations that can be described to a very good extent with a *single* frequency.

This behavior was not observed in previous analyses, when only the fitness class at the nose of the wave was modeled stochastically. Thus, a fully stochastic model is needed to investigate this behavior. We proposed tuned models as convenient starting point for these more elaborate analyses [Hal11b, GRB+12, NH13, GH13]. A different approach was explained in [Fis13], where also a fully stochastic model was treated. There, the use of the timescale τ_{sw} was emphasized. The time τ_{sw} denotes the "sweep time", *i.e.* the time it takes for a sub-population at the nose (with fitness X_c above the mean) to become the largest sub-population, hence $\tau_{sw} \approx X_c/v$.

In tuned models, fluctuations in the nose of the wave are directly amplified by the specific choice of constraint [Fis11]. Thus, fluctuations in population size are not delayed by the time τ_{sw} as in [Fis13], but appear much faster.

We used autocorrelation functions with the definition $A_T \sim \overline{\delta a_t\, \delta a_{t+T}}$ to discern timescales in simulations. As seen in "Fisher's fundamental theorem", the mean of the adaptation speed and the mean of the fitness variance only differ by an additive constant. Thus, relevant timescales in both quantities are identical. The first timescale was the decay of the autocorrelation function, τ_{dec}. A second timescale was τ_{osc}, the time for one oscillation in fitness variance. One of these oscillations depends highly on the (stochastic) realization, but when averaged, a *single* timescale suffices to describe this process. Preliminary results suggest that τ_{osc} is different from the sweep time τ_{sw} above.

One particularly important observation has been summarized in Figure 5.4: the oscillation time τ_{osc} depends on the adaptation speed as

$$\tau_{osc} \approx c\big(v\sigma\big)^{-1/3} \,. \tag{7.5}$$

The coefficient c of this scaling depends on the mutation kernel. We obtained $c^{\text{exp}} \approx 2.8$ and $c^{\text{staircase}} \approx 4.5$. Note that $v\sigma$ has the dimension of a fitness flux [ML10], a crucial quantity in adaptive evolution when also environmental changes are included in the modeling. Conversely, one could also use the dimensionless scaling $\tau_{osc}\sigma \sim \big(\frac{v}{\sigma^2}\big)^{-1/3}$ to state (7.5). Overall, the scaling (7.5) appears almost universal for all three mutation kernels. All our simulation results follow it over roughly eight orders of magnitude by varying the parameters (*cf.* Figure 5.4).

---

[9] In this thesis we treated the two choices $d\chi(x) = du_\star(x)$ and $d\chi(x) = \frac{dx}{N}$.

Measurements of characteristic fitness scales in tuned models have not yet been completed. These measurements of characteristic scales can be compared to the fitness increase during either the oscillation time $v\tau_{osc}$ or the fitness increase during the decay time $v\tau_{dec}$. In appendix D we presented our current (noisy) results on the comparison. Trends are visible, but no matching fitness scale can be distinguished. Nevertheless, we believe that $\mathbf{g}_\star$, the distribution of fixating clones, plays a distinct role in the dynamics of these oscillations, as visual inspection of our simulations shows.

7.4.2. Higher moments in tuned models allow investigation of the complete fluctuation spectrum

Stochastic traveling wave models suffer from the problem many non-linear model have: when stochasticity and non-linearity are combined, they create a hierarchy of moment equations that does not close upon averaging over the noise. Tuned models were constructed essentially in such a way that both effects *exactly* cancel each other. Even more so, as we have shown in section 3.6, the approach can be extended to higher moments.

Originally, tuned models separated the non-linear problem into two equations: a non-linear one for the fixation probability $\mathbf{u}_\star$ and a linear equation for the mean population density $\overline{\mathbf{c}_t}$ [Hal11b]. In the generalization to higher moments one first chooses a maximal n for which the correlation function

$$\overline{\mathbf{C}_t^{(n)}} \equiv \overline{C_t^{(n)}(x_1,\dots,x_n)} = \overline{c_t(x_1)\dots c_t(x_n)} \tag{7.6}$$

should be obtained. Then a closed set of $n+1$ equations for all correlations functions $\overline{\mathbf{C}_t^{(m)}}$ with $m \leq n$ and an additional equation for $\mathbf{u}_\star^{(n)}$ is constructed. Higher moments allow to investigate the fluctuation behavior of those traveling wave models directly. In principle, as we have access to all moments of the population density, the complete fluctuation spectrum can be computed.

Even the localization of the fluctuations can be analyzed: for instance, the fitness dependent covariance of fluctuations is given by

$$\mathrm{Cov}\big[\mathbf{c}_t,\mathbf{c}_t\big] \equiv \overline{C_t^{(2)}(x,y)} - \overline{c_t(x)}\,\overline{c_t(y)}\ . \tag{7.7}$$

From the covariance, the variance is easily obtained by equating the two fitness variables, $x=y$, in (7.7). The coefficient of variation, defined as $\mathrm{Var}\big[\mathbf{c}_t\big]/\overline{\mathbf{c}_t}$, exhibits a maximum just below X_c. There, at the stochastic edge, fluctuations are most pronounced: fitness classes around X_c have the biggest influence on the later dynamics. A large fluctuation at such fitness values will speed up the wave: the elevated occupancy in such a fitness class grows very fast due to selection, and mean fitness increases substantially.

Closing the moment hierarchy at a higher level can also be interpreted as increased stochasticity, as we explained in section 3.6: the original tuned model for $n=1$ had a decreased fixation probability by a factor of 2 compared to unconstrained branching random walks. The derivation presented in appendix C.2 allows the interpretation of

$\mathbf{u}_\star^{(n)}$ as the fixation probability in extended tuned models, where the moment hierarchy is closed at the n^{th} level. Thus, these extended tuned models have a decreased fixation probability by a factor of $n+1$ compared to unconstrained branching random walks. This larger stochasticity was observed in simulations, where the behavior appears much more erratic. Thus, although the higher moments can be obtained directly, the dynamical behavior might get dominated by artifacts for the limit $n \to \infty$. Still, for small n the approach will yield reasonable predictions, and is (to our knowledge) the only approach that can predict them with *exact* equations.

7.4.3. Punctuated equilibrium as macroevolutionary theory also predicts oscillating adaptation speed

In the 1970s, Eldredge and Gould proposed the theory of *punctuated equilibrium* [EG72, Gou02]. The absence of intermediate fossils in the evolutionary history of many species lead them to the conclusion that on a macroevolutionary level, evolution proceeds in discrete steps. On a geological timescale, observed fossils of a single species often show only minor changes, while then suddenly large changes occur [Gou02]. Their view challenged the (at that time) prevailing evolutionary gradualism, that considered accumulation of tiny evolutionary steps as main mode of adaptation.

Microbial evolution experiments can also be investigated in the light of punctuated equilibrium. For a single lineage, adaptation appears almost gradual, as beneficial mutations occur repeatedly. Still, on the larger population level, the situation might look different. In the clonal competition regime, multiple lineages can each accumulate their own cohort of mutations. However, when a "final" beneficial mutation leads to a sweep through the population, adaptation is fast. Overall, the dynamics in these clonal competition scenarios in microbes creates a pattern that appears similar to punctuated equilibrium: a period of stasis is followed by rapid adaptation, at least on the population scale. Oscillations in adaptation speed might be a general mode of evolution.

Clearly, there are differences in the theory of punctuated equilibrium and microbial evolution experiments. In the fossil record only a part of the morphological structure of individuals is preserved: often only fossilized bones (or equivalent stabilizing structures for other organisms) are found. These offer unfortunately only a small window into the evolutionary changes a species experiences. In lab experiments, with a "frozen fossil record" waiting in the next fridge to be thawed, a much more detailed evolutionary history of a lineage can be reconstructed. Therefore, one would expect to detect many more small changes, that were ultimately lost in the history of large organisms. Moreover, the timescales involved in the evolution of whole species reach a duration where climatic variation and/or migration to other geographic locations induce large scale changes in the environments experienced by species. Abrupt changes in environment might also trigger rapid adaptations based on already previously present (neutral) traits. Laboratory conditions are virtually unchanged. Despite these differences, adaptation speed does not appear to be constant for many evolutionary processes.

Genomic evidence suggests that the successful lineage in microbial populations exhibits a different mode of evolution, than previously proposed in the *neutral theory of molecular evolution* by Kimura [Kim84]. This theory applies to "large" organisms. It states that the accumulation of neutral mutations proceeds at a steady pace. Additionally, for such organisms beneficial mutations are rare, resulting in a much larger variance for their substitution rate within the genome. However, for microbes in laboratory populations in the clonal competition regime, this situation seems reversed: multiple beneficial mutations are present at all times, competing for fixation. This leads to a low variance in substitution rates for beneficial mutations. The successful lineage has an almost constant rate of accumulation of beneficial mutations (as it is conditioned on being successful). Neutral (or even deleterious) mutations just hitchhike on fixating lineages and have therefore a larger variance in substitution rates. Note that evidence for this fact is not (yet) conclusive, but first studies point into this direction [BYY+09], where many observed substitutions are supposed to be beneficial.

7.5. Experimental evolution revisited in the light of theoretical results

At the end, one may question the value of the elaborate mathematical treatment we presented on the last pages, now in the light of experimental observations. While our model was constructed to be descriptive at the beginning, a few cases instances exist where it might be predictive. Both cases we present below rely on our observation of the scaling relation (7.5).

7.5.1. Oscillations in adaptation speed allow to infer the mutational scale

The scaling (7.5) can be used to infer a mutational scale in the *long-time evolution experiment* with *E. coli* [Len15]. As a result, we obtained

$$\sigma^{(10000)} \approx 10^{-3}\,, \tag{7.8a}$$
$$\sigma^{(50000)} \approx 10^{-4}\,. \tag{7.8b}$$

These two values for σ are considerably lower than the values published in [BYY+09] for observed mutations: in isogenic strains, *i.e.* only single mutations are introduced in the ancestor's genetic background, fitness measurements yielded values as large as $\sigma \approx \mathcal{O}\big(10^{-1}\big)$ or $\sigma \approx \mathcal{O}\big(10^{-2}\big)$. Even if these values appear relatively large compared to our prediction, these *observed* mutations are conditioned on being observable. Thus, they reached appreciable frequencies in the population and most likely rose to numbers above the threshold such that selection can act on them. On the other hand, our result predicts not the value of fixating mutations, but of the underlying mutational scale. This underlying scale is obscured by the dynamics of adaptation in clonal competition scenarios. Many of the mutations with intermediate fitness effect never arise in a "suitable" genetic background such that they can grow to reach observable frequencies.

The values of the underlying mutational scale, (7.8a) and (7.8b), can be used in the distribution of fixating mutations $\boldsymbol{\rho}_f$ (*cf.* (3.62)). The mean value of this distribution would correspond to the observed mutational scale. Inserting appropriate values in the exponential mutation kernel, a difference of roughly one order of magnitude between the underlying and the observed mutational scale arises. Therefore, the fact that our estimate for σ above is much smaller is not irritating, but reassuring considering the dynamics of adaptation.

7.5.2. Oscillations in adaptation speed allow to infer a reasonable duration of experiments

Moreover, our scaling (7.5) can act as rule of thumb when planning experiments. The coefficient in the scaling $\tau_{osc} \sim v^{-1/3}$ is roughly 50 when neglecting the weak dependence on the mutational scale. After approximately τ_{osc} generations, one would expect the effects of clonal competition to be a major constituent of the dynamics. Starting from a monoclonal population, *i.e.* a single cell, it would build up diversity during a transient time of roughly that length. For standard experiments we thus expect oscillation effects visible after several hundreds or a few thousands of generations. Many experiments with the current "default" length of 1000 generations seem almost too short to observe and measure these oscillations.

However, there is a peculiarity that is yet unresolved in our model. Adaptation seems to be fastest at the beginning, and then decelerates: diminishing returns of novel mutations entering the population do not increase fitness as fast later in the experiment compared to the beginning (see Figures 1.1 and 5.8). Therefore, measuring adaptation speed at the beginning to determine "interesting" timescales with the scaling above might underestimate the actual time needed for effects to be visible.

7.6. Concluding remarks

This closes the analyses presented in this thesis. We investigated traveling wave models based on constrained branching random walks in order to explain the dynamical behavior found in experimental microbial populations. Our minimal model captured the adaptation process that is observed in those experiments: we only included mutations, selection and genetic drift in the model, and ignored any complex interactions between mutations that appear in real populations.

The focus in this thesis was clearly on tuned models, that allowed to treat a fully stochastic version of the governing equations. In exemplaric cases, we could discern the impact of fluctuations on the dynamics of adaptation. These were for instance the effect of the fluctuating population size in section 4.2.3, or the timescales involved in the large scale oscillations in diversity (as laid out in detail in chapter 5).

Still, some recurrent observations in experiments are not yet implemented the model and leave room for future investigations.

In the big picture, however, the applicability of these models to adapting populations in nature is quite limited: simply by looking out of the window, one can assure oneself that evolution produced much more than just ever faster multiplying asexual single celled organisms. On the other hand, microbes are still present outside, proving that their evolutionary strategy is perfectly viable. □

Bibliography

[All03] Allen LJ. *An introduction to stochastic processes with applications to biology.* Pearson Education New Jersey. 2003

[BAD+07] Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler KW, Velculescu VE, Vogelstein B and Nowak MA. Genetic progression and the waiting time to cancer. *PLoS computational biology*, **3**(11):p. e225. 2007

[BAO+10] Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S, Karchin R, Kinzler KW, Vogelstein B and Nowak MA. Accumulation of driver and passenger mutations during tumor progression. *Proceedings of the National Academy of Sciences*, **107**(43):pp. 18545–18550. 2010

[Bar98] Barton NH. The effect of hitch-hiking on neutral genealogies. *Genetical Research*, **72**(02):pp. 123–133. 1998

[BBDL12] Blount ZD, Barrick JE, Davidson CJ and Lenski RE. Genomic analysis of a key innovation in an experimental Escherichia coli population. *Nature*, **489**(7417):pp. 513–518. 2012

[BBL08] Blount ZD, Borland CZ and Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences*, **105**(23):pp. 7899–7906. 2008

[BDFN92] Binney JJ, Dowrick N, Fisher A and Newman M. *The theory of critical phenomena: an introduction to the renormalization group.* Oxford University Press, Inc. 1992

[BEKV13] Barton N, Etheridge A, Kelleher J and Veber A. Genetic hitchhiking in spatially extended populations. *Theoretical population biology*, **87**:pp. 75–89. 2013

[BMBC09] Buckling A, Maclean RC, Brockhurst MA and Colegrave N. The Beagle in a bottle. *Nature*, **457**(7231):pp. 824–829. 2009

[BRW08] Brunet É, Rouzine IM and Wilke CO. The stochastic edge in adaptive evolution. *Genetics*, **179**(1):pp. 603–620. 2008

[BSP06] Bateson W, Saunders ER and Punnett RC. Further experiments on inheritance in sweet peas and stocks: preliminary account. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, **77**(517):pp. 236–238. 1906

[BST14] Bauer B, Siebert R and Traulsen A. Cancer initiation with epistatic interactions between driver and passenger mutations. *Journal of theoretical biology*, **358**:pp. 52–60. 2014

[Bur12] Burke MK. How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proceedings of the Royal Society B: Biological Sciences*, **279**(1749):pp. 5029–5038. 2012

[BYY$^+$09] Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE and Kim JF. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, **461**(7268):pp. 1243–1247. 2009

[CCD$^+$11] Chou HH, Chiu HC, Delaney NF, Segrè D and Marx CJ. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, **332**(6034):pp. 1190–1192. 2011

[CH66] Courant R and Hilbert D. *Methods of mathematical physics*, volume 1. CUP Archive. 1966

[CK70] Crow JF and Kimura M. *An introduction to population genetics theory.* New York, Evanston and London: Harper & Row, Publishers. 1970

[CKL05] Cohen E, Kessler DA and Levine H. Front propagation up a reaction rate gradient. *Physical Review E*, **72**(6):p. 066126. 2005

[Dal87] Dallinger WH. The president's address. *Journal of the Royal Microscopical Society*, **7**(2):pp. 185–199. 1887

[Dar59] Darwin C. *The Origin of Species by Mean of Natural Selection.* 1859

[Des13a] Desai MM. Personal Communication. 2013

[Des13b] Desai MM. Statistical questions in experimental evolution. *Journal of Statistical Mechanics: Theory and Experiment*, **2013**(01):p. P01003. 2013

[DF07] Desai MM and Fisher DS. Beneficial mutation–selection balance and the effect of linkage on positive selection. *Genetics*, **176**(3):pp. 1759–1798. 2007

[DH99] Drake JW and Holland JJ. Mutation rates among RNA viruses. *Proceedings of the National Academy of Sciences*, **96**(24):pp. 13910–13913. 1999

[DNS00] Dahmen K, Nelson D and Shnerb N. Life and death near a windy oasis. *Journal of mathematical biology*, **41**(1):pp. 1–23. 2000

[DSL+09] Daegelen P, Studier FW, Lenski RE, Cure S and Kim JF. Tracing Ancestors and Relatives of Escherichia coli B, and the Derivation of B Strains REL606 and BL21 (DE3). *Journal of molecular biology*, **394**(4):pp. 634–643. 2009

[DWF13] Desai MM, Walczak AM and Fisher DS. Genetic diversity and the structure of genealogies in rapidly adapting populations. *Genetics*, **193**(2):pp. 565–585. 2013

[EG72] Eldredge N and Gould SJ. Punctuated equilibria: an alternative to phyletic gradualism. *Models in paleobiology*, **82**:p. 115. 1972

[Eig71] Eigen M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, **58**(10):pp. 465–523. 1971

[EL03] Elena SF and Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, **4**(6):pp. 457–469. 2003

[EMS89] Eigen M, McCaskill J and Schuster P. The Molecular Quasi-Species. *Advances in chemical physics*, pp. 149–263. 1989

[ENC04] ENCODE Project Consortium. The ENCODE (ENCyclopedia of DNA elements) project. *Science*, **306**(5696):pp. 636–640. 2004

[EWK07] Eyre-Walker A and Keightley PD. The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, **8**(8):pp. 610–618. 2007

[Fel74] Felsenstein J. The evolutionary advantage of recombination. *Genetics*, **78**(2):pp. 737–756. 1974

[FH15] Flot JF and Hallatschek O. 2015

[Fis30] Fisher RA. *The genetical theory of natural selection*. Clarendon. 1930

[Fis37] Fisher RA. The wave of advance of advantageous genes. *Annals of Human Genetics*, **7**(4):pp. 355–369. 1937

[Fis11] Fisher DS. Leading the dog of selection by its mutational nose. *Proceedings of the National Academy of Sciences*, **108**(7):pp. 2633–2634. 2011

[Fis13] Fisher DS. Asexual evolution waves: fluctuations and universality. *Journal of Statistical Mechanics: Theory and Experiment*, **2013**(01):p. P01011. 2013

[FND08] Fogle CA, Nagle JL and Desai MM. Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics*, **180**(4):pp. 2163–2173. 2008

[Gar09] Gardiner CW. *Handbook of stochastic methods*. Springer, Fourth edition. 2009

[GBJ+12] Goyal S, Balick DJ, Jerison ER, Neher RA, Shraiman BI and Desai MM. Dynamic mutation–selection balance as an evolutionary attractor. *Genetics*, **191**(4):pp. 1309–1319. 2012

[GD13] Good BH and Desai MM. Fluctuations in fitness distributions and the effects of weak linked selection on sequence evolution. *Theoretical population biology*, **85**:pp. 86–102. 2013

[GD14] Good BH and Desai MM. Deleterious passengers in adapting populations. *Genetics*, **198**(3):pp. 1183–1208. 2014

[GDL02] Gogarten JP, Doolittle WF and Lawrence JG. Prokaryotic evolution in light of gene transfer. *Molecular biology and evolution*, **19**(12):pp. 2226–2238. 2002

[GH13] Geyrhofer L and Hallatschek O. Stochastic delocalization of finite populations. *Journal of Statistical Mechanics: Theory and Experiment*, **2013**(01):p. P01007. 2013

[Gil10] Gillespie JH. *Population genetics: a concise guide*. JHU Press. 2010

[GL98] Gerrish PJ and Lenski RE. The fate of competing beneficial mutations in an asexual population. *Genetica*, **102**:pp. 127–144. 1998

[Gou02] Gould SJ. *The structure of evolutionary theory*. Harvard University Press. 2002

[GRB+12] Good BH, Rouzine IM, Balick DJ, Hallatschek O and Desai MM. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proceedings of the National Academy of Sciences*, **109**(13):pp. 4950–4955. 2012

[GSS+09] Germonpré M, Sablin MV, Stevens RE, Hedges RE, Hofreiter M, Stiller M and Després VR. Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine and Russia: osteometry, ancient DNA and stable isotopes. *Journal of Archaeological Science*, **36**(2):pp. 473–490. 2009

[GZP+13] Graur D, Zheng Y, Price N, Azevedo RB, Zufall RA and Elhaik E. On the immortality of television sets: "function" in the human genome according to the evolution-free gospel of ENCODE. *Genome biology and evolution*, **5**(3):pp. 578–590. 2013

[Hal27] HALDANE JBS. A mathematical theory of natural and artificial selection, part V: selection and mutation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 23, pp. 838–844. Cambridge University Press. 1927

[Hal11a] HALLATSCHEK O. The accumulation of beneficial mutations in asexual populations – beyond the dichotomy of "clonal interference" versus "multiple mutations". Unpublished notes. 2011

[Hal11b] HALLATSCHEK O. The noisy edge of traveling waves. *Proceedings of the National Academy of Sciences*, **108**(5):pp. 1783–1787. 2011

[Hal13] HALLATSCHEK O. Personal communication ;). 2013

[HD13] HERRON MD AND DOEBELI M. Parallel evolutionary dynamics of adaptive diversification in Escherichia coli. *PLoS biology*, **11**(2):p. e1001490. 2013

[Hen00] HENIG RM. *The monk in the garden: the lost and found genius of Gregor Mendel, the father of genetics.* Houghton Mifflin Harcourt. 2000

[HF14] HALLATSCHEK O AND FISHER DS. Acceleration of evolutionary spread by long-range dispersal. *Proceedings of the National Academy of Sciences*, **111**(46):pp. E4911–E4919. 2014

[HHM$^+$01] HILLMAN G, HEDGES R, MOORE A, COLLEDGE S AND PETTITT P. New evidence of Lateglacial cereal cultivation at Abu Hureyra on the Euphrates. *The Holocene*, **11**(4):pp. 383–393. 2001

[HHRN07] HALLATSCHEK O, HERSEN P, RAMANATHAN S AND NELSON DR. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences*, **104**(50):pp. 19926–19930. 2007

[HJV05] HACCOU P, JAGERS P AND VATUTIN VA. *Branching processes: variation, growth, and extinction of populations.* 5. Cambridge University Press. 2005

[HN08] HALLATSCHEK O AND NELSON DR. Gene surfing in expanding populations. *Theoretical population biology*, **73**(1):pp. 158–170. 2008

[HR66] HILL WG AND ROBERTSON A. The effect of linkage on limits to artificial selection. *Genetical research*, **8**(03):pp. 269–294. 1966

[HRWB02] HERMISSON J, REDNER O, WAGNER H AND BAAKE E. Mutation–selection balance: Ancestry, load, and maximum principle. *Theoretical population biology*, **62**(1):pp. 9–46. 2002

[Hux42] HUXLEY J. *Evolution. The Modern Synthesis.* London: George Alien & Unwin Ltd. 1942

[IS01] Imhof M and Schlötterer C. Fitness effects of advantageous mutations in evolving Escherichia coli populations. *Proceedings of the National Academy of Sciences*, **98**(3):pp. 1113–1117. 2001

[Jai08] Jain K. Loss of least-loaded class in asexual populations due to drift and epistasis. *Genetics*, **179**(4):pp. 2125–2134. 2008

[JG02] Johnson T and Gerrish PJ. The fixation probability of a beneficial allele in a population dividing by binary fission. *Genetica*, **115**(3):pp. 283–287. 2002

[Kam92] van Kampen NG. *Stochastic processes in physics and chemistry*, volume 1. Elsevier. 1992

[KB06] Kassen R and Bataillon T. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nature genetics*, **38**(4):pp. 484–488. 2006

[KDBP11] Kryazhimskiy S, Dushoff J, Bazykin GA and Plotkin JB. Prevalence of epistasis in the evolution of influenza A surface proteins. *PLoS genetics*, **7**(2):p. e1001301. 2011

[KDS$^+$11] Khan AI, Dinh DM, Schneider D, Lenski RE and Cooper TF. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, **332**(6034):pp. 1193–1196. 2011

[Kim62] Kimura M. On the probability of fixation of mutant genes in a population. *Genetics*, **47**(6):p. 713. 1962

[Kim84] Kimura M. *The neutral theory of molecular evolution.* Cambridge University Press. 1984

[KLE$^+$12] Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I and Whitlock MC. Experimental evolution. *Trends in ecology & evolution*, **27**(10):pp. 547–560. 2012

[KPP37] Kolmogorov A, Petrovsky I and Piscounoff N. Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem. *Bull. Univ. Moscow, Ser. Int. A*, **1**(1). 1937

[KRJD14] Kryazhimskiy S, Rice DP, Jerison E and Desai MM. Global Epistasis Makes Adaptation Predictable Despite Sequence-Level Stochasticity. *Science*, **344**(6191):pp. 1519–1522. 2014

[Kus13] Kussell E. Evolution in Microbes. *Annual review of biophysics*, **42**:pp. 493–514. 2013

[KW64] Kimura M and Weiss G. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**(4):p. 561. 1964

[LBBG93] Lynch M, Bürger R, Butcher D and Gabriel W. The mutational meltdown in asexual populations. *Journal of Heredity*, **84**(5):pp. 339–344. 1993

[LBD11] Lang GI, Botstein D and Desai MM. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics*, **188**(3):pp. 647–661. 2011

[LC13] Lenski RE and Crespi S. Podcast interview accompanying publication of [WRL13b]. 2013

[Len15] Lenski RE. The E. coli long-term experimental evolution project site. 2015
URL http://myxo.css.msu.edu/ecoli

[LGPH+12] Le Gac M, Plucain J, Hindré T, Lenski RE and Schneider D. Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with Escherichia coli. *Proceedings of the National Academy of Sciences*, **109**(24):pp. 9487–9492. 2012

[LMTO+04] Lin A, Mann B, Torres-Oviedo G, Lincoln B, Käs J and Swinney H. Localization and extinction of bacterial populations under inhomogeneous growth conditions. *Biophysical journal*, **87**(1):pp. 75–80. 2004

[LRH+13] Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D and Desai MM. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*, **500**(7464):pp. 571–574. 2013

[LRST91] Lenski RE, Rose MR, Simpson SC and Tadler SC. Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2, 000 generations. *American Naturalist*, **138**(6):pp. 1315–1341. 1991

[LVK10] López D, Vlamakis H and Kolter R. Biofilms. *Cold Spring Harbor perspectives in biology*, **2**(7):p. a000398. 2010

[LYGA00] Lev-Yadun S, Gopher A and Abbo S. The cradle of agriculture. *Science*, **288**(5471):pp. 1602–1603. 2000

[ME13] Metzger JJ and Eule S. Distribution of the Fittest Individuals and the Rate of Muller's Ratchet in a Model with Overlapping Generations. *PLoS Comput Biol*, **9**(11):p. e1003303. 2013
URL http://dx.doi.org/10.1371%2Fjournal.pcbi.1003303

[Men66] Mendel G. Versuche über Pflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brunn 4: 3*, **44**. 1866

[ML10] Mustonen V and Lässig M. Fitness flux and ubiquity of adaptive evolution. *Proceedings of the National Academy of Sciences*, **107**(9):pp. 4248–4253. 2010

[MNNM14] Müller MJ, Neugeboren BI, Nelson DR and Murray AW. Genetic drift opposes mutualism during spatial population expansion. *Proceedings of the National Academy of Sciences*, **111**(3):pp. 1037–1042. 2014

[MSH74] Maynard Smith J and Haigh J. The hitch-hiking effect of a favourable gene. *Genetical research*, **23**(01):pp. 23–35. 1974

[Mul32] Muller HJ. Some genetic aspects of sex. *American Naturalist*, pp. 118–138. 1932

[Mul64] Muller HJ. The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **1**(1):pp. 2–9. 1964

[Neh13] Neher RA. Genetic Draft, Selective Interference, and Population Genetics of Rapid Adaptation. *Annual Review of Ecology, Evolution, and Systematics*, **44**:pp. 195–215. 2013

[NH13] Neher RA and Hallatschek O. Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences*, **110**(2):pp. 437–442. 2013

[Now06] Nowak MA. *Evolutionary dynamics: exploring the equations of life*. Harvard University Press. 2006

[NPLK00] Neicu T, Pradhan A, Larochelle D and Kudrolli A. Extinction transition in bacterial colonies under forced convection. *Physical Review E*, **62**(1):p. 1059. 2000

[NS12] Neher RA and Shraiman BI. Fluctuations of fitness distributions and the rate of Muller's ratchet. *Genetics*, **191**(4):pp. 1283–1293. 2012

[NSF10] Neher RA, Shraiman BI and Fisher DS. Rate of adaptation in large sexual populations. *Genetics*, **184**(2):pp. 467–481. 2010

[OF94] O'Hagan A and Forster J. *Kendall's Advanced Theory of Statistics: Bayesian Inference*, volume 2B. New York Halsted Press. 1994

[OKI$^+$14] Ohtomo Y, Kakegawa T, Ishida A, Nagase T and Rosing MT. Evidence for biogenic graphite in early Archaean Isua metasedimentary rocks. *Nature Geoscience*, **7**(1):pp. 25–28. 2014

[Orr98] Orr HA. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*, pp. 935–949. 1998

[Orr03] Orr HA. The distribution of fitness effects among beneficial mutations. *Genetics*, **163**(4):pp. 1519–1526. 2003

[Orr06] Orr HA. The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation. *Journal of theoretical biology*, **238**(2):pp. 279–285. 2006

[Pen55] Penrose R. A generalized inverse for matrices. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 51, pp. 406–413. Cambridge University Press. 1955

[Pen56] Penrose R. On best approximate solutions of linear matrix equations. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 52, pp. 17–19. Cambridge University Press. 1956

[Pen13] Pennisi E. The Man Who Bottled Evolution. *Science*, **342**(6160):pp. 790–793. 2013

[PHLG$^+$14] Plucain J, Hindré T, Le Gac M, Tenaillon O, Cruveiller S, Médigue C, Leiby N, Harcombe WR, Marx CJ, Lenski RE and Schneider D. Epistasis and Allele Specificity in the Emergence of a Stable Polymorphism in Escherichia coli. *Science*, **343**(6177):pp. 1366–1369. 2014

[PK07] Park SC and Krug J. Clonal interference in large populations. *Proceedings of the National Academy of Sciences*, **104**(46):pp. 18135–18140. 2007

[Pri70] Price GR. Selection and covariance. *Nature*, **227**:pp. 520–21. 1970

[Pri72] Price GR. Fisher's 'fundamental theorem' made clear. *Annals of human genetics*, **36**(2):pp. 129–140. 1972

[PSK10] Park SC, Simon D and Krug J. The speed of evolution in large asexual populations. *Journal of Statistical Physics*, **138**(1-3):pp. 381–410. 2010

[PTVF07] Press WH, Teukolsky SA, Vetterling WT and Flannery BP. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press. 2007

[RBW08] Rouzine IM, Brunet É and Wilke CO. The traveling-wave approach to asexual evolution: Muller's ratchet and speed of adaptation. *Theoretical population biology*, **73**(1):pp. 24–46. 2008

[RCR+11] Raney BJ, Cline MS, Rosenbloom KR, Dreszer TR, Learned K, Barber GP, Meyer LR, Sloan CA, Malladi VS, Roskin KM, Suh BB, Hinrichs AS, Clawson H, Zweig AS, Kirkup V, Fujita PA, Rhead B, Smith KE, Pohl A, Kuhn RM, Karolchik D, Haussler D and Kent WJ. ENCODE whole-genome data in the UCSC genome browser (2011 update). *Nucleic acids research*, **39**(suppl 1):pp. D871–D875. 2011

[RDBT12] Ratcliff WC, Denison RF, Borrello M and Travisano M. Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences*, **109**(5):pp. 1595–1600. 2012

[RDVG02] Rozen DE, De Visser J and Gerrish PJ. Fitness effects of fixed beneficial mutations in microbial populations. *Current biology*, **12**(12):pp. 1040–1045. 2002

[RHH+13] Ratcliff WC, Herron MD, Howell K, Pentz JT, Rosenzweig F and Travisano M. Experimental evolution of an alternating uni-and multicellular life cycle in Chlamydomonas reinhardtii. *Nature communications*, **4**. 2013

[Rus93a] Russell H. Bacteriology in its general relation (continued). *American Naturalist*, pp. 1050–1065. 1893

[Rus93b] Russell H. Bacteriology in its general relations. *American Naturalist*, pp. 847–859. 1893

[RWC03] Rouzine IM, Wakeley J and Coffin JM. The solitary wave of asexual evolution. *Proceedings of the National Academy of Sciences*, **100**(2):pp. 587–592. 2003

[Saa03] van Saarloos W. Front propagation into unstable states. *Physics reports*, **386**(2):pp. 29–222. 2003

[SE12] Slatkin M and Excoffier L. Serial founder effects during range expansion: a spatial analog of genetic drift. *Genetics*, **191**(1):pp. 171–181. 2012

[SG10] Sniegowski PD and Gerrish PJ. Beneficial mutations and the dynamics of adaptation in asexual populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**(1544):pp. 1255–1263. 2010

[SGL97] Sniegowski PD, Gerrish PJ and Lenski RE. Evolution of high mutation rates in experimental populations of E. coli. *Nature*, **387**(6634):pp. 703–705. 1997

[SSF+13] Szendro IG, Schenk MF, Franke J, Krug J and de Visser JAG. Quantitative analyses of empirical fitness landscapes. *Journal of Statistical Mechanics: Theory and Experiment*, **2013**(01):p. P01005. 2013

[SSML11] Schiffels S, Szöllősi GJ, Mustonen V and Lässig M. Emergent neutrality in adaptive asexual evolution. *Genetics*, **189**(4):pp. 1361–1375. 2011

[TLK96] Tsimring LS, Levine H and Kessler DA. RNA virus evolution via a fitness-space model. *Physical review letters*, **76**(23):pp. 4440–4443. 1996

[TMM98] Tunnicliffe V, McArthur A and McHugh D. A biogeographical perspective of the deep-sea hydrothermal vent fauna. *Advances in Marine Biology*, **34**:pp. 353–442. 1998

[TRVG+12] Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD and Gaut BS. The molecular diversity of adaptive convergence. *Science*, **335**(6067):pp. 457–461. 2012

[Tun91] Tunnicliffe V. The biology of hydrothermal vents: ecology and evolution. *Oceanography and Marine Biology an Annual Review*, **29**:pp. 319–407. 1991

[VZG+99] de Visser A, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE et al. Diminishing returns from mutation supply rate in asexual populations. *Science*, **283**(5400):pp. 404–406. 1999

[Wäc90] Wächtershäuser G. Evolution of the first metabolic cycles. *Proceedings of the National Academy of Sciences*, **87**(1):p. 200. 1990

[Wat75] Watterson G. On the number of segregating sites in genetical models without recombination. *Theoretical population biology*, **7**(2):pp. 256–276. 1975

[WBT+13] Wielgoss S, Barrick JE, Tenaillon O, Wiser MJ, Dittmar WJ, Cruveiller S, Chane-Woon-Ming B, Médigue C, Lenski RE and Schneider D. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proceedings of the National Academy of Sciences*, **110**(1):pp. 222–227. 2013

[WDDH06] Weinreich DM, Delaney NF, DePristo MA and Hartl DL. Darwinian evolution can follow only very few mutational paths to fitter proteins. *science*, **312**(5770):pp. 111–114. 2006

[WGP03] Whitlock M, Griswold C and Peters A. Compensating for the meltdown: the critical effective size of a population with deleterious and compensatory mutations. In *Annales Zoologici Fennici*, volume 40, pp. 169–183. Helsinki: Suomen Biologian Seura Vanamo, 1964-. 2003

[Wil05] Wilke CO. Quasispecies theory in the context of population genetics. *BMC evolutionary biology*, **5**(1):p. 44. 2005

[Win08] Winitzki S. A handy approximation for the error function and its inverse. 2008

[WK65] Weiss G and Kimura M. A mathematical analysis of the stepping stone model of genetic correlation. *Journal of Applied Probability*, pp. 129–149. 1965

[WKS+11] Wacey D, Kilburn MR, Saunders M, Cliff J and Brasier MD. Microfossils of sulphur-metabolizing cells in 3.4-billion-year-old rocks of Western Australia. *Nature Geoscience*, **4**(10):pp. 698–702. 2011

[WPH+14] Weber MF, Poxleitner G, Hebisch E, Frey E and Opitz M. Chemical warfare and survival strategies in bacterial range expansions. *Journal of The Royal Society Interface*, **11**(96):p. 20140172. 2014

[WRL13a] Wiser MJ, Ribeck N and Lenski RE. Data from: Long-term dynamics of adaptation in asexual populations. *Dryad Digital Repository.* 2013 URL `http://doi:10.5061/dryad.0hc2m`

[WRL13b] Wiser MJ, Ribeck N and Lenski RE. Long-Term Dynamics of Adaptation in Asexual Populations. *Science*, **342**(6164):pp. 1364–1367. 2013

[Zim12] Zimmer C. *Microcosm: E-coli and the New Science of Life*. Random House. 2012

A. Simulation code

Here we present the core of the simulation code upon which we based our simulations, written in C.

The number of individuals in a fitness class `i` is stored in an array `n[i]`, which is allowed to be a fractional number. During a single time step of duration `epsilon`, the number of individuals changes due to reproduction, then the whole array is rescaled to fulfill the population size constraint.

Reproduction

Selection, Mutation and genetic drift

```
void ReproductionMutation() {
  int i,j;
  double tmpn;

  for (i=0; i<space; i++ ) {
    tmpn = (1. + epsilon * x[i])*n[i];

    for (j=0; j<space; j++ ) {
      tmpn += epsilon * mu * mutkernel[i,j] * n[j];
    }
    n[i] = tmpn;

    tmpn += sqrt(2.*epsilon) * (PoissonRNG(n[i]) - n[i]);
    n[i] = tmpn;
  }
}
```

Population constraint

```
void Constraint() {
  int i;
  double sum = 0.;

  for (i=0; i<space; i++ ) {
    sum += u[i] * n[i];
  }
```

```
   for (i=0; i<space; i++ ) {
     n[i] /= sum;
   }
 }
```

Of course, the whole simulation program has to contain much more than only these two snippets of code. The population is moving to higher fitness; at some point the lattice will be too small to represent the population, so the lattice has to be shifted. Input of parameters and output of mean quantities, like population size, fitness variance or mean fitness (or even the complete array `n[i]`) requires additional code. Furthermore, values in arrays that we used tacitly in their obvious interpretations, like `u[i]`, `x[i]` and `mutkernel[i,j]`, have to be initialized. Our simulation code is freely available upon request.

B. Numerical methods

In this chapter, we present various methods to compute numerical profiles in tuned models. These profiles are the fixation probability $\mathbf{u}_\star$ (section B.1) and the mean stationary population density $\overline{\mathbf{c}_\star}$ (section B.2). We also present the code in order to calculate the 2-point-correlation function $\overline{\mathbf{C}_\star^{(2)}}$ in section B.3 for the diffusion mutation kernel. Here, we only briefly mention the underlying concepts for solving these equations, an elaborate treatment on numerical methods is *e.g.* [PTVF07].

In general, note that the numerical code gets unstable for large (huge) population sizes. Then, the range from the smallest value in the profile to the largest value in the profile spans many orders of magnitude.

Note that we only reproduce the mathematical necessities and algorithmic peculiarities here. The full code, written in C, is freely available from the authors.

B.1. Multi-dimensional Newton-Raphson iteration to solve the non-linear constraint equation

With the multi-dimensional Newton-Raphson iteration one typically tries to find a solution to a set of M homogeneous relations in M variables $\mathbf{y} = \{y_1, \dots, y_M\}$. Formally, the set of equations is given by

$$f_i\big(y_1, \dots, y_M\big) = 0 \;, \quad 1 \le i \le M \;. \tag{B.1}$$

For small deviations $\delta\mathbf{y}$, we can expand the expressions for f_i's into a Taylor series,

$$f_i(\mathbf{y} + \delta\mathbf{y}) = f_i(\mathbf{y}) + \sum_{j=1}^{M} \frac{\partial f_i}{\partial y_j} \delta y_j + \mathcal{O}\big(\delta\mathbf{y}^2\big) \;. \tag{B.2}$$

The (matrix of) partial derivatives in (B.2) is a Jacobian, that we define as $J_{ij} = \partial_{y_j} f_i$. Using vector notation (B.2) can be stated as

$$\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) = \mathbf{f}(\mathbf{y}) + J\,\delta\mathbf{y} + \mathcal{O}\big(\delta\mathbf{y}^2\big) \;. \tag{B.3}$$

Starting from the current guess for a solution, now denoted $\mathbf{y}^{\text{old}}$, we want to find $\delta\mathbf{y} = \mathbf{y}^{\text{new}} - \mathbf{y}^{\text{old}}$ such that $\mathbf{f}(\mathbf{y}^{\text{new}}) = \mathbf{0}$ by *linear* extrapolation. Thus, we have

$$\mathbf{0} \overset{!}{=} \mathbf{f}(\mathbf{y}^{\text{old}}) + J\,\delta\mathbf{y} \;. \tag{B.4}$$

A single iteration step in the Newton-Raphson method consists of solving the linear extrapolation (B.4). Using $\mathbf{y}^{\mathrm{new}} = \mathbf{y}^{\mathrm{old}} + \delta\mathbf{y}$, the new guess for the solution is

$$\mathbf{y}^{\mathrm{new}} = \mathbf{y}^{\mathrm{old}} - J^{-1}\,\mathbf{f}(\mathbf{y}^{\mathrm{old}})\ . \tag{B.5}$$

Each iteration step involves the computation of the Jacobian J and its inverse J^{-1}. *If* the algorithm converges, the convergence is quadratic [PTVF07]. However, for success of the method a good initial guess is needed.

The Newton-Raphson iteration will now be applied to obtain solutions for the fixation probability $\mathbf{u}_\star$ in the comoving frame,

$$0 = -v\partial_X u_\star(X) + X u_\star(X) - 2u_\star(X)^2 + \mu \int ds\ m(s)\big(u_\star(X+s) - u_\star(X)\big)\ . \tag{B.6}$$

The comoving frame with speed v was instrumental in tuned models: the adaptation speed v replaces the population size N as parameter (*cf.* section 3.4). In order to solve (B.6), we first discretize it, using an evenly spaced lattice:

$$0 = -v/(2\,dx)\,\big(u_{i-1} - u_{i+1}\big) + x_i u_i - 2u_i^2 - \mu u_i + \mu \sum_j m_j u_{i+j} \tag{B.7}$$

Lattice points have distance dx to each other. We use M to denote the number of lattice points. Start-point and end-point of the lattice have to be chosen far enough apart, such that the different regimes are well represented on the lattice (see section 3.2.1 for a description of the three regimes). As a rule of thumb, we used roughly 100 lattice points between 0 and X_c, although probably fewer would also suffice. Check also Figures 5.5, 5.6 and 5.7, that had each $M = 1000$ lattice points. The depicted range covers the majority of the numerical solution. In addition, we usually take $\sigma/dx = 20$ for the exponential mutation kernel and $\sigma/dx = 50$ in the staircase mutation kernel.

The equations (B.7) represent the original relations in (B.1) with u_i being equivalent to the M variables y_i before. The Jacobian J is given by

$$J = \begin{pmatrix} \ddots & \vdots & & & & \\ \cdots & x_{i-1} + \gamma - 4u_{i-1} & \mu m_1 & \mu m_2 & \mu m_3 & \\ & \mu m_{-1} & x_i - \gamma - 4u_i & \mu m_1 & \mu m_2 & \\ & \mu m_{-2} & \mu m_{-1} & x_{i+1} - \gamma - 4u_{i+1} & \mu m_1 & \\ & \mu m_{-3} & \mu m_{-2} & \mu m_{-1} & x_{i+2} - \gamma - 4u_{i+2} & \cdots \\ & & & & \vdots & \ddots \end{pmatrix} . \tag{B.8}$$

The symbol $\gamma = \mu(m_0 - 1)$ is an abbreviation for the amount of neutral and outgoing mutations. In (B.8), only a part of J is stated explicitly. However, the general structure of the $M \times M$ matrix is easily deduced from (B.8).

As initial guess, we use $u_j = j\,dx/2$ for $j > 0$ and $u_j = dx$ for $j \leq 0$. This is roughly the shape for the fixation probability $\mathbf{u}_\star$ we would expect from the asymptotic solutions in regime (I) and (III). The aim of the algorithm is to find the solution to the intermediate regime (II): it is the most important range in clonal competition scenarios.

In addition, we have to consider boundary conditions. For positive fitnesses X, we analytically compute the correction one would expect for the mutation term, if the asymptotic solution $u_\star^{(III)}(X) \approx X/2$ extends to infinity. This term is added while evaluating (B.6), as if the sum over the mutation kernel extends over $j \in \mathbb{Z}$. For negative fitnesses X, the asymptotic solution $u_\star^{(I)}(X) \approx 0$ is sufficient, such that we can ignore the lower boundary. Note that these additional boundary correction terms only apply to the exponential mutation kernel. The staircase and diffusion mutation kernel have a very limited range, hence elaborate corrections are not needed.

Moreover, we have to deal with the discretized derivative from the comoving frame, that includes the adjacent lattice points in each equation. At the last point of the lattice, we use $X/2$ as approximation for the next point that is not on the lattice. At the lower boundary, we can ignore the coupling to the previous point that is not on the lattice, as there $u_\star^{(I)}(X) \approx 0$ is sufficient as approximation.

Usually, the algorithm converges in less than 1000 iterations. However, for some (extreme) parameter values, the algorithm fails to converge and the solution starts oscillating on a *part of the lattice*, around $X \approx 0$. This is similar to the “badly behaved upside-down Gaussian”-solution described in [Fis13]: we obtain approximately $|u_j| \sim \exp\big(\alpha(j\,dx)^2\big)$, but with alternating signs. For large, but not extreme, values decreasing dx and increasing M helps to circumvent this problem, but for extreme parameter values, the algorithm does not converge. These “extreme” parameter values usually correspond to very large population sizes N.

The numerical simulation code, written in C, is freely available upon request from the authors.

B.2. Numerical methods for the mean stationary population density

For the mean stationary population density $\overline{\mathbf{c}_\star}$, the multi-dimensional Newton-Raphson iteration described above does not work as presented. This is due to the fact that the governing (integro-)differential equation is *linear* in $\overline{\mathbf{c}_\star}$. Thus, the M relations (B.1) in the M variables y_i can be stated as

$$0 = f_i\big(\mathbf{y}\big) = \sum_j F_{ij} y_j\ , \quad 1 \le i \le M\ . \tag{B.9}$$

For this linear system of equations, the Jacobian J has a particular simple form, $J_{ij} = F_{ij}$. Inserting J into (B.5) yields the trivial solution, $\mathbf{y}^{\text{new}} = \mathbf{y}^{\text{old}} - J^{-1} F \mathbf{y}^{\text{old}} = \mathbf{0}$. After a single iteration step the algorithm collapses.

Therefore, we have to use different methods to solve the equation for $\overline{\mathbf{c}_\star}$. Recall, that the mean stationary population density is governed by (3.59). The discretized version is

$$0 = v/(2\,dx)\big(c_{i+1} - c_{i-1}\big) + \big(x_i - 2u_i - \mu\big)c_i + \mu \sum_j m_j c_{i-j}\ . \tag{B.10}$$

The desired solution c_i is in the nullspace of the coefficient matrix F (see (B.9)).

In the following we describe algorithms for this task.

Staircase mutation kernel

For the staircase mutation kernel we employ the Moore-Penrose-pseudoinverse [Pen55] to obtain $\overline{\mathbf{c}_\star}$. The pseudoinverse can be used to construct a least-squares approximation $\mathbf{y}$ to the linear equation $A\mathbf{y} = \mathbf{b}$, where A is not necessarily a square matrix [Pen56].

We define A as the matrix F from (B.9) with an additional row at the bottom. The entries in this last row are given by the solution for the fixation probability u_j. In addition, $\mathbf{b}$ is defined as the vector $\big(0, \dots, 0, 1/dx\big)$. Thus, the last row indicates the discretized version of the constraint $\big\langle \mathbf{u}_\star \big| \overline{\mathbf{c}_\star} \big\rangle = 1$. As it is simultaneously imposed onto the solution we arrive at the correct scaling for c_i.

Note that results start to acquire numerical noise at density values that are several orders of magnitude (roughly $\sim \mathcal{O}\big(10^{15}\big)$) below the maximum of the bell-shaped density, at $X \approx 0$. In most figures shown in this thesis these effects would not be visible (or only barely at boundaries). Still, we replaced this obvious noise with the semi-analytical approximations $\sim \exp\big(-\alpha X^2\big)$ for negative fitness and $\sim \exp\big(-\beta X\big)$ for positive fitness. The values of α and β were adjusted to fit the non-noisy part of the solution.

Using the Moore-Penrose-pseudoinverse is the most general method to obtain a solution c_i. It is also applicable to the other mutation kernels. However, we use more elaborate algorithms there, that do not suffer from the numerical errors as this method described here.

Diffusion mutation kernel

For the diffusion mutation kernel, we use a "local" version of the Newton-Raphson iteration. It yields more reliable results than the Moore-Penrose pseudoinverse; numerical oscillations do usually not occur. As mentioned above, the full Newton-Raphson method yields the trivial solution. The "local" version of this method does not use the inverse of the Jacobian, but expands the relation $f_i(\mathbf{y})$ *only* in the variable y_i. Carrying out the same manipulations as above, we arrive at

$$y_i^{\text{new}} = y_i^{\text{old}} - \frac{f_i(\mathbf{y}^{\text{old}})}{\partial_{y_i} f_i(\mathbf{y}^{\text{old}})} \,. \tag{B.11}$$

Note, that the Jacobian in (B.5) is now replaced with the ("local") derivative with respect to the variable y_i in (B.11).

In particular for the diffusion mutation kernel, the two relations $f_i(\mathbf{c})$ and $\partial_{c_i} f_i(\mathbf{c})$ are

$$\begin{aligned} f_i(\mathbf{c}) &= v/(2\,dx)\big(c_{i+1} - c_{i-1}\big) + D/(dx^2)\big(c_{i+1} - 2c_i + c_{i-1}\big) \\ &\quad + \big(x_i - 2u_i\big)c_i \,, && \text{(B.12a)} \\ \partial_{c_i} f_i(\mathbf{c}) &= \big(-2D/(dx^2) + x_i - 2u_i\big)c_i \,. && \text{(B.12b)} \end{aligned}$$

For boundary conditions it usually suffices to set the next and previous lattice point that are not in the lattice to zero. Convergence is much slower than in the full Newton-Raphson iteration using the Jacobian in the number of steps needed for a reasonably accurate solution.

Exponential mutation kernel

For the exponential mutation kernel we can use the Moore-Penrose pseudoinverse. However, as above we rely on the more accurate "local" Newton-Raphson iteration here. However, the simple implementation used in the diffusion mutation kernel does not work here: the derivation with respect to c_i, similar to (B.12b), changes its sign on the lattice. $\partial_{c_i} f_i(\mathbf{c})$ is negative for negative fitness, changes its sign below $X = 0$ to positive, and again to negative after $X = 0$. This effect can be seen in the graph of the function $\big(x - 2\mathbf{u}_\star\big)$, depicted in Figure 3.3B. The mutation term in the diffusion mutation kernel in the algorithm is $-2D/(dx^2)$, which renders $\partial_{c_i} f_i(\mathbf{c})$ negative for all i. On the other hand, the mutation term in the exponential mutation kernel is $\gamma = \mu(m_0 - 1)$, which is tiny compared to $\big(x - 2\mathbf{u}_\star\big)$ for reasonable mutation rates μ. Therefore, the large peak around $X = 0$ remains (*cf.* Figure 3.3B).

An alternating sign in the derivative $\partial_{c_i} f_i(\mathbf{c})$ leads to large numerical oscillations in the solution and the algorithm does not converge.

Therefore, we apply the "local" Newton-Raphson iteration to equation C.8 derived below. It does not suffer the problem of alternating signs.

B.3. Solving the 2-point-correlation-function

As of writing this thesis, we only implemented the code for the diffusion mutation kernel. These results are shown in Figure 3.5. The other two mutation kernels suffer from the fact that we cannot implement the "local" Newton-Raphson iteration. In both kernels, the derivative $\partial_{c_i} f_i(\mathbf{c})$ changes sign on the lattice. Moreover, for reasonable lattice sizes ($M \approx 500 - 1000$), the resulting matrix (that scales as M^4) for the Moore-Penrose pseudoinverse is too large. Other methods have to be found.

C. Analytical addendum

C.1. Integral transforms for exponential mutation kernels

In this section we present additional analytics for the exponential mutation kernel. We show that an integral transform removes the integration, while at the same time introduces another derivation with respect to fitness. A linearized equation (neglecting the quadratic term) can be solved analytically, yielding a result for the shape of $\mathbf{u}_\star$ far below the critical fitness X_c. In this regime (denoted regime (I) in the main text) the methods employed so far could not obtain proper solutions. The derivation here is based on unpublished notes [Hal11a]. It is included here for completeness, as we use these results at various places in this thesis.

Recall, that $\mathbf{u}_\star$ is governed by

$$v\partial_X u_\star(X) = X u_\star(X) - 2u_\star(X)^2 + \mu/\sigma \int\limits_0^\infty ds\ e^{-s/\sigma}\big(u_\star(X+s) - u_\star(X)\big)\ . \tag{C.1}$$

Next, we introduce $k(X) = \sigma^{-1}\int ds\ e^{-s/\sigma} u_\star(X+s)$, such that we can split (C.1) into two coupled equations,

$$\begin{aligned} v\partial_X u_\star(X) &= (X-\mu)u_\star(X) + \mu k(X) - 2u_\star(X)^2\ , && \text{(C.2a)}\\ \sigma\partial_X k(X) &= k(X) - u_\star(X)\ . && \text{(C.2b)} \end{aligned}$$

Subsequently, we use $U_\star(X) = u_\star(X)e^{-X/\sigma}$ and $K(X) = k(X)e^{-X/\sigma}$ to arrive at

$$\begin{aligned} v\partial_X U_\star(X) &= (X-\mu-v/\sigma)U_\star(X) + \mu K(X) - 2U_\star(X)e^{X/\sigma}\ , && \text{(C.3a)}\\ \sigma\partial_X K(X) &= -U_\star(X)\ . && \text{(C.3b)} \end{aligned}$$

After a derivative with respect to X of (C.3a), we can insert (C.3b), which yields

$$vU_\star''(X) = \big(X-\mu-v/\sigma\big)U_\star'(X) + \big(1-\mu/\sigma\big)U_\star(X) - 2e^{X/\sigma}\Big(\big(U_\star(X)^2\big)' + U_\star(X)^2/\sigma\Big)\ . \tag{C.4}$$

Finally, transforming back from $U_\star(X)$ to $u_\star(X)$, we arrive at

$$0 = -vu_\star''(X) - \big(X-\mu+v/\sigma\big)u_\star'(X) + \big(X/\sigma-1\big)u_\star(X) - 2u_\star(X)^2 + 2\left(u_\star(X)^2\right)'\ , \tag{C.5}$$

with primes $'$ denoting derivatives with respect to the argument X.

The main reason to state these questions is to analyze the behavior of $\mathbf{u}_\star$ for small (and negative) fitness X. To this end, we investigate the transformed solution $U_\star(X)$ in (C.4). In the range of interest, the exponential in (C.4) dampens the non-linear terms, such that we will neglect them. The resulting differential equation is similar to Hermite's equation and can be cast into the form

$$f''(Y) - Yf'(Y) - \lambda f(Y) = 0 \,, \tag{C.6}$$

with the identification $Y = \big(X - \mu - v/\sigma\big)/\sqrt{v}$ and $\lambda = 1 - \mu/\sigma$. An Ansatz of the form $f(Y) = \sum_n a_n Y^n$ (see [CH66]) shows that there are two independent solutions, an odd function and an even function. The coefficients are given by the recursion relation, $a_{n+2} = (n+\lambda)(n+2)^{-1}(n+1)^{-1}a_n$. Relabeling indices to steps of 1, we arrive at an expression using Kummer's confluent hypergeometric function ${}_1F_1$,

$$\begin{aligned} U_\star(X) = u_\star(X)e^{-X/\sigma} \;\; &= \;\; \alpha\, {}_1F_1\left(\frac{1-\mu/\sigma}{2}, \frac{1}{2}; \frac{\big(X - v/\sigma\big)^2}{2v}\right) \\ &\quad + \beta\big(X - v/\sigma\big)\, {}_1F_1\left(\frac{2-\mu/\sigma}{2}, \frac{3}{2}; \frac{\big(X - v/\sigma\big)^2}{2v}\right) \,, \end{aligned} \tag{C.7}$$

In the argument we assumed a small mutation rate, such that $v/\sigma + \mu \approx v/\sigma$. The upper line indicates the even function, whereas the lower line is an odd function (in the variable Y). As boundary condition we require that the two coefficients α and β are chosen such that the two independent solutions stay positive, but grow slower than the exponential $\exp\big(-X/\sigma\big)$ on the left hand side: ultimately we want $u_\star(X) \overset{X\to-\infty}{\longrightarrow} 0$.

A immediate observation in (C.7) is that the behavior changes qualitatively at $X - v/\sigma = 0$. Below that point, the upward curvature of the even function and the downward curvature of the odd function have to cancel each other, due to boundary conditions. Therefore, we expect the main contribution to the solution $\mathbf{u}_\star$ to be the exponential, as claimed in the main text (see section 3.2.1). For positive $X - v/\sigma$ both, the even and odd, solutions bend upwards (a negative solution is not allowed), leading to the regime we had assigned the inverse Gaussian $\sim \exp\big(X^2/(2V)\big)$ as solution in the main text. Overall, we expect the transition from regime (II) to regime (I) for the exponential mutation kernel to be at $X_c^{(II)\to(I)} \approx v/\sigma$.

These observations also comply with numerical solutions to the full non-linear equation for $\mathbf{u}_\star$, as can be checked *e.g.* in Figure 5.5.

Similarly, we can use $\widetilde{k}(X) = \int ds/\sigma e^{-s/\sigma}\overline{c_\star}(X-s)$ to cast the equation for the stationary population density $\overline{\mathbf{c}_\star}$ from an integro-differential equation to a simple ODE:

$$0 = v\overline{c_\star}''(X) + \big(X - \mu + v/\sigma\big)\overline{c_\star}'(X) + \big(X/\sigma + 1 - 2u_\star(X)/\sigma\big)\overline{c_\star}(X) - 2\big(u_\star(X)\overline{c_\star}(X)\big)' \,. \tag{C.8}$$

The final two equations, (C.5) and (C.8), are now much more amenable to numerical methods, as integration often complicates the analysis enormously.

For completeness, we also state the solution to the linearized version of (C.8). We neglect the last two terms involving our constraint $\mathbf{u}_\star$ to arrive at

$$\begin{aligned}\overline{c_\star}(X)e^{X^2/(2v)} &= \alpha\,{}_1F_1\left(-\frac{\mu}{2\sigma},\frac{1}{2};\frac{\left(X-v/\sigma\right)^2}{2v}\right)\\ &\quad+\beta\big(X-v/\sigma\big)\,{}_1F_1\left(\frac{1-\mu/\sigma}{2},\frac{3}{2};\frac{\left(X-v/\sigma\right)^2}{2v}\right)\,, \qquad \text{(C.9)}\end{aligned}$$

The main contribution in (C.9) is the expected Gaussian shape for the population density. The coefficients α and β have to be adjusted such that the overall solution decays to zero on both sides, $\overline{c_\star(X)} \to 0$ for $X \to \pm\infty$.

C.2. Fixation probability for moment closures at higher levels

In section 3.6 we introduced constrained BRWs where the moment hierarchy is closed an arbitrary, but chosen, level n and not at the first moment. The dynamics for the correlation function $\overline{\mathbf{C}_t^{(n)}}$ in these extended tuned model is given by (3.70). All lower correlation functions $(m < n)$ can be obtained by the marginalization

$$\overline{\mathbf{C}_t^{(m)}} = \big\langle u_\star^{(n)} | \overline{\mathbf{C}_t^{(m+1)}} \big\rangle_i\,, \quad 1 \le i \le m+1\,. \qquad \text{(C.10)}$$

Here, the index i denotes the variable in the correlation function that is integrated over and therefore vanishes.

In this section, we want to investigate fixation probabilities of individuals in this modified time evolution. Similar to section 3.3.1 we define a sub-population $\mathbf{w}_t$ and track its fate. On long times, it will either take over the population and fixate, or go extinct. However, when closing the moment hierarchy at $n > 1$, the dynamics of the (mean) sub-population cannot be investigated directly as it couples to higher moments, in particular $\overline{\mathbf{w}_t\mathbf{c}_t}$:

$$\partial_t\overline{\mathbf{w}_t} = \big(\mathcal{L} - 4\mathbf{u}_\star^{(n)}\big)\overline{\mathbf{w}_t} - 2(n-2)\big\langle \mathbf{u}_\star^{(n)}\,\mathbf{u}_\star^{(n)} | \overline{\mathbf{w}_t\mathbf{c}_t}\big\rangle\,. \qquad \text{(C.11)}$$

In section 3.3.1 we used the Green's operator and imposed the infinite-time limit onto it. However, a Green's operator is defined for *linear* operators. Therefore, the lowest moment, where the dynamics is governed by a linear operator, is the n^{th} moment – exactly at the closure of the moment hierarchy. Thus, we have to investigate the dynamics of the *correlator*

$$\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}} \equiv \overline{w_t(x_1)c_t(x_2)\dots c_t(x_n)}\,. \qquad \text{(C.12)}$$

Its dynamics is given by

$$\partial_t\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}} = \sum_i \big(\mathcal{L} - 2n\mathbf{u}_\star^{(n)}\big)_i\,\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}} + \sum_{i\neq j} 2\delta(x_i - x_j)\left\langle \mathbf{u}_\star^{(n)} | \overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}}\right\rangle_i\,. \qquad \text{(C.13)}$$

Now we can turn our attention to the long-time behavior of the correlator: we know that the sub-population either takes over the population or goes extinct. Thus, at a time τ far ahead of t we have

$$\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}} \overset{\tau-t\to\infty}{\longrightarrow} F\,\overline{\mathbf{C}_\star^{(n)}} + (1-F)\,\mathbf{0}\,. \tag{C.14}$$

Here, F denotes the fixation probability of the sub-population $\mathbf{w}_t$. Similar to section 3.3.1, we state the dynamics of the Green's operator for the *linear* dynamics (in the comoving frame) of the correlator,

$$\mathcal{G}\,\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}} = \int d\sigma(\lambda)e^{\lambda(\tau-t)}\left\langle\mathbf{\Phi}_{\lambda,t}^{(n),(L,v)}\middle|\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}}\right\rangle\mathbf{\Phi}_{\lambda,\tau}^{(n),(R,v)}\,, \tag{C.15}$$

where $\mathbf{\Phi}_{\lambda,t}^{(n),(L,v)}$ and $\mathbf{\Phi}_{\lambda,\tau}^{(n),(R,v)}$ are the left (L) and right (R) eigenvectors with the appropriate indices (eigenvalue λ, at time t or τ, in comoving frame with speed v, closure at n^{th} moment). However, as the constant fixation probability drags the population density along, we can deduce that $\lambda = 0$ is again the largest eigenvalue for the operator in the comoving frame (as $\lambda = 0$ is the larges eigenvalue for $n = 1$, see Section 3.3.1). In the long-time limit, all other eigenvalues decay exponentially and we are left with the stationary correlation function of the population density as the right eigenvector and an unknown left eigenvector, both to eigenvalue $\lambda = 0$:

$$\mathbf{\Phi}_{0,\tau}^{(n),(R,v)} = \overline{\mathbf{C}_\star^{(n)}}\,, \tag{C.16a}$$

$$\mathbf{\Phi}_{0,t}^{(n),(L,v)} \overset{!}{=} \mathbf{B}^{(n)}\,. \tag{C.16b}$$

Here, we introduced the much simpler notation $\mathbf{B}^{(n)}$ to denote the eigenvector we seek. In this notation, the Green's operator converges to

$$\mathcal{G}\,\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}} \overset{\tau-t\to\infty}{\longrightarrow} \left\langle\mathbf{B}^{(n)}\middle|\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}}\right\rangle\overline{\mathbf{C}_\star^{(n)}}\,. \tag{C.17}$$

In order to obtain the fixation probability, we have to find $\mathbf{B}^{(n)}$.

The left eigenvector is a solution to the adjungated form of the operator. To this end, we have to adjungate the integro-differential operator acting on the correlator $\overline{\mathbf{w}_t\mathbf{C}_t^{(n-1)}}$ in (C.13). The first term, $\left(\mathcal{L}-2n\mathbf{u}_\star^{(n)}\right)_i$ can readily be adjungated to $\left(\mathcal{L}^\dagger-2n\mathbf{u}_\star^{(n)}\right)_i$. In the second term the Dirac-δ causes (minor) problems at first, but we can find a weak solution using partial integration.

$$\begin{aligned}
&\int d(x)_n B^{(n)}\big((x)_n\big)\left[\sum_{i\neq j}2\delta(x_i-x_j)\left\langle\mathbf{u}_\star^{(n)}\middle|\overline{\mathbf{C}_t^{(n)}}\right\rangle_i\right]\\
=\;&\sum_{i\neq j}2\int d(x)_n dy B^{(n)}\big((x)_n\big)\delta(x_i-x_j)u_\star^{(n)}(y)\overline{C_t^{(n)}}(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n)\\
=\;&\sum_{i\neq j}\int d(x)_n dy B^{(n)}(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n)\delta(y-x_j)u_\star^{(n)}(x_i)\overline{C_t^{(n)}}\big((x)_n\big)
\end{aligned}$$

$$
\begin{aligned}
& +\sum_{i\neq j}\int d(x)_n dy B^{(n)}(x_1,\dots,x_{j-1},y,x_{j+1},\dots,x_n)\delta(x_i-y)u_\star^{(n)}(x_j)\overline{C_t^{(n)}}\big((x)_n\big) \\
= & \sum_{i\neq j}\int d(x)_n \overline{C_t^{(n)}}\big((x)_n\big)B^{(n)}(x_1,\dots,x_{i-1},x_j,x_{i+1},\dots,x_n)u_\star^{(n)}(x_i) \\
& +\sum_{i\neq j}\int d(x)_n \overline{C_t^{(n)}}\big((x)_n\big)B^{(n)}(x_1,\dots,x_{j-1},x_i,x_{j+1},\dots,x_n)u_\star^{(n)}(x_j) \\
= & \int d(x)_n \overline{C_t^{(n)}}\big((x)_n\big)\left[\sum_{i\neq j}B^{(n)}\big((x)_{n;i\mapsto j}\big)u_\star^{(n)}(x_i)+B^{(n)}\big((x)_{n;j\mapsto i}\big)u_\star^{(n)}(x_j)\right] \quad \text{(C.18)}
\end{aligned}
$$

If appropriate, several variables have been denoted by a single symbol to simplify notation, $(x)_n = \{x_1, x_2, \dots, x_n\}$. From (C.18) the adjungated (time-reversed) differential equation follows as

$$
-\partial_t \mathbf{B}^{(n)} = \sum_i \left(\mathcal{L}^\dagger - 2n\mathbf{u}_\star^{(n)}\right)_i \mathbf{B}^{(n)} + \sum_{i\neq j}\left(u_\star^{(n)}(x_i)\mathbf{B}_{i\mapsto j}^{(n)} + u_\star^{(n)}(x_j)\mathbf{B}_{j\mapsto i}^{(n)}\right) , \quad \text{(C.19)}
$$

where the subscripted arrows $(i \mapsto j)$ denote the replacement of x_i with x_j in $\mathbf{B}^{(n)}$. We propose the solution

$$
B^{(n)}(x_1,\dots,x_n) = u_\star^{(n)}(x_1)\dots u_\star^{(n)}(x_n) \; . \quad \text{(C.20)}
$$

Upon inserting this expression into (C.19), the terms in the second sum all have the form

$$
u_\star^{(n)}(x_1)\dots u_\star^{(n)}(x_i)u_\star^{(n)}(x_i)\dots u_\star^{(n)}(x_n) \; .
$$

Incidentally, this is the same structure as the second term in the first sum. Inserting into the (backward-in-time) dynamics, (C.19), we arrive at

$$
\begin{aligned}
-\partial_t B^{(n)}\big((x)_n\big) &= \sum_i \big(\mathcal{L}_i^\dagger - 2nu_\star^{(n)}(x_i) - (n-1)u^{(n)}(x_i)\big)B^{(n)}\big((x)_n\big) \\
&= \sum_i \big(\mathcal{L}_i^\dagger - (n+1)u_\star^{(n)}(x_i)\big)\big(u_\star^{(n)}(x_1)\dots u_\star^{(n)}(x_n)\big) \; . \quad \text{(C.21)}
\end{aligned}
$$

Each term involves the condition of moment closure, $0 = \big(\partial_t + \mathcal{L}^\dagger - (n+1)\mathbf{u}_\star^{(n)}\big)\mathbf{u}_\star^{(n)}$. In turn, the proposed solution (C.20) fulfills the dynamics (C.19) (although other solutions might exist).

With the left eigenvector, we can evaluate (C.14) and (C.17) above. It follows that the fixation probability of the sub-population $\mathbf{w}_t$ in the case of the closure at the n-th moment can be computed as

$$
F = \left\langle \mathbf{B}^{(n)} | \overline{\mathbf{w}_t \mathbf{C}_t^{(n-1)}} \right\rangle_{1\dots n} = \left\langle \mathbf{u}_\star^{(n)} | \overline{\mathbf{w}}_t \right\rangle_1 \; . \quad \text{(C.22)}
$$

Using a single individual as starting condition for the sub-population, $w_t(x) = \delta(x - x_0)$, the spatial fixation probability is simply $u_\star^{(n)}(x_0)$, generalizing the result $u_\star(x_0)$ from $n = 1$ in section 3.3.1.

C.2.1. Measurement of fixation probabilities in simulations

Similar the procedure we employed in the main text for $n = 1$ (see section 3.3.3), we can try to "measure" the fixation probability for higher moments in stochastic simulations. To this end, we take 100 independent starting configurations of the population density, reset time to $t = 0$ and label a sub-population $\mathbf{w}_0$ of the most fit individuals in the nose of the wave such that

$$\left\langle \mathbf{u}_\star^{(n)} \middle| \mathbf{w}_0 \right\rangle = 0.5 \; . \tag{C.23}$$

Then, we run stochastic simulations with the constraint $\mathbf{u}_\star^{(n)}$. If the sub-population grows or shrinks, such that it passes one of the thresholds, $\left\langle \mathbf{u}_\star^{(n)} \middle| \mathbf{w}_t \right\rangle = 10^{-4}$ or $\left\langle \mathbf{u}_\star^{(n)} \middle| \mathbf{w}_t \right\rangle = 1 - 10^{-4}$, we stop the simulation and record the event as extinction or fixation, respectively. After we recorded 200 events from each initial condition, we use Bayesian inference to estimate the 95%-confidence interval for the expected fixation probability. The derivation using Bayes' theorem is presented in the main text (in section 3.3.3). If 0.5 is contained within the interval, our initial hypothesis is most likely correct: $\mathbf{u}_\star^{(n)}$ is the fixation probability in extended tuned models for closure at n.

In Figure C.1 we present results for all three mutation kernels. The 95%-confidence interval is depicted for $n = 1$, $n = 2$ and $n = 3$. Results are rather noisy. All three mutation kernels show different trends. In the exponential mutation kernel the confidence intervals scatters for larger n, but no clear trend is discernible. In the staircase mutation kernel some initial conditions appear to clearly favor fixation for larger n, while most other intervals are still in the expected range. In the diffusion mutation kernel extinction events occur more often than expected.

While some of the scattering can be explained by discreteness effects in simulations compared to the analytic, future investigations might establish a profound interpretation of these results.

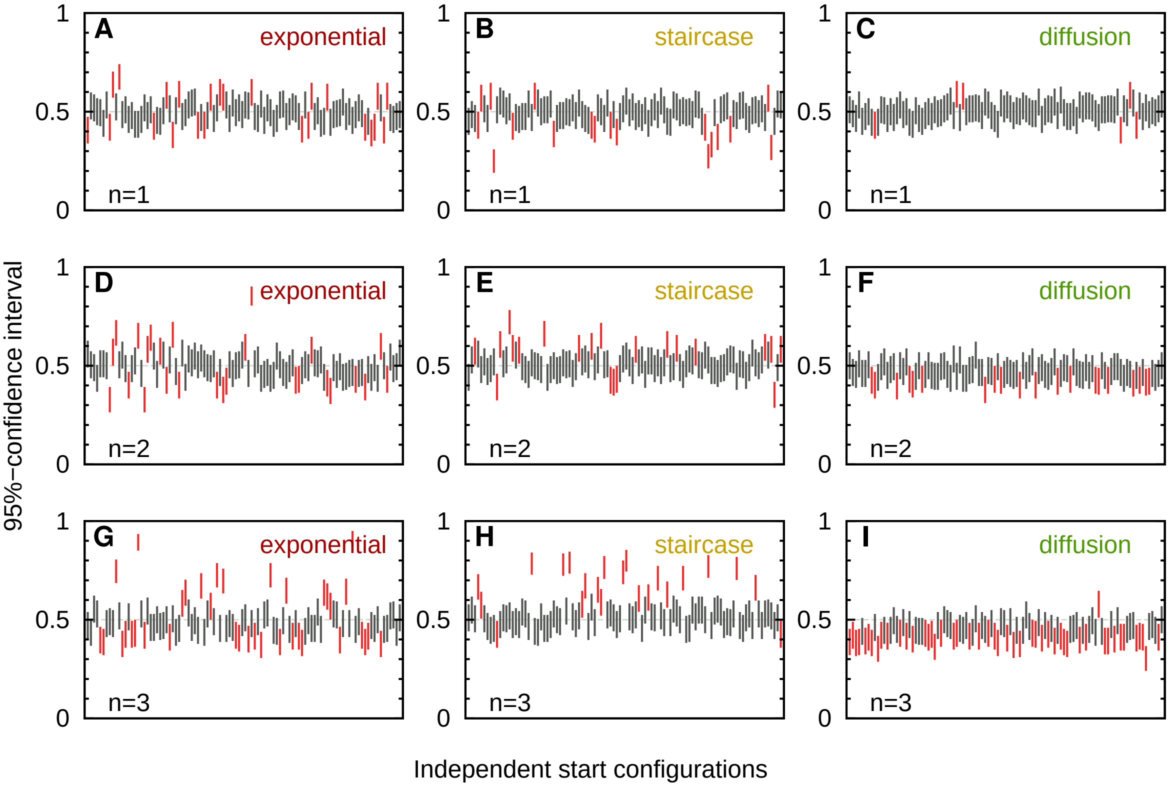

Figure C.1.: **Measurement of fixation and extinction events for closures at $n = 1$, $n = 2$ and $n = 3$.** Here, the 95%-confidence interval for the inferred fixation probability is depicted as black line if the expected value of 0.5 is contained within the interval, or as red line if the expected value is not contained within the interval. The interval is obtained by Bayesian inference, as explained in detail in section 3.3.3. Results for $n = 1$ are identical to those presented in Figure 3.2. While for $n = 1$ the original assumption for $\mathbf{u}_\star \equiv \mathbf{u}_\star^{(1)}$ being the fixation probability appears to hold, no clear trend can be discerned for higher n (that is independent of the mutation kernels). Although we always observe both, fixation and extinction, some initial conditions appear to be clearly biased to one direction. Decreasing the time-step ϵ in the simulation reduces numerical noise, and more confidence intervals contain 0.5. Still, deviations from expectations exist.

D. Oscillations compared to fitness scales in tuned models

In the main text, we introduced the autocorrelation function (ACF) of fitness variance in chapter 5. We briefly mentioned that tuned models can be used to compare timescales extracted from simulation data to exact solutions. These timescales in the ACF were the decay time τ_{dec} and the oscillation time τ_{osc}, as detailed in section 5.2.3. Here, we present additional numerical results for profiles obtained with tuned models. To be able to compare fitness scales in profiles to the timescales, we first convert times to fitness by multiplying it with adaptation speed. The resulting quantities are then the average fitness the traveling wave has moved within the decay time or the oscillation time:

$$v\tau_{dec} \quad \text{and} \quad v\tau_{osc} \, . \tag{D.1}$$

In the next section, we explain various crucial distributions in traveling waves. After that, in section D.2, we explicitely compare the relevant scales in the ACF to the distributions in all three mutations kernels. We state our conclusions in section D.3.

D.1. Fitness distributions in traveling wave models

Traveling wave models have multiple distributions that can be used to infer crucial features of the dynamics. While those distributions are often hard to estimate in general traveling wave models, we can compute them exactly in tuned models. In the main text, we introduced already some of those distribution in fitness space. Section 3.3 dealt with the fixation probabilities $\mathbf{u}_\star$. In section 3.5 we presented the governing equation and the associated solution for the stationary population density $\overline{\mathbf{c}_\star}$. Moreover, the mutation kernel $\mathbf{m}$ is also one of the important distributions, even if it is a parameter of the model.

These three distributions, $\mathbf{u}_\star$, $\overline{\mathbf{c}_\star}$ and $\mathbf{m}$, can be used to derive other dynamically relevant ones. The first of those has been introduced in section 3.5.2: the *distribution of fixating clones* $\mathbf{g}_\star$, defined as the probability of a clone to reach fixation. Note that this is *not* identical to the fixation probability $\mathbf{u}_\star$ of an *individual*, as $\mathbf{g}_\star$ weights fixation probabilities with actual occupancies,

$$\mathbf{g}_\star = \mathbf{u}_\star \overline{\mathbf{c}_\star} \, . \tag{D.2}$$

This distribution does not directly depend on the mutation kernel. The influence of the mutation kernel is implicit in the solution to the fixation probability $\mathbf{u}_\star$ and subsequently in the mean stationary population density $\overline{\mathbf{c}_\star}$. Thus, $\mathbf{g}_\star$ denotes only accumulated fitness effects of *multiple* beneficial mutations. It does not discern the number of mutations leading to the current fitness value.

In the following, we present additional distributions that describe the effects of a *single* mutation on the dynamics. All of these distributions include $\mathbf{u}_\star$, $\overline{\mathbf{c}_\star}$ and $\mathbf{m}$, but combine them in different ways. All of them include the occurrence of a novel mutation that fixates, however.

The *distribution of fixating mutations* $\boldsymbol{\rho}_f$ was already used in [GRB+12]. In this thesis it was introduced in section 3.5.3. It is defined as the probability of fixation for an individual after having an fitness increase by x ($\sim \int dy\ \overline{c_\star(y-x)}u_\star(y)$), weighted by the actual probability of having a mutation of effect x:

$$\rho_f(x) \sim m(x) \int dy\ \overline{c_\star(y-x)}u_\star(y)\ . \tag{D.3}$$

The *distribution of backgrounds* $\boldsymbol{\rho}_{bg}$ focuses on the possible genetic backgrounds: given a fitness x, the probability of fixation is averaged over all possible fitness effects of a novel mutation, then the resulting distribution is weighted with the current population density:

$$\rho_{bg}(x) \sim \overline{c_\star(x)} \int dy\ m(y)u_\star(x+y)\ . \tag{D.4}$$

The *distribution of "winning tickets"* $\boldsymbol{\rho}_{win}$ adds the effect of a single novel mutation to its previous fitness background, before the resulting individual goes to fixation:

$$\rho_{win}(x) \sim u_\star(x) \int dy\ m(y)\overline{c_\star(x-y)}\ . \tag{D.5}$$

D.2. Comparing numerical profiles to relevant timescales

All distributions in the previous section help to discern dynamic effects in the adaptation process as modeled by traveling waves. In order to establish a relation with the two timescales of the ACF, we compute the mean fitness effect and the variance in fitness for each of those distributions ($\mathbf{g}_\star$, $\boldsymbol{\rho}_f$, $\boldsymbol{\rho}_{bg}$, $\boldsymbol{\rho}_{win}$). Then, we compare those fitness scales with the (converted) timescales of the simulations. Moreover, we also use the stochastic edge X_c as additional fitness scale. Results are shown in Figures D.1, D.2 and D.3.

Exponential mutation kernel

In the exponential mutation kernel, all four distributions are different,

$$\begin{aligned}
\mathbf{g}_\star &= \mathbf{u}_\star\overline{\mathbf{c}_\star}\ , && \text{(D.6a)}\\
\boldsymbol{\rho}_f &\sim \mathbf{m}\big(\mathbf{u}_\star * \overline{\mathbf{c}_\star}\big)\ , && \text{(D.6b)}\\
\boldsymbol{\rho}_{bg} &\sim \overline{\mathbf{c}_\star}\big(\mathbf{u}_\star * \mathbf{m}\big)\ , && \text{(D.6c)}\\
\boldsymbol{\rho}_{win} &\sim \mathbf{u}_\star\big(\mathbf{m} * \overline{\mathbf{c}_\star}\big)\ . && \text{(D.6d)}
\end{aligned}$$

Note that the mathematical operations denoted by $*$ are related to convolution operations, but have different sign conventions for each of the integrals. Refer to (D.3), (D.4) and (D.5) for the exact definitions.

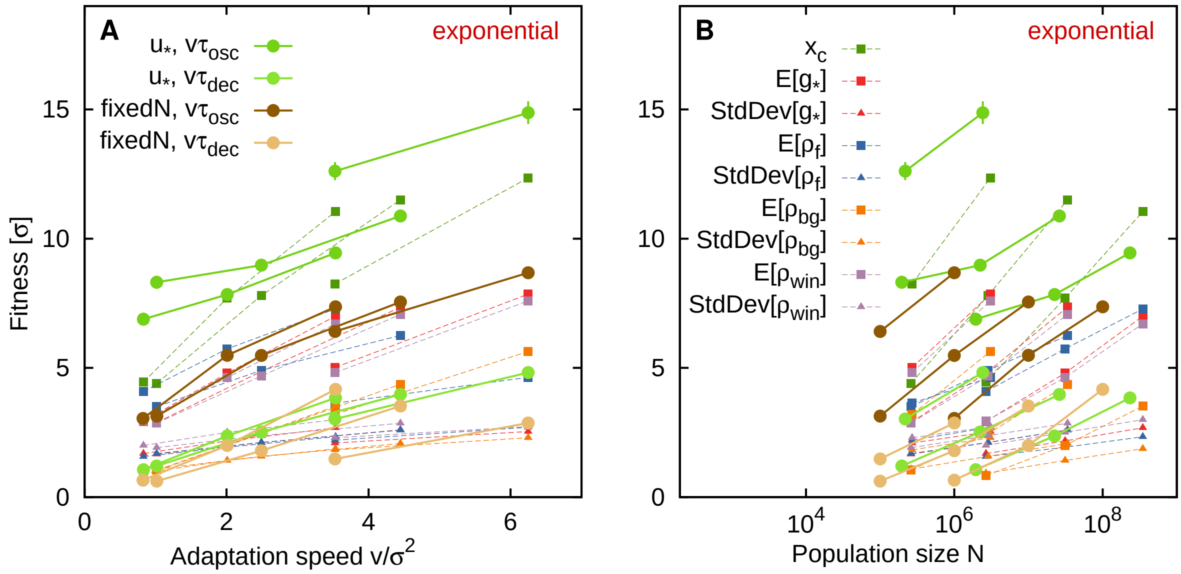

Figure D.1.: **Fitness increase during the oscillation and decay timescales, $v\tau_{osc}$ and $v\tau_{dec}$, compared to various characteristic fitness scales of distributions in tuned models with the exponential mutation kernel.** Linked points have identical mutation rates. From top to bottom, values are $\mu = 10^{-4}$, $\mu = 10^{-5}$ and $\mu = 10^{-6}$. The mean mutational effect is $\sigma = 10^{-3}$. Data in (A) and (B) are from identical simulations and identical numerical profiles, but shown with a different dependence on simulation parameters. The profiles $\mathbf{u}_\star$, $\overline{\mathbf{c}_\star}$ and $\mathbf{g}_\star$ are shown in Figure 5.5. Simulation data on the adaptation speed was presented in Figure 4.3. The autocorrelation function of fitness variance is exhibited in Figure 5.3, together with the two time scales τ_{osc} and τ_{dec} extracted from its shape.

Staircase mutation kernel

In the staircase mutation kernel, the mutation kernel $m(s) = \delta\big(s - \sigma\big)$ simplifies some of those distributions,

$$\begin{aligned} \mathbf{g}_\star &= \mathbf{u}_\star\overline{\mathbf{c}_\star}\ , && \text{(D.7a)}\\ \rho_f(x) &\sim \delta\big(x-\sigma\big)\ , && \text{(D.7b)}\\ \rho_{bg}(x) &\sim \overline{c_\star(x)}u_\star(x+\sigma)\ , && \text{(D.7c)}\\ \rho_{win}(x) &\sim u_\star(x)\overline{c_\star(x-\sigma)} \sim \rho_{bg}(x-\sigma)\ . && \text{(D.7d)} \end{aligned}$$

Diffusion mutation kernel

In the diffusion mutation kernel, we do not have an expression for the function $\mathbf{m}$. The dynamical equations follow from a Kramers-Moyal expansion of this mutation kernel,

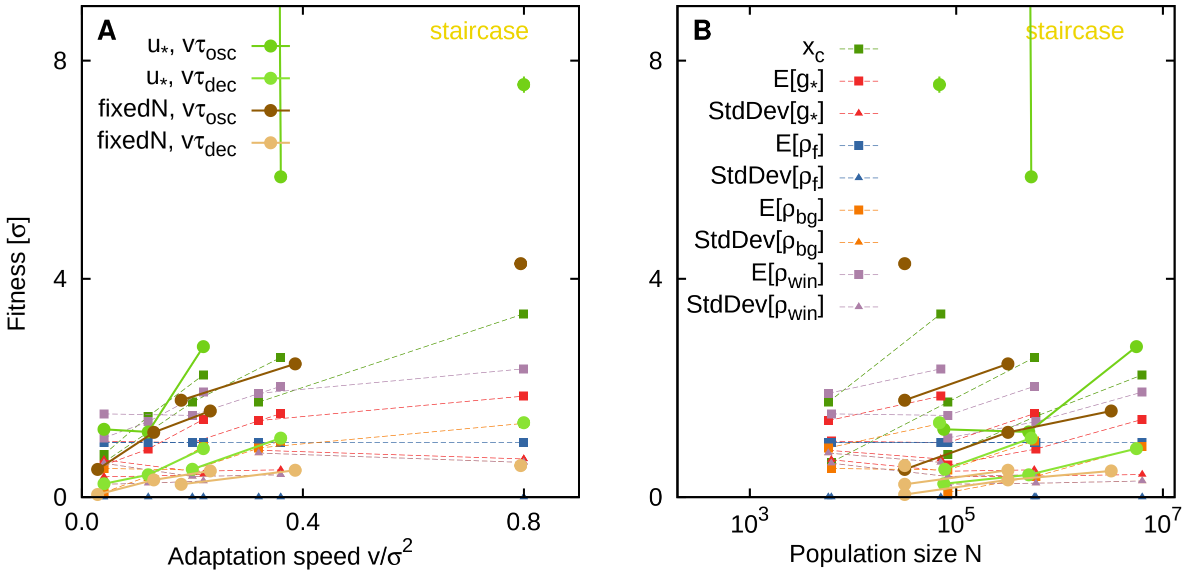

Figure D.2.: **Fitness increase during the oscillation and decay timescales, $v\tau_{osc}$ and $v\tau_{dec}$, compared to various characteristic fitness scales of distributions in tuned models with the staircase mutation kernel.** Linked points have identical mutation rates. From top to bottom, values are $\mu = 10^{-4}$, $\mu = 10^{-5}$ and $\mu = 10^{-6}$. The mean mutational effect is $\sigma = 10^{-3}$. Data in (A) and (B) are from identical simulations and identical numerical profiles, but shown with a different dependence on simulation parameters. The profiles $\mathbf{u}_\star$, $\overline{\mathbf{c}_\star}$ and $\mathbf{g}_\star$ are shown in Figure 5.6. Simulation data on the adaptation speed was presented in Figure 4.3. The autocorrelation function of fitness variance is exhibited in Figure 5.3, together with the two time scales τ_{osc} and τ_{dec} extracted from its shape.

and only the second order term is kept. However, recall that in the diffusion mutation kernel, the limit $\mu/\sigma \gg 1$ is treated. We can state that a huge amount of almost neutral mutations enter the population. Thus, we use the limit $\mathbf{m} \to \boldsymbol{\delta}$: the mutation kernel converges approximately to a Dirac-delta centered at $x = 0$. Then, the distributions simplify to

$$\begin{aligned} \mathbf{g}_\star &= \mathbf{u}_\star \overline{\mathbf{c}_\star}\,, && \text{(D.8a)} \\ \boldsymbol{\rho}_f &\approx \boldsymbol{\delta}\,, && \text{(D.8b)} \\ \boldsymbol{\rho}_{bg} &\approx \mathbf{g}_\star\,, && \text{(D.8c)} \\ \boldsymbol{\rho}_{win} &\approx \mathbf{g}_\star\,. && \text{(D.8d)} \end{aligned}$$

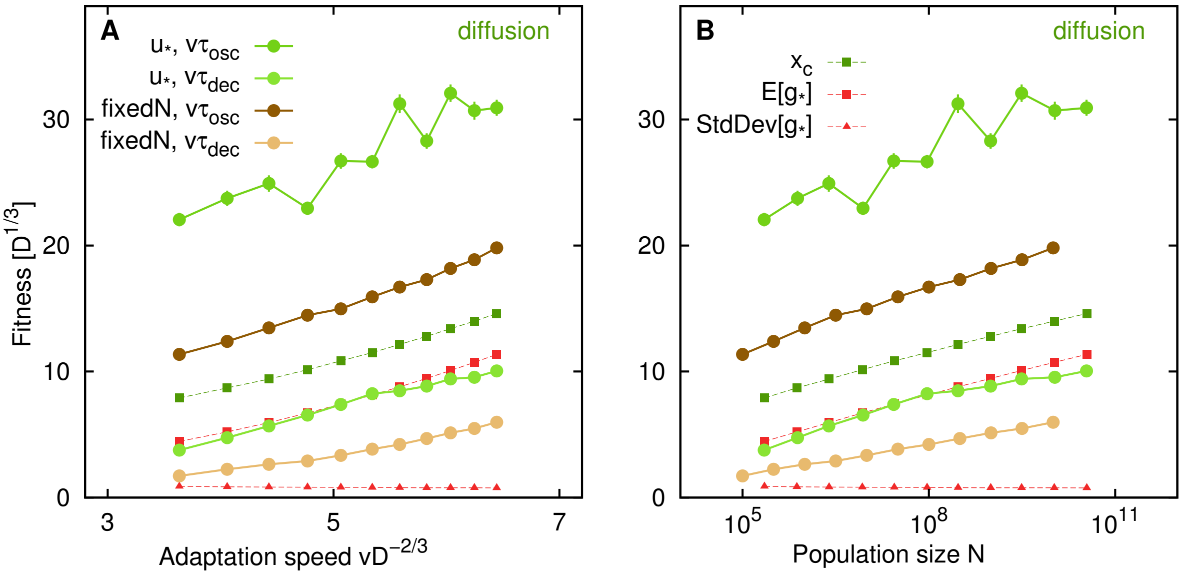

Figure D.3.: **Fitness increase during the oscillation and decay timescales, $v\tau_{osc}$ and $v\tau_{dec}$, compared to various characteristic fitness scales of distributions in tuned models in the diffusion mutation kernel.** The diffusion constant was $D = 10^{-9}$. Data in (A) and (B) are from identical simulations and identical numerical profiles, but shown with a different dependence on simulation parameters. The profiles $\mathbf{u}_\star$, $\overline{\mathbf{c}_\star}$ and $\mathbf{g}_\star$ is shown in Figure 5.7. Simulation data on the adaptation speed was presented in Figure 4.3. The autocorrelation function of fitness variance is exhibited in Figure 5.3, together with the two time scales τ_{osc} and τ_{dec} extracted from its shape.

D.3. Evaluation of fitness scales

Several observations are glaring when inspecting Figures D.1, D.2 and D.3. We could not distinguish a single characteristic fitness scale that corresponds undoubtedly to timescales in the ACF for all mutation kernels. However, the fitness increase during the decay time τ_{dec} or during the oscillation time τ_{osc} is clearly in the same range as characteristic scales of all profiles. Simulation data is still quite noisy and extensive simulations might result in better agreement.

It should be noted that in tuned models the timescales for decay or oscillations is considerably larger than in the fixed population size model. While the fitness scales in the distributions should correspond rather to the simulation results from tuned models, they often appear be in the range of fixed-N simulations.

Only in the diffusion mutation kernel we could detect a clear trend. The fitness increase during one oscillation period, as well as the fitness increase during the decay of the ACF, can be captured in the relations

$$\left(\frac{v\tau_{dec}}{D^{1/3}}\right)^{\text{diffusion}} \approx \alpha \log\left(D^{1/3}\,\overline{N}\right)\ , \tag{D.9a}$$

$$\left(\frac{v\tau_{osc}}{D^{1/3}}\right)^{\text{diffusion}} \approx 2\alpha \log\left(D^{1/3}\,\overline{N}\right) + \beta\ . \tag{D.9b}$$

The numerical coefficients were estimated as $\alpha \approx 0.39$ and $\beta \approx 8.1$.

Even if our comparison is not conclusive, we conjecture that the adaptation process depends on the properties of the *distribution of fixating clones* $\mathbf{g}_\star$. All other distributions depend on effects caused by a *single* mutation. However, in the clonal competition regime of adaptation, one clone has most likely accumulated several beneficial mutations. The observed large scale oscillations in fitness variance involve most likely multiple mutations.

E. Experimental data

E.1. Additional tests to examine experimental data

In the main text (*cf.* section 5.4), we estimated the oscillation time τ_{osc} from published data in the Lenski experiment [WRL13b, Len15]. However, the interval between consecutive fitness measurements is not constant over the complete time of the datasets. At the beginning of the time-series considerably more measurements were made (or at least published).

We used the minimal spacing between data points that spanned the whole time series to calculate the ACF. This resulted in the intervals $\Delta t = 500$ for the shorter dataset over the first 10000 generations [Len15], and the interval $\Delta t = 2000$ for the longer dataset over the first 50000 generations [WRL13b].

In this section, we investigate the impact of choosing different intervals for calculating the ACF of adaptation speed. Thus, we either cannot use the whole time series, or the underlying data is really sparse (and probably not completely reliable anymore). The main goal is to try to corroborate the existence of the first anti-correlated (negative) peak in the ACF. We ultimately used this first anti-peak to estimate a full oscillation in fitness variance (or, equivalently, adaptation speed): at this minimum, half of the time has passed. However, this estimate is heavily influenced by the choice of the interval we use to calculate the ACF.

In Figure E.1 we present ACF of adaptation speed calculated for additional intervals. As in the main text, we first calculated the adaptation speed as difference in fitness increase over the difference in time for *each* data point. Then, we used only values of the adaptation speed with the correct interval and discard the rest. In the first dataset we checked the intervals $\Delta t = 500, 1000, 2000$. Note that for the last point ($T = 8000$, $\Delta t = 2000$) in Figure E.1A only two values are used in the average, such that results are probably not reliable. The ACF with the interval $\Delta t = 1000$ is similar to our original choice $\Delta t = 500$, inasmuch as the first data point is also anti-correlated and all others scatter around zero. Hence, we consider the additional analysis of the shorter dataset not very conclusive: but note that we want to investigate a highly stochastic phenomenon that takes roughly 1000 generations in a dataset spanning only 10000 generations. More data is needed to corroborate the effect of oscillating fitness variance.

This additional data might be found in the longer dataset over the first 50000 generations. Up to generation 5000, the data intervals are 500 generations and up to generation 20000, fitness has been measured every 1000 generations. Only from generation 20000 to generation 50000 the interval is only 2000 generations. In Figure E.1B we present the ACF of adaptation speed for the intervals $\Delta t = 500, 1000, 2000, 4000$. Except for the

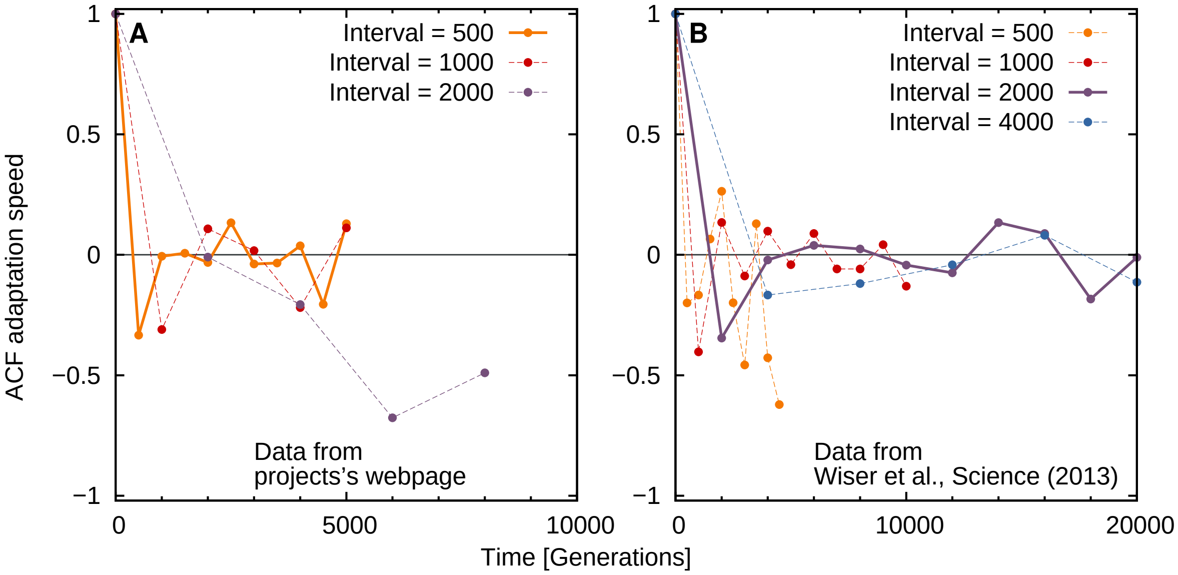

Figure E.1.: **Different intervals for calculating the autocorrelation function of adaptation speed in the datasets of the Lenksi experiment.** In the main text, we presented the ACF of adaptation speed in Figure 5.8. There we used $\Delta t = 500$ for the data from the project's webpage [Len15] and for $\Delta t = 2000$ for the extended dataset from a recent publication [WRL13b] (both highlighted as solid lines). Here, we show additional time intervals. When increasing the interval, the amplitude of anti-correlation at the first interval starts to decrease. The result is not entirely conclusive, but hints in the direction that the anti-correlation is a genuine feature of the ACF of adaptation speed: for the smallest interval ($\Delta t = 500$) in (B), the first *two* points are negative, indicating that oscillations on a timescale of the order of a few thousand generation are real. In the main text we estimated $\tau_{osc} \approx 1000$ for the data in A and $\tau_{osc} \approx 4000$ for the data shown in B.

largest interval, the first peak is discernibly anti-correlated. However, for the smaller intervals, we also expect an additional effect: data for these intervals are only from the beginning of the experiment, where adaptation proceeds faster. Thus, also the oscillations should be faster. Note, however, that for the smallest interval $\Delta t = 500$ the first *two* data points are negative, such that we can assume the oscillations with a period of roughly $\tau_{osc} \approx 1000 - 1500$ to be genuine. This value also agrees with the one obtain from the smaller dataset from the webpage [Len15].

Acknowledgments

The final effort to write a thesis is left to oneself, but *many* people shape and influence the process of its creation via suggestions, discussions, comments or support. Here is the place to thank them.

The largest share of the credit goes of course to my supervisor Oskar Hallatschek. Many ideas presented here were originally seeded by him and evolved to their final form as written in this thesis over the course of many stimulating, beneficial and interesting discussions.

I am particularly indebted to Albert, Bernhard, Jean-Francois and Jens for providing uncountable and valuable comments on various incomplete versions of this thesis.

A great deal of the contents of this thesis has been discussed with the members of the *Biological Physics and Evolutionary Dynamics* group. Thanks to all its current and past members for a fantastic time, in particular Christopher, Clemens, Erik, Fabian, Hedvika, Jean-Francois, Jens, Jörn, Marc, Marijn, Matthias, Nils, Oskar, Paulo and Sven.

I am very grateful for the nice and welcoming atmosphere at the *Max Plank Institute for Dynamics and Self-Organization*. After the experience here, it is hard to believe that there are other places with such a high concentration of delightful people. The support provided by Barbara was invalueable for all those little bureaucratic issues at the institute.

I really enjoyed organizing the GOEvol symposia and appreciate all the great times during and also outside the organization. Thanks to Alex, Alex, Gisela, Ines, Jean-Francois, Jens, Jule and Nico.

There have been too many grand people with whom I have interacted over the last few years here in Göttingen, therefore it is hard to write a comprehensive list of names, as it will always be too short. I thank you all for a really wonderful time here in Göttingen.

Moreover, I want to thank the DFG Schwerpunktsprogramm "Probabilistic Structures in Evolution" (DFG SPP 1590) for funding, travel grants and enjoyable hospitality at various conferences and meetings.

Finally, I want to thank my parents Karin and Manfred for their continuous support throughout the years. Without them, all this would not have been possible.

Curriculum Vitae

Name	**Lukas Geyrhofer**
born on	December, 29th 1983

Academic Education

2003 - 2007	**Food Science and Biotechnology** (BSc. equivalent, *Bakk. techn.*) University of Natural Resources and Life Sciences, Vienna
2004 - 2010	**Physics** (Diploma, *Mag. rer. nat*) University of Vienna
2011 - 2014	**Physics** (PhD Candidate) MPI for Dynamics and Self-Organization, and University of Göttingen

Publications

Tscheliessnig R, Geyrhofer L, Wendland M, and Fischer J. Adsorption from Oversaturated Aqueous Solution: Mean Force Molecular Simulations. *AIChE Journal* **54**(9):pp. 2479-2486. 2009

Zagordi O, Geyrhofer L, Roth V, and Beerenwinkel N. Deep Sequencing of a Genetically Heterogeneous Sample: Local Haplotype Reconstruction and Read Error Correction. *Journal of Computational Biology*, **17**(3):pp. 417-428. 2010

Geyrhofer L, and Hallatschek O. Stochastic delocalization of finite populations. *Journal of Statistical Mechanics: Theory and Experiment*, **2013**(01):p. P01007. 2013