Emergence of Antibiotic Resistance in Microbial Populations with Spatial Heterogeneity

by

Maxwell G. De Jong

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in the University of Michigan
2021

Doctoral Committee:

Professor Kevin Wood, Chair
Professor Charles Doering
Professor Qiong Yang
Associate Professor Luis Zaman
Professor Michal Zochowski

Maxwell G. De Jong

mgdejong@umich.edu

ORCID iD: 0000-0001-9088-664X

DEDICATION

I am dedicating this thesis to my family. First and foremost, to my parents, Mel and Chet, who have unconditionally supported and encouraged my educational pursuits since my first day of preschool. Next, to my wife Gracia, who has been a bright spot during many long research days. And finally, to our beloved son Jack—may all of my work be trivial to you.

ACKNOWLEDGMENTS

If it truly takes a village to raise a child, I can never hope to thank everyone who has positively impacted me as my formal educational journey comes to end. Undoubtedly, I am the product of many outstanding teachers and educators ranging from elementary school to graduate school as well as numerous supportive colleagues across this span of time. Although I cannot express gratitude to all of the people responsible for my dissertation, there are several individuals not yet named whom I would be remiss to not mention.

While completing my undergraduate education at Caltech, I was exceedingly fortunate to have met a handful of very intelligent lifelong friends: Cedric Flamant, Ellen Price, Jacob Shenker, and Charles Tschirhart. I learned more from them than any professor or textbook while at Caltech, and I would not have undertaken graduate school without everything that I gained from them.

Upon coming to Michigan, I stumbled into Kevin Wood's office looking for an opportunity to apply statistical physics to biological problems. While I did receive this opportunity, I had fortuitously found a much more valauble opportunity—the chance to work with a brilliant scientist, a patient teacher, and a mentor of exemplary character. I benefited tremendously from his mathematical acumen, scientific intuition, and collaborative approach to research during our time together.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF APPENDICES

ABSTRACT

The emergence of antibiotic resistance in microbial populations is a complex evolutionary process dependent on a large number of factors, both mechanistic and external. We have modeled spatial heterogeneity as one such external factor and have found that it can greatly affect the times to resistance. We formed various models governed by simple atomistic rules in which antibiotic resistance can be treated as an emergent phenomenon and studied using classical tools from statistical mechanics, including a classical mean first passage time calculation and a kinetic Monte Carlo simulation. We have found that spatial heterogeneity can speed or slow the emergence of resistance, depending on the specific system and the relative rates of mutation and migration in spatially-extended systems. Finally, we developed a robust Bayesian inference pipeline to precisely estimate the relevant parameters from turbidostat evolution experiments. We believe that this work establishes environmental spatial heterogeneity as a critical factor in understanding the emergence of antibiotic resistance and may the groundwork for drug dosing strategies spatially-optimized to slow the emergence of resistance.

CHAPTER 1

Introduction

Many of the medical accomplishments from the last century are due to the successful clinical application of antibiotics to treat infections. Since the modern era of antibiotics began with the discovery of penicillin in 1928, antibiotics have transformed society and become ubiquitous in medicine [1, 2]. However, this dependence on drugs has introduced significant evolutionary pressure on the emergence of drug resistance, a rapidly growing public health threat and a central impediment to the continued treatment of cancer, viruses, and microbial infections [3, 4, 5, 6]. Although bacteria naturally acquire resistance to antibiotics through evolution, this process has been greatly accelerated by our modern reliance on a relatively small number of antibiotics [7]. Unfortunately, the solution to the antibiotic crisis is not as simple as expecting new antibiotics from the pharmaceutical and biotechnology industries. Due to the cost and time required to develop new drugs, the antibiotic pipeline will soon not be able to keep pace with new resistant strains [8, 9]. Antibiotic resistance is a multi-faceted challenge with global medical, social, and economic impacts that will require solutions on many levels, including novel scientific discoveries, thoughtful public health initiatives, increased public education, and directed policy work [7, 8, 10, 11].

One of the basic scientific goals is to understand how resistance develops in bacteria. The past decade has seen tremendous advances in understanding the molecular mechanisms through which resistance develops as a result of breakthroughs in systems biology, genomics, and structural biology [12, 13, 14, 15]. Certainly, understanding these varied molecular mechanisms governing both spontanteous mutation and horizontal gene transfer will be critical for controlling antibiotic resistance. However, new resistance mechanisms are constantly being discovered, preventing definitive mechanism cataloging. Additionally, resistance mechanisms have already been discovered for all current antibiotics [12]. And finally, there are significant challenges with translating each new mechanism into a pharmaceutical solution [16, 17]. So although understanding the molecular mechanisms of resistance is necessary, this is not sufficient for a complete understanding of antibiotic resistance.

Fundamentally, resistance is an emergent phenomenon that arises from stochastic processes governed by the complex interplay between microbial ecology and evolutionary selection. Follow-

1

ing this philosophy, many recent developments have focused on evolution-based treatment strategies combining pathogen genotyping with evolutionary genetics to slow resistance emergence at the genotype-level [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29]. This reliance on evolutionary genetics requires a mature understanding of the evolutionary pathways taken though genotype-space to implement such strategies. In addition to the pathways themselves, a complete understanding of the external factors shaping the evolution of the system is required. The environment has long been recognized as impacting the dynamics of pathogens [30]. Although much of the existing work has focused on evolution in spatially-homogeneous environments, a number of recent experimental and theoretical studies have demonstrated that spatial heterogeneity in drug concentration can significantly alter the evolutionary dynamics leading to resistance [31, 32, 33, 34, 35, 36, 37, 38, 29].

In a natural or even clinical setting, evolution takes place in environments characterized by spatial fluctuations in drug concentration, nutrients, temperature, pH, and other factors, all of which can potentially impact the growth rates of multiple gentoypes. As one example, there are many clinial motivations for drug concentration varying spatially [39, 40], such as heterogeneous drug permeability in biofilms [41] and tumor stromas [42]. Understanding evolution and ecology in such spatially extended systems is a challenging and long-studied problem spanning many fields [43, 44, 45, 46, 47]. Early work investigating spatial structure came from population genetics [48], and today this problem most neatly fits under the umbrella of landscape ecology, which aims to combine biotic and abiotic factors of an ecosystem to better understand how resistance collectively emerges [49]. However, the emergence of resistance in spatially-extended systems continues to be a multi-disciplinary problem benefiting from a number of fields. In recent years, theoretical tools orginating from statistical physics have proven valuable for understanding the spatiotemporal evolutionary dynamics in spatially-extended systems [50, 51, 52], including dispersion models [53, 54, 55, 56, 57], and Moran meta-population models [58, 59, 60].

In this work, we continue to extend existing statistical physics approaches to studying the emergence of antibiotic resistance in spatially-structured communities. We focus on understanding the impact of spatial profiles of selection pressure on the emergence of resistance because unlike the genotypic fitness landscape, which is an inherent property of each cell, selection pressure is an external tunable profile that we can hope to modulate spatially. Previous work has already established the importance of understanding the role of spatial heterogeneity in the form of spatial selection pressure landscapes on evolutionary dynamics, but a number of important questions remain unanswered. Both experimental and theoretical work have established that spatial profiles of selection pressure induced from drug concentration gradients affect evolutionary dynamics [36, 38, 32], but the variation in times to resistance across selection pressure profiles conserving the spatially-averaged selection pressure is not well-understood. Using simple models inspired by results from evolution, population dynamics, and evolutionary game theory, we seek to more fully understand

the effect of spatial distribution of selection pressure on the resistance emergence without the added confounding variable of total selection pressure. Additionally, much of the existing work on understanding resistance under spatially heterogeneous selection profiles has focused on simple gradients. We are interested in understanding the evolutionary dynamics under more complex and potentially non-monotonic spatial selection profiles.

Methodologically, we aim to answer these questions using simple models of evolution that capture fundamental relationships between evolutionary dynamics and spatial selection profiles. These models represent huge simplifications from the underlying biology, neglecting many considerations regarding mutational pathways, drug mechanisms, and drug-cell interations [61]. Although the importance of these factors is well-established, they are often specific to both the microbial species and drug of interest. Motivated by recent successes of low fidelity models to uncover fundamental relationships governing evolution independent of these important factors [62, 63], we hope to use these models to reveal general principles of spatial heterogeneity that are not specific to any particular mechanism or drug class. Instead of merely relying on hope, however, we focus on generating interpretable results that can be intuitively understood in general systems and using these models to craft experimentally-verifiable predictions. We believe that this approach maximizes the utility of toy models of evolution to inform a more general understanding of the relationship between spatial heterogeneity and antibiotic resistance.

Our ultimate goal is to not only have a more mature understanding of antibiotic resistance in spatially-structure populations but also to leverage this understanding to slow the emergence of resistance using spatial heterogeneity. Much of the current work exploring the effects of heterogeneity has highlighted its potential to rapidly increase the emergence of resistance. Powerful experimental illustrations of rapid onset of resistance have greatly increased awareness of the importance of spatial heterogeneity introduced from simple drug gradients [36, 38]. However, little work has been performed to demonstrate the ability of spatial heterogeneity to lead to slower resistance in spatially-structured populations. Spatial profiles of selection could be a powerful tool against the emergence of resistance because such spatial profiles are external to the molecular and genomic properties of the cell and relatively easily modulated. One practical application of a more mature understanding of the impact of spatial heterogeneity on the emergence of resistance would be spatially-optimized drug dosing strategies, in which the drug concentration or other environmental factor is modulated spatially to create selection profiles that slow down the emergence of resistance across the entire population.

CHAPTER 2

Tuning Spatial Profiles of Selection Pressure to Modulate Resistance Emergence in a Simplified System

2.1 Introduction

The battle against drug resistance has been largely fought at the molecular level, leading to an increasingly mature understanding of its underlying biochemical and genetic roots. At the same time, evolutionary biologists have long recognized resistance as a fundamentally stochastic process governed by the complex interplay between microbial ecology and evolutionary selection. The last decade, in particular, has seen a significant surge in efforts to develop and understand evolution-based treatment strategies to forestall resistance [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29]. While the vast majority of this work focuses on spatially homogeneous environments, a number of recent studies, both theoretical and experimental, have demonstrated that spatial heterogeneity in drug concentration can dramatically alter the evolutionary dynamics leading to resistance [31, 32, 33, 34, 35, 36, 37, 38, 29]. On a practical level, the picture that emerges is somewhat bleak, as resistance evolution is dramatically accelerated in the presence of spatial gradients in drug concentration [32, 33, 34, 36, 37, 38] or heterogeneous drug penetration [31, 35]. Interestingly, however, recent work shows that this acceleration can be tempered or even reversed when the mutational pathway (i.e., the genotypic fitness landscape) leading to resistance contains fitness valleys [32], which are known to inhibit evolution [64, 65, 66, 67]. Unfortunately, because the fitness landscape is a genetic property of the cells themselves, the potential for accelerated evolution appears to be "built in", making it difficult to combat in a treatment setting. However, these results raise the question of whether non-monotonic profiles of tunable properties of the system—for example, the spatial selection pressure—might also have the potential to slow evolution, even when the mutational pathway lacks the requisite fitness valleys.

Evolution in natural or clinical settings takes place in heterogeneous environments characterized

by spatial fluctuations in multiple factors, including drug concentrations, nutrients, temperature, and pH, all of which potentially affect cellular growth. Recent studies investigating evolution and ecology in spatially extended systems have demonstrated rich dynamics when inter-cellular interactions are defined on heterogeneous complex networks [68, 69, 70], where spatial structure can (for example) promote invasive strategies in tumor models [69] or modulate fixation times on random landscapes [68]. Remarkably, in the weak selection limit, evolutionary dynamics can be solved for any population structure [70], providing extensive insight into game-theoretic outcomes on complex networks. In addition, theoretical tools from statistical physics have proven useful for understanding spatiotemporal dynamics in spatially structured populations in a wide range of contexts, including biologically-inspired Monte Carlo models [32], toy models of source-sink dynamics [33], stepping-stone models of spatial pattern formation [71], models of dispersion [53, 54, 55, 56, 57], and Moran meta-population models [58, 59, 60].

Motivated by the successful application of statistical physics to understanding antibiotic resistance evolution, We investigated the emergence of resistance in environments with spatial heterogeneity using tools more commonly used in statistical physics. We have constructed simple models of evolution that aim to teach us about the fundamental role of the environment on the specific evolutionary trajectory taken. These toy models of evolution represent huge simplifications from the true underlying biology, as evolution is a complex phenomenon depending on a large number of pathways, mechanisms, and drug-cell interactions unique to each species and drug under consideration [61]. However, by treating these systems in generality, we can hope to understand fundamental relationships that govern evolution independent of these other, important factors [62, 63]. In this work, we rely on simple models inspired from fields such as evolution, ecology, population dynamics, and evolutionary game theory, and seek to use them to uncover fundamental relationships between the emergence of evolution and the underlying environment.

2.2 Developing a Simple Model of Evolution

2.2.1 Calculating Mean Fixation Times

Motivated by the previous success in applying statistical physics to antibiotic resistance, here we use stochastic models of evolution along with theoretical tools from statistical physics to investigate the effects of spatially heterogeneous fitness pressure on the evolution of resistance. In contrast to previous models defined on heterogeneous networks at the single-cell level, here we consider meta-populations connected via simple topologies and investigate the effects of spatial structure imposed by arbitrary distributions of selection pressure.

To investigate resistance evolution on a spatially heterogeneous landscape, we consider a
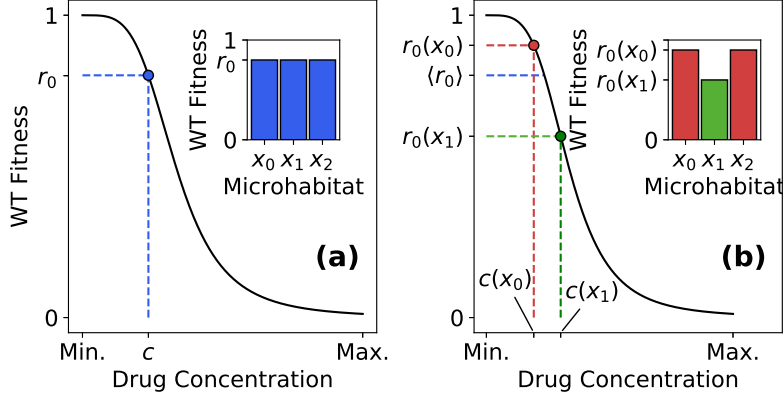
Figure 2.1: (a) The drug concentration c is chosen to yield to a desired fitness r_0 using to the wild-type dose-response curve. Inset: resulting homogeneous fitness landscape. (b) With spatial heterogeneity the drug concentrations $c(x_0)$ and $c(x_1)$ are chosen to yield fitness values $r_0(x_0)$ and $r_0(x_1)$, respectively. Inset: resulting fitness landscape with spatial fitness valley.

stochastic Moran-like model [72] of a finite population (N) consisting of ($N - n^*$) wild-type cells with fitness $r_0 \leq 1$ and n^* drug-resistant mutants with fitness r^*, which we set to unity without loss of generality. Note that this model does not include a fitness cost to resistance (i.e., $r^* \geq r_0$ for all conditions). We motivate our interest in this problem with a simple thought experiment illustrated in Figure 2.1. One can imagine recording a dose-response curve for wild-type cells sensitive to some drug present that allows us to select the drug concentration c to achieve the desired fitness r_0 in all of the microhabitats (Figure 2.1a). One could also allow each microhabitat to have its own drug concentration. In Figure 2.1b, an example is shown with spatially-varying drug concentrations $c(x_0) = c(x_2)$ in the extrema microhabitats and $c(x_1)$ in the center microhabitat. Crucially, the drug concentrations could be chosen such that resulting fitness values have the same mean fitness $\langle r_0 \rangle$ as the fitness value r_0 used in the homogeneous experiment.

To understand the effect of spatial heterogeneity in the fitness landscapes on the evolutionary trajectories in this simple system, we form a simple model of the evolutionary dynamics. At each time step cells are randomly selected for birth and death, with cells of higher fitness (in this case, resistant cells) chosen preferentially for division. Wild-type cells can mutate to become drug resistant at rate μ; we neglect reverse transitions to the drug-sensitive state. To incorporate spatial heterogeneity, we consider a simple spatially extended system with M distinct microhabitats, each containing N cells; cells are allowed to migrate at rate β between connected microhabitats. At each time t, the state of the system is characterized by the vector $n^*(x_i)$ whose components correspond to the number of mutants in each discrete microhabitat $x_i = 0, 1, \ldots, M - 1$. The system evolves according to a continuous time master equation

$$\frac{dP_m}{dt} = \sum_{m'} \Omega_{mm'} P_{m'}, \tag{2.1}$$

where m and m' denote different states of the system and Ω is a $N^M \times N^M$ matrix whose entries depend on the wild-type fitness value $r_0(x_i)$ at each spatial location x_i. For tractability, we restrict our analysis to $M = 3$, which is the simplest model that allows for potentially non-monotonic fitness landscapes, such as fitness peaks and fitness valleys. In what follows, we refer to the vector $s(x_i) \equiv 1 - r_0(x_i)$ as the spatial profile of selection pressure, as it measures the difference in fitness between resistant and wild-type cells in each microhabitat x_i. Intuitively, large values of $s(x_i)$ correspond to regions where the resistant mutant has a significant evolutionary advantage over the wild-type cells (e.g., regions of high drug concentration).

Although Equation 2.1 is difficult to solve explicitly, the stationary distribution is actually trivial. Because the model has a single absorbing state—the fully resistant state ($n^*(x_i) = N$ for all x_i)—the system will asymptotically approach the fully resistant state. However, other quantities describing the evolution of resistance in various spatial profiles are straightforward to calculate and also more insightful. We characterize the speed of fixation in the presence of different spatial profiles $s(x_i)$ by calculating mean first passage times (MFTPs) between states, which obey [73, 74]

$$-1 = \sum_{m' \neq m_f} T(m_f|m') \Omega_{m',m_i}, \tag{2.2}$$

where $T(m_f|m_i)$ is the mean time required for a system initially in state m_i to first reach state m_f. While several elegant approaches exist for studying these models in particular limits (e.g., with a center manifold reduction) [58, 59, 60], here we instead use a classical mean first passage time approach based on adjoint equations to reduce the calculation of mean fixation times to a simple collection of linear equations that can be easily solved for arbitrary spatial distributions of selection pressure. This method also allows us to find the fixation times from arbitrary initial states, which are often difficult to compute using other methods. We take m_f to be the fully resistant state and solve the coupled set of linear equations for $\tau_f^j \equiv T(m_f|j)$, where j is an index that runs over all initial states. In particular, when j is the fully wild-type population ($n^*(x_i) = 0$ for all x_i), we refer to the MFPT as the mean fixation time τ_f.

Before we can characterize τ_f for different systems, we need to specify a form for the transition rates populating Ω. Consider, as an example, the four processes that lead to an increase in the number of mutants in the center microhabitat: a wild-type cell can mutate, an existing mutant can replicate, a mutate can migrate from its left-neighboring microhabitat, and a mutant can migrate from its right-neighboring microhabitat. Mathematically, the transition rate for increasing the

7
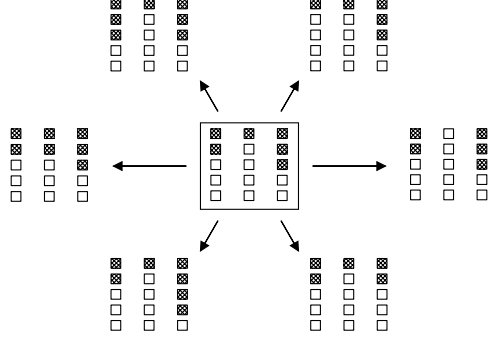
Figure 2.2: An illustration of the new states that can be reached in a single timestep by a general state of the system, shown in the center. Due to the assumption of a Moran-like process, the total number of cells of either type can change by at most one. The transition rate in Equation (2.3) governs the transition to the state directly to the left of the initial state if checkered boxes represent mutant cells.

number of mutants in microhabitat x_1 given some initial state (j, k, ℓ) takes the form

$$
\begin{aligned}
T^+[x_1](j, k, \ell) = {} & (N - k) \underbrace{\frac{(N - k)w(x_1)r_0(x_1)}{(N - k)r_0(x_1) + kr^*(x_1)}\mu}_{\text{WT mutates}} \\
& + (N - k) \underbrace{\frac{kw(x_1)r^*(x_1)}{(N - k)r_0(x_1) + kr^*(x_1)}(1 - 2\beta)}_{\text{Mutant Replicates}} \\
& + (N - k) \underbrace{\frac{jw(x_0)r^*(x_0)}{(N - j)r_0(x_0) + jr^*(x_0)}\beta}_{\text{Mutant Migration}} \\
& + (N - k) \underbrace{\frac{\ell w(x_2)r^*(x_2)}{(N - \ell)r_0(x_2) + \ell r^*(x_2)}\beta,}_{\text{Mutant Migration}} \quad (2.3)
\end{aligned}
$$

where $w(x_i)$ is a microhabitat-specific weighting parameter that accounts for differences in the global fitness across habitats. We can enumerate over these different mechanisms to form a complete set of transition rates for our system. Because the number of mutants can only change by at most one in a single microhabitat, there exist few new states that can be reached from a general initial state of the system, shown schematically in Figure 2.2, which additionally illustrates the sparsity in Ω. The complete set of transition rates can be found in Appendix A.1.

Under these conditions, one could weight each microhabitat by its relative fitness (i.e., mean fitness in the habitat normalized by the mean fitness of the entire multi-habitat system), reflecting the fact that more division events occur in fast-growing habitats. This weighting scheme could be

8
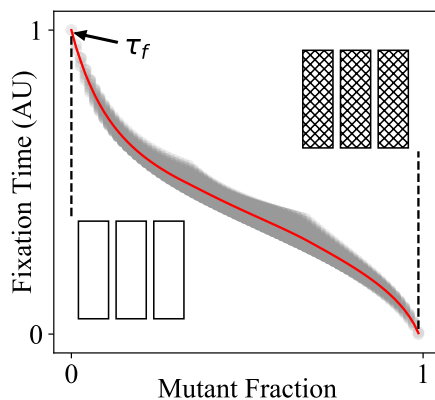
Figure 2.3: The time to fixation can be calculated by solving the coupled MFPT equations from every unique initial state of the system. Of primary interest is τ_f, defined as the time to fixation with no initial mutants.

particularly relevant if the global fitness in each habitat is expected to vary significantly–for example, if the different demes were held at different temperatures. In what follows, we choose $w_{x_i} = 1$ for simplicity, though we have found qualitatively similar results using more complex weighting schemes (see Appendix A.1.1). With the transition rates enumerated, Ω can be formed and the MFPTs can be solved according to Equation 2.2. As an illustration of the MFPT calculation, the full fixation time vector $T((N, N, N)|(j, k, \ell))$ for initial state (j, k, ℓ) using a particular parameter choice is shown in Figure 2.3. Although not all states with a given mutant fraction take the same amount of time to reach fixation, the fixation time generally decreases with additional mutants as expected. Our primary interest is understanding the fixation time τ_f, the single time to fixation from an initial mutant fraction of zero.

2.3   Modulating Fixation Times with Spatial Heterogeneity

In the case of a single microhabitat, the mean fixation time τ_f increases as selection pressure decreases. In the spatially extended case, τ_f would also be expected to increase when the selection pressure is globally decreased, though it should also depend on the spatial structure of the specific selection profile $s(x_i)$. To investigate the impact of spatial structure alone, we compared τ_f across different selection profiles $s(x_i)$, all of which were characterized by the same spatially averaged selection pressure, $\langle s \rangle = \sum_i s(x_i)/M$. For simplicity, we begin with a symmetric profile characterized by a background selection pressure s_0 in the edge habitats and a relative peak of height δs in the center habitat (Figure 2.4a). This toy landscape has an average selection pressure of $\langle s \rangle = s_0 + \delta s/M$, and the parameters s_0 and δs are constrained by the fact that $0 \leq s(x_i) \leq 1$ at
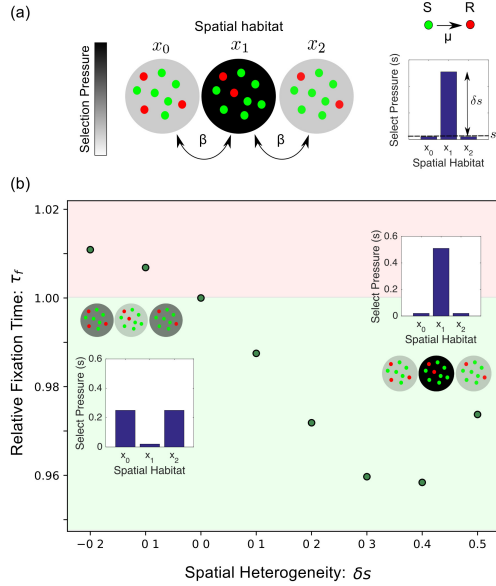
Figure 2.4: (a) Stochastic model for emergence and spread of resistant cells (red) in a spatially extended population of sensitive cells (green). Each spatial habitat (x_i) contains N total cells. Cells migrate at a rate β between neighboring habitats, and sensitive cells mutate at a rate μ to resistant cells. The spatial distribution of selection pressure is characterized by a background value (s_0) and a peak height (δs). (b) Example plot of the mean fixation time for different landscapes with $\mu = 5 \times 10^{-3}$, $\beta = 0.08$, $N = 25$, and $\langle s \rangle = 0.167$. The time to fixation can be either faster (green) or slower (red) than the spatially homogeneous landscape with $\delta s = 0$. Inset: selection landscapes for $\delta s = -0.2$ and $\delta s = 0.5$.

all spatial locations. We vary δs systematically to explore different selection profiles, which can include a single selection pressure valley ($\delta s < 0$), a homogeneous landscape ($\delta s = 0$), or a single selection pressure peak ($\delta s > 0$).

Interestingly, we find that modulating heterogeneity (δs) can increase or decrease τ_f for certain choices of migration and mutation rates, even when $\langle s \rangle$ is held constant (Figure 2.4b). More generally, we find that the $\beta - \mu$ plane can be divided into three non-overlapping regions where the homogeneous landscape 1) leads to the smallest value of τ_f, 2) leads to the largest value of τ_f, or 3) does not correspond to an extremum τ_f (Figure 2.5a-b). In the latter region, heterogeneity often modulates the fixation time by only a few percent, but we do find larger effects in the high and low migration limits (i.e., on the edges) of the intermediate regime (see Appendix A.2). In addition, as we increase β for a fixed value of μ, τ_f smoothly transitions from being minimized at $\delta s = 0$ to being maximized near $\delta s = 0$. We find empirically that the fixation time can be dominated by τ_1, the time required to achieve a small population of mutants (Figure 2.5c, rightmost panel) or τ_2, the time required for this small population to achieve fixation (Figure 2.5c, leftmost panel). However, in many cases—particularly those close to the intermediate region where fixation can

10
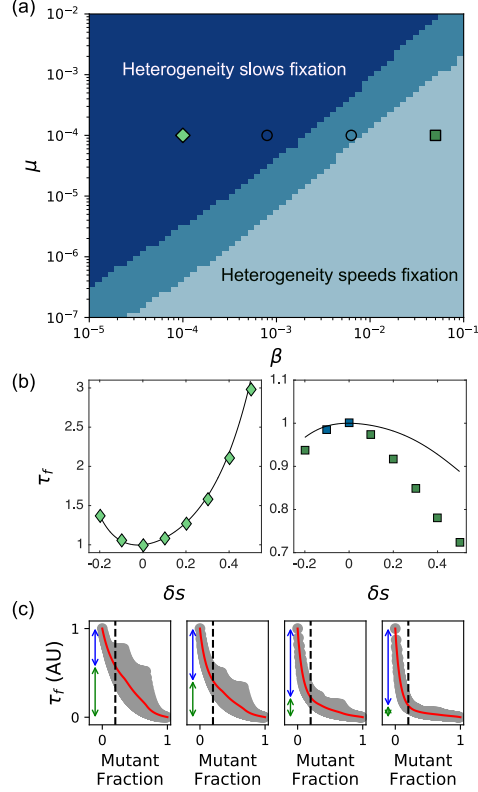
Figure 2.5: Spatial heterogeneity can speed or slow fixation depending on the rates of migration (β) and mutation (μ). (a) Phase diagram illustrates region of parameter space where the homogeneous landscape leads to a maximum (light blue), minimum (dark blue) or intermediate (medium blue) value of the in fixation time. MFPT calculations were performed for the indicated values of β and μ and for $-0.2 \leq \delta s \leq 0.5$ in steps of 0.1. (b) Sample fixation curves in the regions where heterogeneity slows fixation (left panel, diamonds; $\beta = 10^{-4}$, $\mu = 10^{-4}$) or accelerates fixation (right panel, squares; $\beta = 5 \times 10^{-2}$, $\mu = 10^{-4}$). Solid curves indicate analytical approximations. (c) Gray shaded region indicates fixation time τ_f from every initial state $(n^*(x_0), n^*(x_1), n^*(x_2))$, where $n^*(x_i)$ is the initial number of mutants at position x_i. Red curves show mean fixation time over all initial states with a given total mutant fraction. Vertical arrows represent time to achieve a total mutant fraction of $1/5$ (τ_1, blue) and time to go from that fraction to fixation (τ_2, green). Left to right panels: increasing β at a fixed value of $\mu = 10^{-4}$; plots correspond to symbols on phase diagram in panel (a). $N = 25$ and $\langle s \rangle = 0.167$ in all panels.

be accelerated or slowed by heterogeneity—both timescales contribute to the dynamics. While we restrict ourselves primarily to $N = 25$, $\langle s \rangle = 1/6$, and to symmetric landscapes, we find qualitatively similar results (i.e., 3 distinct regions) for other values of $\langle s \rangle$ and N (Appendix A.3), as well as for alternative topologies, such as permuting the selection profiles and assuming globally connected profiles (Appendix A.4), and monotonic (gradient) selection profiles (Appendix A.5).

11

2.3.1 Analytical Approximation

To intuitively understand these results, we developed a simple analytical approximation for τ_f. In a single microhabitat with selection pressure s, the fixation time τ_f in the limit $\mu \ll 1$ can be approximated as

$$\tau_f \approx \frac{1}{\lambda(s)}, \tag{2.4}$$

where $\lambda(s)$ is the scale of an exponential process for the arrival time of the first mutant. With a single microhabitat, the only source of mutants is *de novo* mutations, so that

$$\lambda(s) = \mu N \mathcal{P}(s), \tag{2.5}$$

where $\mathcal{P}_{\text{fix}}(s)$ is the probability of a single mutant fixing. To calculate this quantity, we consider genetic drift for a state with i initial mutants and N total cells governed by transition rates

$$T^-(i) = \frac{(N-i)r_0}{(N-i)r_0 + i} \frac{i}{N}, \tag{2.6}$$

$$T^+(i) = \frac{i(1-\epsilon)}{(N-i)r_0 + i} \frac{N-i}{N}, \tag{2.7}$$

for increasing and decreasing the number of mutants, respectively. Fixation in a single microhabitat is a long-solved problem so that the probability of fixation can be solved for all states (Appendix A.6). We will restrict our analysis to the probability of fixation after a single mutation,

$$\mathcal{P}(s) = \frac{s}{1 - (1-s)^N}. \tag{2.8}$$

We extend this idea to systems with multiple microhabitats by again assuming that the arrival of the first mutant in each vial dominates the time to fixation, valid in the limit $\mu, \beta \ll 1$. Mutants can now be acquired through a *de novo* mutation event (at rate μN) or through a migration event (at rate $N\beta n_{\text{fix}}$, where n_{fix} is the number of neighboring fixated microhabitats). This is essentially a coarse-graining of the different states of the model, so that different microhabitats effectively achieve fixation sequentially. Fixation within each microhabitat is now an exponential process governed by rate

$$\lambda(s, n_{\text{fix}}) = N(\mu + \beta n_{\text{fix}})\mathcal{P}(s), \tag{2.9}$$

where $\mathcal{P}(s)$ is still given by Equation 2.8. In general, the analytical approximation for τ_f is algebraically cumbersome (Appendix A.7). However, in the limit $\beta \ll \mu$, the approximation reduces

to the expected maximum of three independent exponential random variables, leading to

$$\tau_f \approx \tau_{max} = \lambda_0^{-1} + \lambda_1^{-1} + \lambda_2^{-1} - (\lambda_0 + \lambda_1)^{-1}$$
$$- (\lambda_0 + \lambda_2)^{-1} - (\lambda_1 + \lambda_2)^{-1} + (\lambda_0 + \lambda_1 + \lambda_2)^{-1} \quad (2.10)$$

using the shorthand $\lambda_i \equiv \lambda(s(x_i), 0)$. In this limit, the three-vial system acts effectively as three independent systems, with the overall fixation time corresponding to the slowest fixation. After rewriting τ_{max} in terms of $\langle s \rangle$ and δs, it is straightforward to show that $(\partial \tau_{max}/\partial \delta s)|_{\delta s=0} = 0$ and $(\partial^2 \tau_{max}/\partial \delta s^2)|_{\delta s=0} > 0$, indicating that the homogeneous landscape ($\delta s = 0$) minimizes the fixation time, consistent with results of the exact calculation (Figure 2.5b, left panel). Intuitively, increasing heterogeneity reduces the minimum selection pressure in the spatial array, which in turn slows the expected maximum fixation time among the three habitats.

By contrast, in the limit $\mu \ll \beta$, τ_f reduces to the expected minimum of three independent exponential processes, leading to $\tau_f \approx \tau_{min} = \lambda_{\text{eff}}(s)^{-1}$, where $\lambda_{\text{eff}}(s) \equiv \lambda_0 + \lambda_1 + \lambda_2$. In this limit, the fixation time is dominated by dynamics in the vial that first achieves fixation; the remaining vials then rapidly achieve fixation due to fast migration. For large but finite N, the fixation time τ_{min} is maximized at $\delta s = 0$, indicating that heterogeneity always accelerates fixation, again consistent with the exact calculation (Figure 2.5b, right panel). In this limit, the effective rate of fixation $\lambda_{\text{eff}}(s)$ is increased for all $\delta \neq 0$, as heterogeneity decreases fixation time in the vial with the fastest average fixation.

2.4 Dosing Resistant Subpopulations

Our results indicate that a judicious choice of selection pressure profile can potentially slow fixation of *de novo* mutants. In addition, selection pressure profiles can be optimized to mitigate the effects of resistance once it has emerged. One advantage of the MFPT approach (i.e., solving Equation 2.2) is that it provides fixation times starting from all possible initial states (see Figure 2.3), making it straightforward to apply to cases where a resistant subpopulation already exists. Specifically, consider a situation where a resistant subpopulation has arisen at a particular spatial location. Is it possible to choose the spatial distribution of selection pressure—for example, by spatially dosing the drug—to minimize the time to fixation from this state? Intuitively, the goal is to delay the onset of treatment failure as long as possible. As an illustrative example, we consider a population consisting of $N/2$ mutants in the center microhabitat and calculate the mean time to fixation for different spatial profiles of selection pressure (Figure 2.6a). The specific choice of spatial profile significantly impacts the time to fixation from the initial resistant subpopulation. We then find the optimal value for δs—that is, the heterogeneity corresponding to the spatial landscape
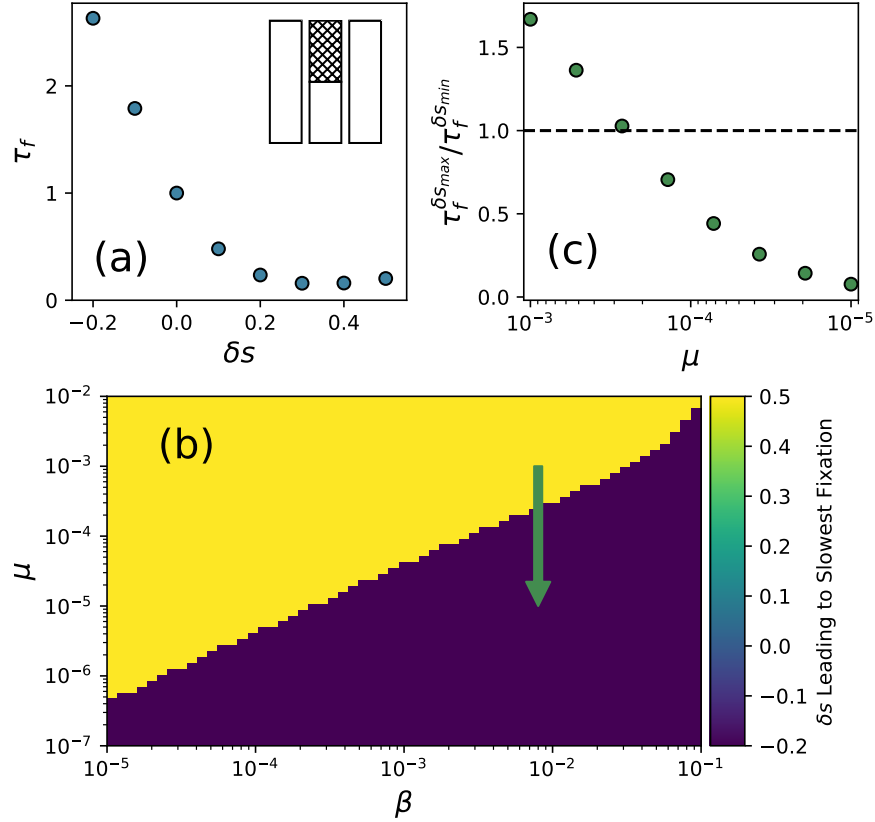
Figure 2.6: (a) Heterogeneity can significantly speed or slow fixation starting from an initial resistant subpopulation consisting of $N/2$ cells in the center habitat, shown schematically in the inset. Parameters were $\mu = 10^{-5}$, $\beta = 8 \times 10^{-3}$. (b) The optimal spatial heterogeneity (δs) leading to the slowest mean fixation time from an initial state of $(0, N/2, 0)$. Depending on the specific parameter regime, the optimal selection pressure profile is the one with the largest possible valley consistent with $\langle s \rangle$ (purple) or the one having the largest possible peak (yellow). (c) Relative magnitude of $\tau_f^{\delta s_{max}}$ (mean fixation time at maximum value of δs) and $\tau_f^{\delta s_{min}}$ (mean fixation time at minimum value of δs) as mutation rate decreases at constant migration rate (green arrow, panel (b)). $N = 24$ and $\langle s \rangle = 0.167$ in all panels.

with the slowest fixation time—in different regions of parameter space (Figure 2.6b). We observe two distinct regions of parameter space that lead to two very different dosing regimes. For μ sufficiently large relative to β, slowest fixation occurs when we maximize the amount of drug in the center microhabitat ($\delta s = 0.5$, yellow region). On the other hand, at large migration rates fixation is optimally slowed by maximizing the amount of drug in the two microhabitats without any initial mutants ($\delta s = -0.2$). In contrast to the case with no initial mutants (e.g., Figure 2.5), fixation time is never maximized by choosing the homogeneous profile. To further characterize these two regimes, we compare the fixation times from a maximally peaked landscape (δs is maximized) to
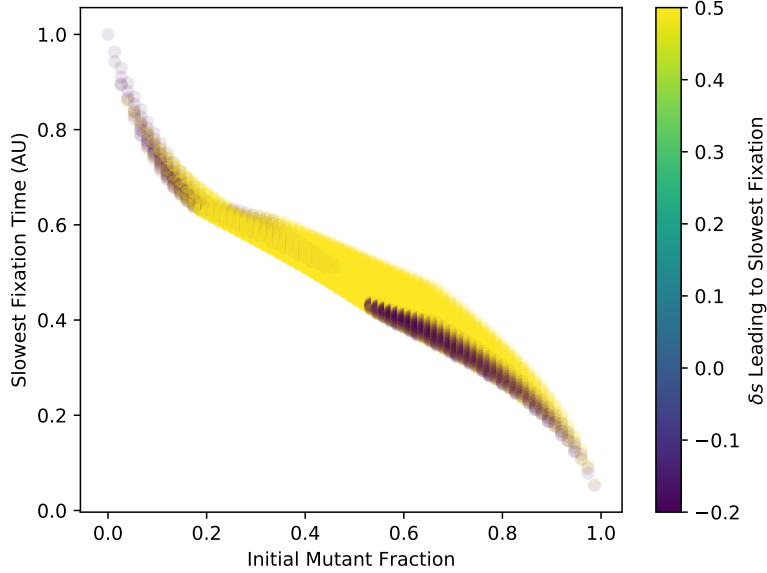
Figure 2.7: The value of δs leading to the slowest fixation time for each possible initial state and its corresponding fixation time is plotted for $\mu = 5 \times 10^{-3}$, $\beta = 0.08$, $N = 25$, and $\langle s \rangle = 1/6$. In addition to the parameters of our system, the initial state is important for determining the optimum dosing even in our simple model of microbial evolution.

that from a landscape with a large valley (δs is minimized). The selection landscape that leads to the slowest fixation rapidly becomes sub-optimal as mutation rate is decreased at constant β (Figure 2.6c).

In addition to μ and β, the initial mutant subpopulation affects the optimal selection landscape for the system. To explore more of the possible initial states of the system, the optimal value of δs is plotted for all possible initial states in Figure 2.7 with $\beta = 0.08$ and $\mu = 5 \times 10^{-3}$ and $\langle s \rangle = 0.167$. For most of the possible states, the slowest fixation is achieved with either the largest possible selection peak ($\delta s = 0.5$) or largest possible selection valley ($\delta s = -0.2$). Interestingly, for a fixed mutant fraction, the optimal selection landscape is often either of these boundary values for δs. Clearly, determining the optimal dosing policy is a complicated problem affected by a large number of factors both environmental and inherent to the specific microbe, even in the very simple toy model that we have developed.

2.5 MFPT Tensor Structure

Thus far, our analysis has focused on understanding the nature of τ_f, the single time to fixation from an initial state with zero mutants. We briefly considered an alternative initial state, looking
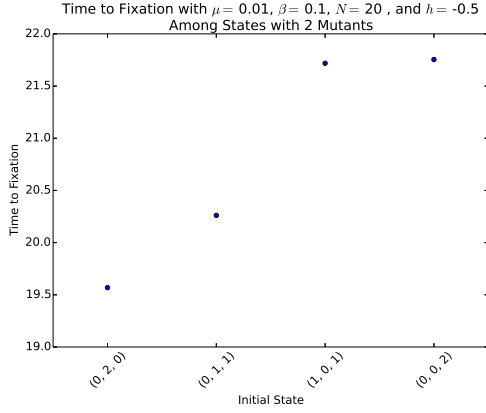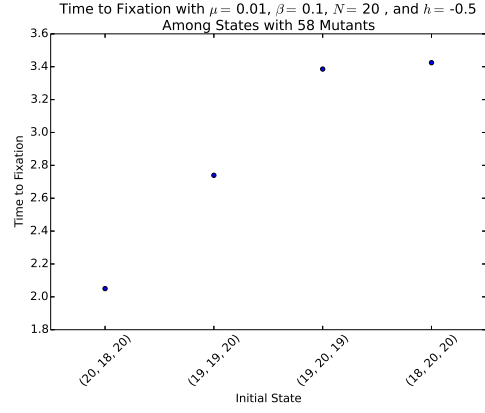
(a) 2 Total Mutants   (b) 2 Total Wild-Type

Figure 2.8: Comparing the time to fixation for different states with (a) 2 total mutants and (b) 2 total wild-type cells.

at the fixation times starting with an initial mutant subpopulation in Figure 2.7. Recall that these additional times to fixation are necessarily calculated to arrive at a value for τ_f, so studying the time to fixation from alternative initial states permits us to better understand the evolutionary dynamics without requiring any additional calculations.

We can continue with this approach, looking at particular choices of initial states to learn about the convergence towards fixation. As one concrete example, we can look at specific fixation times for a landscape with $\delta s = 0.5$ in a system with $\mu = 0.01$, $\beta = 0.1$, and $N = 20$ with very few mutants and with very few remaining wild-type cells. In states with two initial mutants, we see in Figure 2.8a that fixation is achieved most quickly when both mutants are in the center microhabitat, which offers the greatest selection pressure. This matches our intuition, since maximizing the number of mutants in the highest selection pressure microhabitat minimizes the likelihood of the small initial mutant population going extinct due to genetic drift. Conversely, minimizing the number of mutants in this microhabitat maximizes the probability of extinction due to genetic drift. If we look at initial states with only two remaining wild-type cells, shown in Figure 2.8b, we observe the opposite trend. Now that the probability of extinction is negligible, maximizing the number of mutants in the lowest selection pressure microhabitat leads to the slowest fixation and minimizing the number of mutants in this microhabitat leads to the fastest fixation.

2.5.1 Tensor Decomposition

Selecting individual initial states of interest is informative but far from exhaustive. Ideally, we could understand something more general in the calculated fixation times. Recall that the result

16
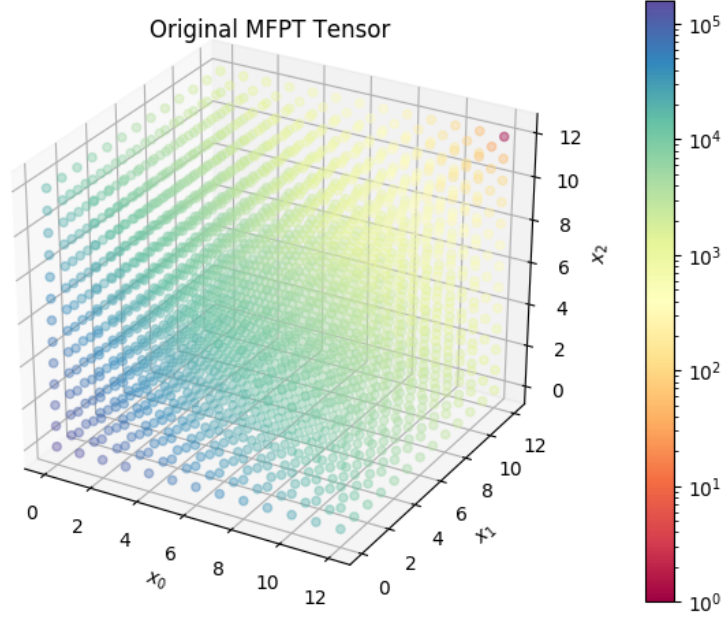
Figure 2.9: Full MFPT tensor $\mathcal{T}$ calculated using $N = 12$, $\mu = 10^{-6}$, $\beta = 0.1$, $\langle s \rangle = 1/6$, and $\delta s = 0.2$.

of solving Equation 2.2, in its most natural formulation, is an order-3 tensor, $\mathcal{T}^{N_0^* \times N_1^* \times N_2^*}$, with the indices $0 \leq i_m \leq N_m^*$ representing the time to fixation with i initial mutants in microhabitat m. An example of the resulting tensor is shown in Figure 2.9, where we use $N = 12$ for ease of visualization. This tensor encodes important information for understanding the trajectory taken by the system as it reaches fixation. We now look at the underlying structure in the 3rd order MFPT tensor itself rather than focus on individual scalar values. To this effect, we decompose the MFPT tensor into a sum of the outer product of r rank-1 tensors (vectors) according to

$$\mathcal{T} = \sum_{i=1}^{r} \lambda_r \mathbf{a}_{0,i} \otimes \mathbf{a}_{1,i} \otimes \mathbf{a}_{2,i}, \tag{2.11}$$

where λ_r is a scalar weight and $\mathbf{a}_{m,i} \in F^{N_m^*}$ for $m \in [0,2]$. Intuitively, this is the multilinear algebra generalization of the singular value decomposition from linear algebra. This decomposition allows us to work with a series of vectors rather than higher order tensors. If we assign a single vector to each of the three microhabitats, we can truncate the sum to a single term expressed by

$$\hat{\mathcal{T}} = \lambda \mathbf{t}_0 \otimes \mathbf{t}_1 \otimes \mathbf{t}_2, \tag{2.12}$$

17

where $\mathbf{t}_i$ is the vector representing the dependence on the number of mutants in the ith microhabitat on the fixation times. Our goal then is to determine $\mathbf{t}_i$ and λ to minimize the approximation error

$$\mathcal{E} = \left\| \mathcal{T} - \hat{\mathcal{T}} \right\|_F, \tag{2.13}$$

characterized by the Frobenius norm of the difference between the true tensor $\mathcal{T}$ and the reconstructed tensor $\hat{\mathcal{T}}$. The general procedure for determining these factors is a long-solved problem [75], that goes by several different names, including CANDECOMP [76], PARAFAC [77], and polyadic decomposition. We will refer to this procedure as simply CP decomposition. In practice, this decomposition is performed using an alternating least-squares approach, in which all but one vector is fixed, with the free vector determined by solving a least squares system [78]. For example, if $\mathbf{t}_1$ and $\mathbf{t}_2$ are fixed, $\mathbf{t}_0$ is the solution to the least squares problem

$$\left\| \mathbf{T}_{(1)} - \hat{\mathbf{t}}_0 (\mathbf{t}_2 \odot \mathbf{t}_1)^T \right\|_F, \tag{2.14}$$

where $\mathbf{T}_{(1)}$ is mode-1 matricization of $\mathcal{T}$. This has the solution

$$\hat{\mathbf{t}}_0 = \mathbf{T}_{(1)} \left[(\mathbf{t}_2 \odot \mathbf{t}_1)^T \right]^\dagger. \tag{2.15}$$

By alternating between the free vector, we converge to the component vectors that produce $\hat{\mathcal{T}}$ minimizing the reconstruction error. Implementing the standard CP decomposition, we produce the reconstruction tensor shown in Figure 2.10a, with residuals in Figure 2.10b. The reconstructed tensor formed from a single set of vectors captures 99.4% of the Frobenius norm of the original tensor. The reconstruction error is particularly impressive considering the amount of information discarded with a rank-1 truncation of the decomposition. In general, this approach compresses an $(N+1)^M$ tensor into M vectors of length $N+1$. In this example with $N = 12$, the reconstruction tensor requires storing less than 2% of the number of scalar values in the original tensor.

We can actually perform a modified version of CP decomposition for MFPT tensors calculated using symmetric selection landscapes that is faster and slightly improves upon the reconstruction error. Due to the underlying symmetry of the selection landscape, we necessarily have $\mathbf{t}_0 = \mathbf{t}_2$, up to a normalization factor. This reduces the number of vectors needed to be solved in the alternating least squares approach, speeding up the runtime. The reconstruction error can be improved by considering the single fixation time $T((N, N, N)|(N, N, N))$. By definition, the time require to first reach fixation from the state associated with fixation is zero, and $\mathcal{T}_{N,N,N}$ satisfies this property. However, in the reconstructed tensor, this element is calculated according to

$$\hat{\mathcal{T}}_{N,N,N} = \lambda \mathbf{t}_0[N] \mathbf{t}_1[N] \mathbf{t}_2[N], \tag{2.16}$$

18

(a) Reconstructed Tensor $\hat{\mathcal{T}}$ (b) Residuals $\hat{\mathcal{T}} - \mathcal{T}$
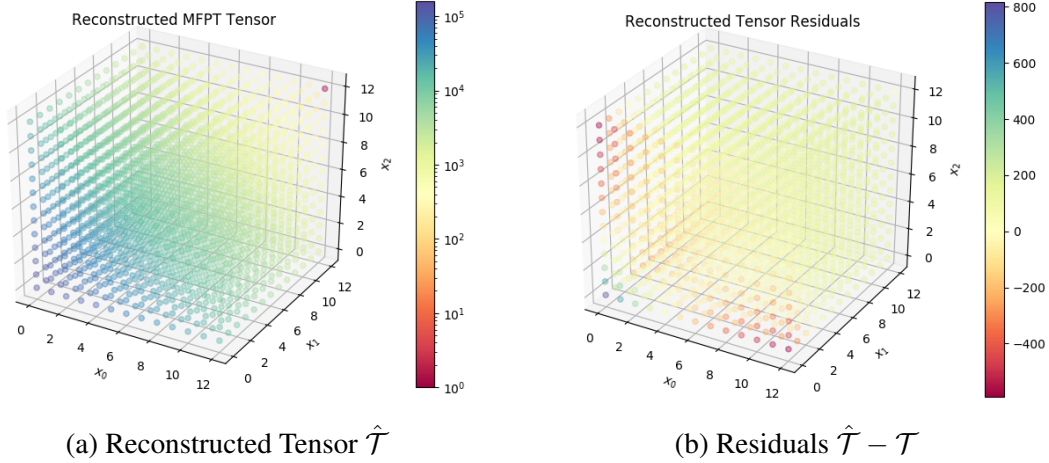
Figure 2.10: (a) The reconstructed tensor $\hat{\mathcal{T}}$ formed from a single set of vectors and (b) its associated residuals.

which will not be zero unless one of these vectors has a value of 0 for the state with N mutants. However, having a value of 0 for this single element would lead to several undesired zero elements in $\hat{\mathcal{T}}$ upon performing the outer product. So not only will the reconstructed tensor not accurately calculate $\hat{\mathcal{T}}_{N,N,N}$, but failing to account for the fact that this value is identically equal to zero will add an artificial constraint that biases the elements in our reconstruction vectors. To remedy these concerns, we treat the fixation time from the (N, N, N) state as a missing value according to [79], so that it does not impact the least squares solution. The basic idea is to modify the Frobenius norm by introducing a mask

$$\mathcal{M}_{ijk} = \begin{cases} 0 & i = j = k = N, \\ 1 & \text{else}, \end{cases} \tag{2.17}$$

so that we instead work with the Hadamard product $M * \mathcal{T}$. For our system, we find that the improvement in reconstruction error is relatively small, but this modification allows us to be confident in the calculated reconstruction vectors.

2.5.2 Probing Reconstruction Fixation Vectors

With a robust method for decomposing the MFPT tensor into simple vectors, we direct our attention to learning more about evolution from these simple vectors. First, we are interested in understanding the similarity of the reconstructed fixation time vectors between the center and ex-

trema microhabitats. This similarity is quantified according to

$$\gamma^{\text{recon}} = \mathbf{t}_0 \circ \mathbf{t}_1, \tag{2.18}$$

where $\circ$ denotes the cosine similarity, defined as

$$\mathbf{x} \circ \mathbf{y} = \frac{\mathbf{x} \odot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}. \tag{2.19}$$

The natural comparison for this system is the time to fixation vector for the corresponding single microhabitat system. The single microhabitat fixation time vectors τ_{x_0} and τ_{x_1} are calculated using the same system parameters and with the same selection pressure values found in x_0 and x_1, respectively. The similarity between the fixation times in the center and extrema microhabitats from the single microhabitat fixation time calculation is calculated according to

$$\gamma^{\text{single}} = \tau_{x_0} \circ \tau_{x_1}. \tag{2.20}$$

Together, γ^{recon} and γ^{single} allow us to understand the relative importance of mutants in different microhabitats for both sets of vectors. We can also look at the similarity of the vectors produced from these two different methods. To this effect, we define

$$\xi^{\text{extrema}} = \mathbf{t}_0 \circ \tau_{x_0}, \tag{2.21}$$

$$\xi^{\text{center}} = \mathbf{t}_1 \circ \tau_{x_1}. \tag{2.22}$$

The reconstructed and single vial vectors for our MFPT tensor are shown in Figure 2.11, which also shows the values of the four similarity metrics that we have defined. We see that for this particular landscape and choice of parameters, there is a significant difference both between the center and extrema microhabitat vectors and also between the vectors formed from decomposition and those calculated from the single-microhabitat MFPT equation. Therefore, both the heterogeneity as well as the existence of multiple microhabitats play an important role in the dynamics for this particular system as it reaches fixation.

In general, we want to look more broadly across the parameters of our system. Our analysis of this more general behavior is far from complete, but to highlight the richness of the MFPT structure, we will restrict ourselves to the homogeneous selection pressure landscape. Even in this simple landscape, we observe several complex features. As on example, we first look at the reconstruction error assuming $\langle s \rangle = 1/6$ and $\delta s = 0$ as a function of μ and β in Figure 2.12. Although the reconstruction error is quite small in many regions of parameter space, including the reference point $\mu = 10^{-6}$ and $\beta = 0.1$ that we have been using, there are regions of parameter
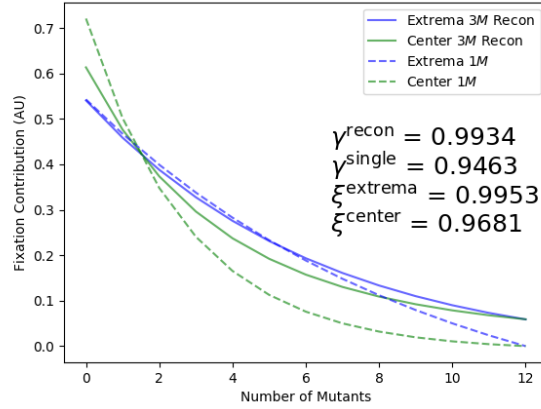
Figure 2.11: Comparing the single microhabitat vectors formed from CP decomposition to those obtained from solving the single-microhabitat MFPT equation.

space with reconstruction error a full order of magnitude larger than that of our reference point. Furthermore, the boundaries separating these different regions in the reconstruction error plot are neither simple nor immediately intuitive. We observe similar complexities when examining γ^{recon} in Figure 2.13 and ξ^{center} in Figure 2.14. More work is needed to understand how the reconstruction error and similarity metrics are affected by the system parameters μ and β, but this approach has the potential to simplify the rich evolutionary dynamics observed in our simple toy model of evolution.

2.6 Conclusion

Our model is a dramatic oversimplification of the biological dynamics leading to drug resistance. Practical applications will require analysis of more realistic models and may call for spatial optimizations with different constraints–for example, limits on the maximum allowable local selection pressure. Nevertheless, the simplicity of our model allows for a thorough characterization of fixation time over a wide range of parameters, and its behavior is surprising rich. Importantly, our results do not require a fitness cost of resistance or a genetic fitness valley, and they predict that spatial heterogeneity in drug concentrations would impact populations of motile and non-motile cells in opposing ways, even when mutations rates are relatively similar. While heterogeneity is likely to accelerate evolution for populations of motile bacteria, similar to what is observed in experiments with *E. coli* [38, 36], our results predict slowed evolution for less motile cells (e.g., the nosocomial pathogen *E. faecalis* [80]) or cells with rapid mutation rates. Perhaps most interestingly, our results suggest counter-intuitive, spatially optimal profiles for slowing the spread of resistance sub-populations. In the long term, these results may lay the groundwork for optimized,
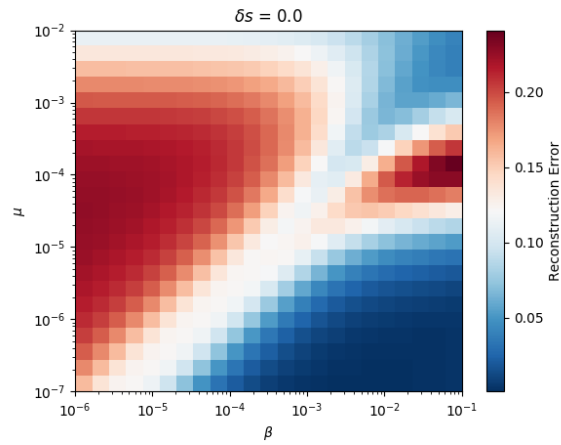
Figure 2.12: The reconstruction error between $\mathcal{T}$ and $\hat{\mathcal{T}}$ shows a complex dependence on μ and β even with a spatially homogeneous selection pressure landscape.
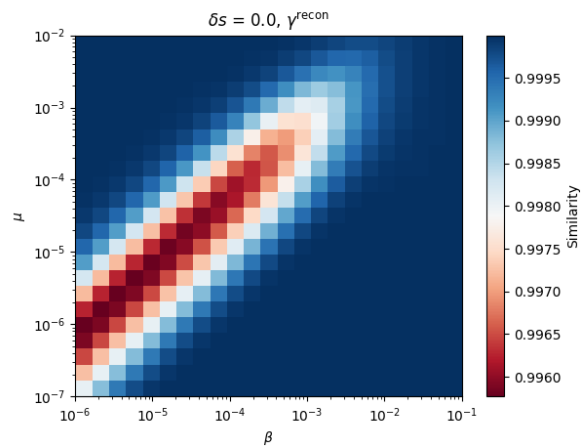
Figure 2.13: The cosine similarity of the reconstructed center microhabitat and extrema microhabitat fixation vectors calculated using tensor decomposition shows complex structure with the system parameters μ and β.
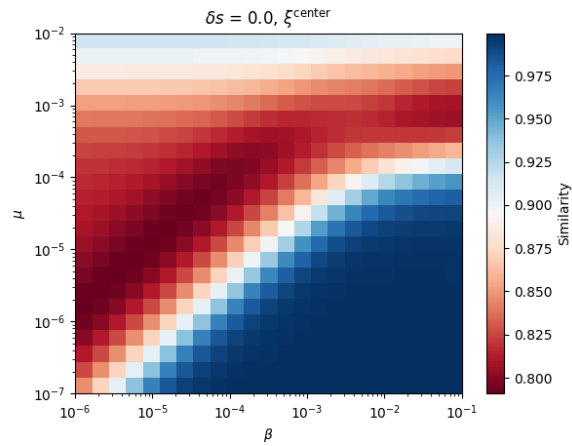
Figure 2.14: The cosine similarity of the center microhabitat fixation vector calculated from tensor decomposition and calculated from the single-microhabitat MFPT also shows complex structure with the system parameters.

spatially-resolved drug dosing strategies for mitigating the effects of drug resistance.

CHAPTER 3

Modulating Resistance in a Simplified System with Imperfect Mutants and with Larger Systems

3.1 Introduction

To more fully understand the emergence of antibiotic resistance in spatially-extended systems, we build upon our previous work using a toy model of Moran-like stochastic evolution in small systems [81]. A large number of simplifying assumptions went into this model to permit semi-analytical solutions for calculating metrics quantifying the emergence of resistance in spatially-extended systems. Principally, we assumed that the system evolved under Moran-like evolution in a way that is amenable to a master equation approach. Within this framework, we were able to use a mean first passage time (MFPT) calculation to determine the fixation time from arbitrary initial states. The MFPTs revealed rich dynamics emerging from the system parameters and spatial heterogeneity despite the many simplifying assumptions. Because this model was able to support non-trivial behavior and yet also permit semi-analytical calculations, we use this framework of Moran-like, stochastic evolution characterized by a master equation as we generalize some of the initial assumptions. Although using this framework does constrain the space of models that we can consider, there are generalizations to be made within this framework that will extend the applicability of our previous results and help establish fundamental relationships between spatial heterogeneity and evolutionary dynamics in more general systems. In this work, we focus specifically on more faithfully capturing several important properties observed in *in vitro* microbial populations.

First, we note that mutations conferring resistance frequently involve a fitness cost—that is, such mutations often lower the fitness of the mutant in the absence of drug [82, 83]. We previously were able to observe that spatial heterogeneity affected the fixation times in systems even when the spatially conserved selection pressure was held constant across different selection pressure landscapes and the mutants were assumed to be "perfect"—they achieved the maximum possible growth rate both in the presence and absence of drug. We now relax this assumption to allow for imperfect mutants. We can expect to observe qualitatively new behavior, since imperfect mutants

allow for the possibility of microhabitats in which the mutants are selected against rather than the wild-type cells.

Second, we look to generalize our model to systems with more than three connected microhabitats, since this number is both arbitrary and artificially small for many evolution scenarios of interest. We want to study systems small enough that finite size effects are not trivial—that is, we want to avoid the continuous limit in which the dynamics are described by differential equations or a mean-field theory approach. Calculating fixation times in larger, discrete systems in which genetic drift, migration, and *de novo* mutation continue to be stochastic processes allows us to better understand how the results generalize to different finite-sized system.

3.2 Accounting for Imperfect Mutants

To account for imperfect mutants, we incorporate a fitness cost to the mutant cells. To investigate the impact of fitness cost, we assume that mutant cells have a fitness given by $r^* = 1 - \epsilon$ for $\epsilon \geq 0$. Physically, these constraints assume that the mutant cells cannot grow faster than the wild-type cells in a drug-free environment but may replicate more quickly than the wild-type cells in the presence of a drug. The transition rates incorporating a fitness cost can be found in Appendix B.1.

3.2.1 Single Microhabitat Dynamics

We start with the simple case of a single microhabitat and denote the fitness of wild-type cells as r_0. For simplicity, we will assume that $\mu \ll 1$ so that the timescale of *de novo* mutation is separable from the genetic drift timescale during which a mutant population either fixates or goes extinct. The single microhabitat transition rates for genetic drift are formed by modifying the rates in Equation 2.7 to incorporate a fitness cost ϵ. The resulting transition rates can be used to write

$$\mathcal{P} = \frac{(1 - \epsilon) - r_0}{1 - \epsilon} \left[1 - \left(\frac{r_0}{1 - \epsilon} \right)^N \right]^{-1}, \tag{3.1}$$

where $\mathcal{P}$ is the probability of fixation in single microhabitat with imperfect mutants. When $r_0 < 1 - \epsilon$, the wild-type cells never have a fitness advantage, regardless of drug concentration, and the problem reduces to the cost-free case. However, when $r_0 > 1 - \epsilon$, the mutant cells are selected against. Our primary interest is understanding this latter case, and we will examine the effect of a fitness cost on the probability of a single mutant cell fixating when it is selected against. Note that with $\epsilon = 0$, the mutant cells are equivalent to the wild-type cells and the system evolves according to neutral genetic drift so that $\mathcal{P} = 1/N$. Without loss of generality, let $r_0 = 1$; as the fitness cost is increased,the mutant cells are selected against and the probability of the mutant cell fixating
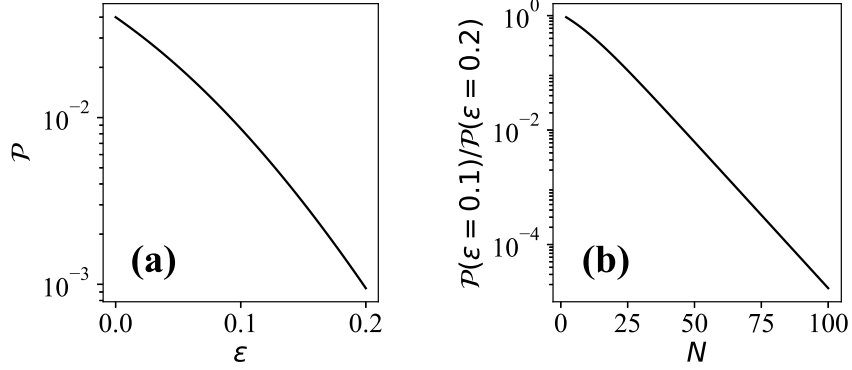
Figure 3.1: (a) Probability of fixation of a single mutant cell, $\mathcal{P}$, with $N = 25$ total cells calculated in a single microhabitat with a given fitness cost ϵ. (b) The ratio of the probability of a single mutant fixating with fitness cost $\epsilon = 0.1$ to the probability of fixation with $\epsilon = 0.2$ decreases exponentially with population size N.

decreases (Figure 3.1(a)). Note that doubling the fitness cost from $\epsilon = 0.1$ to $\epsilon = 0.2$ reduces the fixation probability by a full order of magnitude, and as expected, the effect depends strongly on system size N (Figure 3.1(b)).

3.2.2 Preliminary Spatially-Extended Results

Our goal is to determine how a non-zero fitness cost impacts fixation times under different spatially-extended fitness profiles. From our previous results without a fitness cost, we anticipate that the relative magnitude of $\langle r_0 \rangle$ should have important consequences for dynamics our system. If $r^*(x_i)$ is much larger than the mean wild-type fitness $\langle r_0 \rangle$, we expect qualitatively similar results to what was observed without a fitness cost. But when the mutant and wild-type fitness values are comparable, qualitatively new dynamics could be observed since there could exist microhabitats in which the mutants are selected against. To test this intuition, we first calculate the MFPTs across a broad range of μ and β values across different fitness landscapes satisfying $\langle r_0 \rangle = 3/5$ and with a fitness cost of $\epsilon = 0.2$ (Figure 3.2(a)). In this regime, we expect similar results to those obtained without a fitness cost, and we see that the spatial distribution of fitness can again affect the fixation time, even though $\langle r_0 \rangle$ is conserved across different landscapes. We observe the existence of the same regions of parameter space that were present without a fitness cost, indicating that the specific effect of spatial heterogeneity depends on the system parameters as before. We next form the phase diagram assuming $\langle r_0 \rangle = 5/6$ and $\epsilon = 0.2$ (Figure 3.2(b)), so that qualitatively new behavior can potentially arise due to the similarity between mean wild-type and mutant fitness values. However, the phase diagram appears to be very similar to that generated without any fitness cost. So although spatial heterogeneity still impacts the fixation times with a fitness cost in this regime,
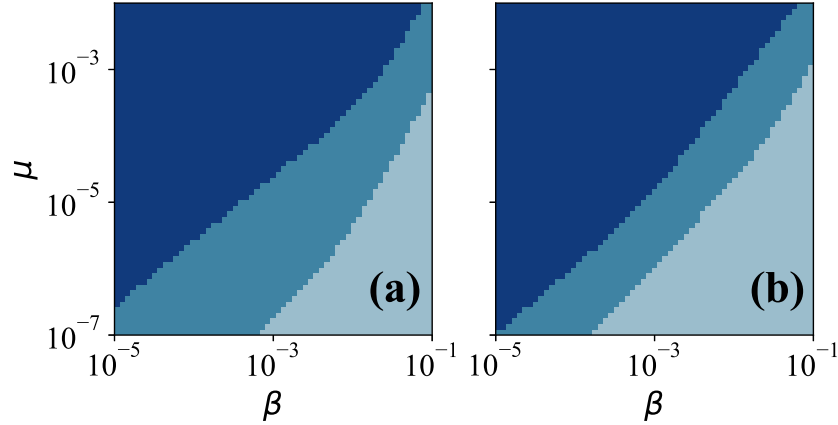
Figure 3.2: (a) The fixation times τ_f calculated with $\langle r_0 \rangle = 3/5$ and $\epsilon = 0.2$ again produce regions of parameter space in which the homogeneous spatial landscape leads to a maximum (light blue), minimum (dark blue), or intermediate (medium blue). (b) The fixation times calculated with $\langle r_0 \rangle = 5/6$ and $\epsilon = 0.2$ produce qualitatively similar regions. All plots were generated using $N = 25$.

any new dynamics introduced with a fitness cost do not manifest in the different phases in our parameter space.

However, introducing a fitness cost does affect how specific landscapes reach fixation within these regions and the magnitude of the effect of spatial heterogeneity. We demonstrate this by modifying the thought experiment introduced in Figure 2.1, so that the fitness landscape is fixed bu the fitness cost ϵ is varied. We parameterize the wild-type fitness according to $r_0(x_0) = r_0(x_2) = B$ and $r_0(x_1) = B + h$, for some background fitness B and offset h. Upon calculating the fixation times for this thought experiment, we find that heterogeneity exclusively slows fixation in the limit $\beta \ll \mu$ (Figure 3.3(a)) and generally, but not always, slows fixation in the limit $\mu \ll \beta$ (Figure 3.3(b)).

To understand the results of this thought experiment, we first return to the previous results obtained without any consideration of a fitness cost ($\epsilon = 0$). In the limit $\beta \ll \mu$, migration is negligible and any spatial heterogeneity leads to slower fixation times. We see this behavior in Figure 3.3(a) looking at $\epsilon = 0$. As the fitness cost increases, the fixation time increases as the genetic drift favors the mutants less. However, the $h = 0$ landscape continues to lead to the fastest fixation, since any spatial heterogeneity results in a microhabitat in which the genetic drift becomes more biased against the mutants as the fitness cost increases. In the limit $\mu \ll \beta$, we saw that the slowest fixation time was achieved with $h = 0$ without a fitness cost, and we see this again looking at Figure 3.3(b) with $\epsilon = 0$. In this limit, *de novo* mutations are rare and any spatial heterogeneity leads to microhabitat in which the mutants are less likely to go extinct. As the fitness cost increases, however, the $h = 0$ landscape no longer leads to the slowest fixation times. While
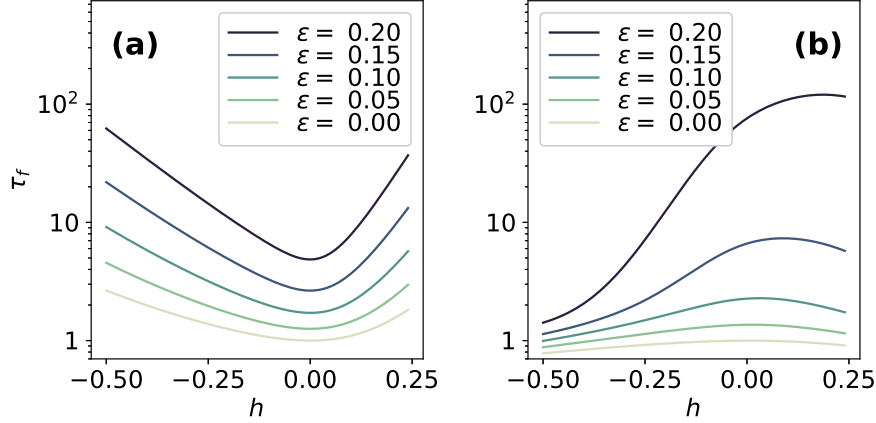
Figure 3.3: (a) Fixation times calculated for a given fitness cost ϵ for all landscapes parameterized by center microhabitat fitness offset h with $\mu = 0.01$ and $\beta = 10^{-6}$. (b) Fixation times calculated using $\mu = 10^{-6}$ and $\beta = 0.01$. Both panels use $\langle r_0 \rangle = 5/6$ and $N = 25$.

any spatial heterogeneity continues to reduce the probability of mutant extinction in at least one microhabitat, landscapes can be constructed in which the mutants are selected against in at least one microhabitat with a non-zero fitness cost. These two competing behaviors result in the value of h leading to the slowest fixation time shifting with fitness cost.

3.2.3 Constraints Introduced from Selection Pressure Conservation

The results of this simple, idealized experiment give us clues into the dynamics we could expect to observe in a spatially-extended system with a fitness cost. Although this experiment seems very natural to carry out in this manner, there is an important confounding factor for which we would have failed to account—by fixing the wild-type fitness values and varying the fitness cost, the different systems no longer conserve the mean selection pressure $\langle s \rangle$. If we assume the selection pressure in microhabitat x_i is given by $s(x_i) = r^*(x_i) - r_0(x_i)$, we cannot simply fix the wild-type fitness landscape and vary the fitness cost while conserving $\langle s \rangle$ across different landscapes. Rather, the underlying fitness landscape for a given h must also change with the fitness cost to conserve $\langle s \rangle$.

To properly understand the amplified effect of spatial heterogeneity observed in a system with a fitness cost, the fixation times for a fixed $\langle s \rangle$ can be examined for different fitness costs ϵ. For simplicity, we will start with the homogeneous landscape. The transition rates for our model (Appendix B.1) can be rewritten so that the wild-type fitness $r_0(x_i)$ and mutant fitness $1 - \epsilon$ do not appear as separate terms but only as a ratio. In the homogeneous landscape, the fitness ratio is
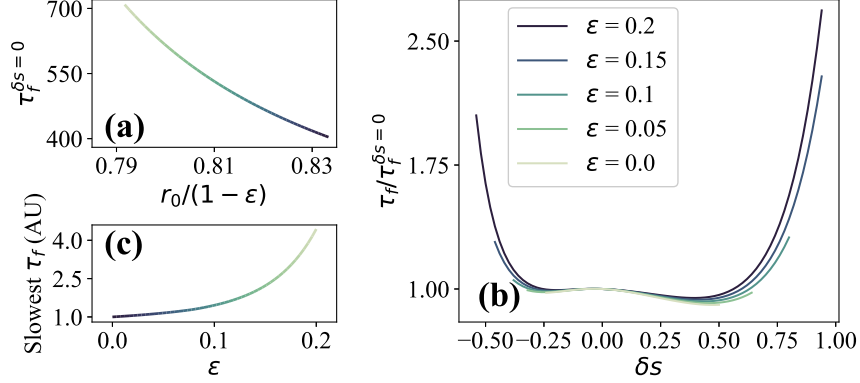
Figure 3.4: (a) The fixation times τ_f plotted for all physical selection pressure landscapes satisfying $\langle s \rangle = 1/6$ for a given fitness cost ϵ. As the fitness cost is increased, the range of valid δs values increases as does the range of fixation times observed. (b) The fixation time for the homogeneous landscape scales almost linearly with the fitness ratio $r_0/(1-\epsilon)$. (c) Slowest fixation times observed across all landscapes for different fitness costs ϵ while conserving $\langle s \rangle + \epsilon = 7/30$. All fixation times calculated using $\mu = 10^{-4}$, $\beta = 0.01$, and $N = 25$.

constant across all three microhabitats. We define $f \equiv r_0/(1-\epsilon)$, which is given by

$$f(\epsilon, \langle s \rangle) = \frac{1 - (\epsilon + \langle s \rangle)}{1 - \epsilon}, \tag{3.2}$$

after solving for the wild-type fitness value r_0 needed to ensure $\langle s \rangle$ is conserved. Because the transition rates of our model only depend on f, the fixation times can be equivalent across different fitness costs provided that the ratio of the wild-type fitness to the mutant fitness remains fixed. However, we cannot impose both this constraint and the constraint of conserving $\langle s \rangle$. If we conserve $\langle s \rangle$, we see in Figure 3.4(a) that the calculated homogeneous fixation times scale almost linearly with f, which changes relatively little for $\epsilon \in [0, 0.2]$.

We quantify the difference between the transition rates upon adding a non-zero fitness cost by defining $\Delta_\epsilon^+[x_i](j, k, \ell) \equiv T_\epsilon^+[x_i](j, k, \ell) - T_{\epsilon=0}^+[x_i](j, k, \ell)$ and $\Delta_\epsilon^-[x_i]$ analogously. Both are complicated expressions due to the nonlinearity of the transition rates. Upon performing a Taylor expansion of f about $\epsilon = 0$, we find that $\Delta_\epsilon^+[x_i] \geq 0$ and $\Delta_\epsilon^-[x_i] \leq 0$, and that the magnitude of each scale linearly with ϵ (Appendix B.2). Since the transition rate increasing the number of mutants is bounded below by the transition rate obtained with $\epsilon = 0$ and the transition rate decreasing the number of mutants is bounded above by the transition rate obtained with $\epsilon = 0$, we necessarily obtain fixation more quickly upon the addition of a fitness cost with the homogeneous landscape.

Moving beyond the homogeneous landscape, we can calculate the fixation times for arbitrary landscapes conserving $\langle s \rangle$ as a function of δs. The normalized fixation times are plotted across sev-

eral fitness cost values ϵ in Figure 3.4(b), where all selection pressure landscapes have $\langle s \rangle = 1/6$. Note that different fitness costs permit different ranges of δs resulting in a physical fitness landscape. After accounting for the shift in the homogeneous fixation times, the fixation times are qualitatively similar across values of δs that produce physical landscapes across all fitness costs. Upon moving outside of this range of landscapes, systems with larger fitness costs support selection profiles exhibiting significantly greater variation in the resulting fixation times. Spatial heterogeneity impacts the fixation times in all of these systems, but the effect of spatial heterogeneity is amplified by a fitness cost, at least for this choice of parameters. The range of physical values of δs is constrained by the requirement $0 \leq r_0(x_i) \leq 1$ in every microhabitat x_i (i.e., that the wild-type fitness must lie between 0 and 1). Imposing these constraints, we find

$$\delta s_{min} = \max \left\{ -\frac{3}{2}(\langle s \rangle + \epsilon), -3(1 - \langle s \rangle - \epsilon) \right\}, \tag{3.3}$$

$$\delta s_{max} = \min \left\{ \frac{3}{2}(1 - \langle s \rangle - \epsilon), 3(\langle s \rangle + \epsilon) \right\}. \tag{3.4}$$

See Appendix B.3 for the derivation. Note that $\langle s \rangle$ and ϵ only appear together as a summed quantity, indicating that they play the same role in determining the valid range of δs values for the system. However, they do not have the same effect on the fixation times. This is illustrated in Figure 3.4(c), which plots the slowest fixation times observed for a given fitness cost while keeping $\langle s \rangle + \epsilon$ constant. Although the range of valid δs values is only a function $\langle s \rangle + \epsilon$, the resulting fixation times are less simple. Note that in the single microhabitat system, the fitness cost can be modeled as an effective mean selection pressure, so that adding a fitness cost produces no new behavior. But in the spatially-extended system, no self-consistent rescaling of the mean selection pressure exists, indicating that qualitatively new behavior can be observed in multi-microhabitat systems.

To better understand the qualitatively new behavior that can be present with a fitness cost, we look at when the system shifts from genetic drift biased against the wild-type cells to against the mutant cells in individual vials. Neutral genetic drift occurs in the center microhabitat with $\delta s = -\frac{3}{2}\langle s \rangle$, independent of ϵ. However, the resulting fitness landscape is only physical when the system satisfies

$$-\frac{3}{2}\epsilon \leq \langle s \rangle \leq \frac{2}{3}(1 - \epsilon). \tag{3.5}$$

A similar expression can be found for the condition on $\langle s \rangle$ to achieve neutral selection pressure in the extrema microhabitats. These constraints determine when selection landscapes can be constructed that will begin to favor the wild-type cells in a single microhabitat, but they do not predict

the magnitude of the genetic drift biased against the mutant cells. To achieve the slowest fixation times, the single microhabitat results suggest that the genetic drift bias against mutant cells should be maximized to slow their proliferation. The largest possible selection advantage of the wild-type cells is simply given by ϵ.

3.2.4 System-Size Expansion of Master Equation

To better understand the significantly slower fixation times possible with a fitness cost, we examine the system in the deterministic limit $N \to \infty$. We form coupled differential equations for the mutant population in each of the three microhabitats using a system-size expansion to write [74]

$$\frac{dn_i}{dt} = \langle T^+[x_i] \rangle - \langle T^-[x_i] \rangle, \tag{3.6}$$

where n_i is the mutant fraction in microhabitat i. The coupled nonlinear differential equations have no analytical solution, but numerical linear stability analysis provides information about the existence and stability of fixed points. We find that for certain choices of parameters, the state of fixation can be an unstable fixed point and a new state emerges as a stable fixed point. Without a fitness cost the genetic drift can never be biased against the mutant cells in any microhabitat, so the state corresponding to fixation is always a trivial fixed point. Due to the nonlinearity of the coupled differential equations, even the existence and nature of these fixed points cannot be written down analytically in terms of the system parameters. Instead, we form a simplified picture using the results from the single microhabitat system and neglecting migration between microhabitats. We approximate the spatially-extended deterministic system as having a nontrivial fixed point whenever $r_0(x_i) > 1 - \epsilon$ for any microhabitat x_i. In general, the wild-type fitness of a microhabitat can be written as

$$r_0(x_i) = d[x_i](\delta s, \langle s \rangle) + (1 - \epsilon), \tag{3.7}$$

for some fitness deviation $d[x_i]$ from the mutant fitness. Within this simplified framework, the sign of this deviation function determines when a stable non-trivial fixed point can be expected to exist. Interestingly, this term has no explicit dependence on ϵ, so the existence of a nontrivial fixed point depends explicitly only on the selection landscape and its topology. However, upon enforcing the physical constraint $r_0(x_i) \leq 1$, we require $d[x_i] \leq \epsilon$, so that the fitness cost does constrain the

existence of non-trivial stable fixed points. For our system, we find

$$d[x_1](\delta s, \langle s \rangle) = -\frac{3\langle s \rangle + 2\delta s}{3}, \tag{3.8}$$

$$d[x_0, x_2](\delta s, \langle s \rangle) = -\frac{3\langle s \rangle - \delta s}{3}. \tag{3.9}$$

See Appendix B.3.3 for details. For a given selection landscape characterized by $\langle s \rangle$ and δs, we expect a non-trivial stable fixed point when $d[x_i] > 0$ in the deterministic limit. In Figure 3.5(a), the different physical values of δs are first plotted as a function of $\langle s \rangle + \epsilon$ in the black rectangular region. Note that the overall shape of this region is fixed with respect to our underlying assumptions on the fitness values permitted. Across the range of valid δs values, there exist different regions in which the existence and nature of fixed points depends on the fitness cost. With a fitness cost of $\epsilon = 0.1$, the blue shaded region shows where we would expect a non-trivial fixed point to exist due to $d[x_1] < 0$, while the red shaded region shows where we would expect a non-trivial fixed point to exist due to $d[x_0, x_2] < 0$. The dashed blue and red lines mark the locations of the boundaries with $\epsilon = 0.2$. Also shown explicitly by the black horizontal line is the class of physically valid landscapes with $\delta s = 0.5$.

The different regions, determined by the nature of the fixed points within them, were calculated by ignoring migration, which was treated as three independent microhabitats, as well as the system parameters β and μ. We numerically integrate the full coupled differential equations across all physical landscapes with $\delta s = 0.5$ in Figure 3.5(b), which used $\mu = 10^{-4}$ and $\beta = 10^{-4}$. We observe a clear bifurcation in the stable mutant population of the extrema populations from the numerical solutions as expected. For $\langle s \rangle + \epsilon < (\langle s \rangle + \epsilon)^*$, which corresponds to the blue shaded region in Figure 3.5(a), the fixed point corresponding to mutant fixation is unstable (dashed line) and there is a stable fixed point (solid line) near the state corresponding to having mutants fixed only in the center microhabitat. As $\langle s \rangle + \epsilon$ is increased past the critical value, there is only a single stable fixed point at complete mutant fixation. The simple single microhabitat approximation examining the sign of $d[x_1]$ predicts a discontinuous bifurcation at the location of the vertical blue dotted line. The approximation neglecting migration correctly predicts the existence of the bifurcation, but the expected location of the bifurcation will only be accurate in the limit $\beta \lesssim \mu$. In Figure 3.5(c) and Figure 3.5(d) the trajectories taken through state-space are plotted for the two reference points labeled in the earlier panels. In Figure 3.5(c), we have a non-trivial fixed point so that all initial states other than mutant fixation converge to a point near $(n_0, n_1) = (0, 1)$. In Figure 3.5(d), we are in the region where $d[x_i] > 0$ so that all initial trajectories converge to the state associated with mutant fixation.

Finally, we now relate these results obtained in the deterministic limit back to the finite-sized system. In the deterministic limit, there is no guarantee of reaching fixation in finite time for
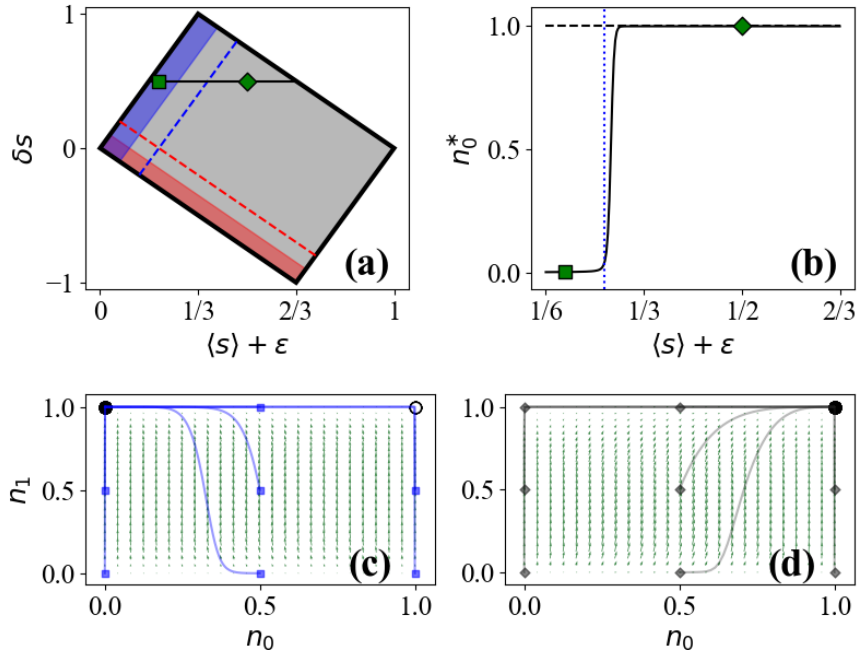
Figure 3.5: (a) The range of valid δs values, outlined in black, encapsulates regions where we expect a non-trivial fixed point due to genetic drift favoring the wild-type cells in the center (red) or extrema (blue) microhabitats. The regions for $\epsilon = 0.1$ are shown, with the boundaries for $\epsilon = 0.2$ represented by dashed lines. (b) For physical landscapes with $\delta s = 0.5$, we observe a bifurcation in stable fixed points (solid line) and unstable fixed points (dashed line) for the mutant fraction in the extrema microhabitat upon numerically integrating the coupled differential equations with $\mu = 10^{-4}$ and $\beta = 10^{-4}$. Dotted vertical line shows the predicted bifurcation point when neglecting migration. (c) Numerically integrating the coupled differential equations for different initial states (blue squares) to obtain trajectories (blue lines) taken across a vector field (green arrows) for a system with $\epsilon = 0.1$ and $\langle s \rangle = 0.1$. Stable and unstable fixed points are shown as black and white circles, respectively. (d) Numerically integrating from initial states (gray diamonds) for a system with $\epsilon = 0.1$ and $\langle s \rangle = 0.4$. Reference points: $\langle s \rangle + \epsilon = 0.2$ (green square) and $\langle s \rangle + \epsilon = 0.5$ (green diamond).

33

an arbitrary landscape with a given non-zero fitness cost. We observe finite fixation times from the MFPT calculation due to the inherently stochastic evolution in a finite-sized system. This stochasticity in the evolutionary dynamics is necessary to reach the state of fixation in landscapes leading to a non-trivial fixed point in the deterministic limit.

For a given selection landscape in the three microhabitat system, we can look at the sign and magnitude of $d[x_i]$ to anticipate when the fixation times calculated with a fitness cost will deviate significantly from those calculated without a fitness cost. Upon imposing that different landscapes must conserve $\langle s \rangle$, however, $d[x_i]$ becomes simply a rescaling of δs. If we trust the results obtained in the deterministic system that the fixation times can be approximated as only a function of $d[x_i]$, and not the fitness cost ϵ, we would be led to believe that the fixation times in the finite-sized system should be independent of the fitness cost ϵ after accounting for different offsets from the homogeneous landscape. Indeed, this is exactly what we observe in Figure 3.4(b). By relating the sign of $d[x_i]$ to δs, the deterministic limit allows us to understand the reasoning behind the rescaled fixation times for a given δs being almost independent of ϵ. We also observed that landscapes supporting large ranges of δs can significantly amplify the effect of spatial heterogeneity on the resulting fixation times. In the deterministic limit of the system, this can be understood as the magnitude of $d[x_i]$ leading to non-trivial fixed points. In the discrete system, evolution escapes these states through stochastic dynamics, and the degree of stochasticity required is determined by δs.

3.3 Generalizing to Large, Finite Systems

3.3.1 Extending Analytical Approximation

We wish to extend our previous results with $M = 3$ microhabitats to systems with more microhabitats. Since solving for the exact fixation time using a mean first passage time calculation requires solving $\mathcal{O}(N^M)$ coupled equations, studying larger systems will rapidly become infeasible with the current approach. Instead, we will extend the approximation we previously developed to understand fixation in the $M = 3$ system without a fitness cost. The main idea of this approximation was to coarse-grain over the states of the system so that entire microhabitats represent the smallest states of the system. Using this ambitious approximation, the $\mathcal{O}(N^M)$ states of the system reduce to M binary states, where each microhabitat is composed either entirely of wild-type cells (unfixated) or mutant cells (fixated). Highlighting the details of the approximation using a more general

notation, the fixated time was approximated according to

$$\hat{\tau}_f \approx \tau_{\min} + \sum_{k=0}^{M-1} Q(x_k)\tau_{\max}^{\beta}(x_k), \tag{3.10}$$

where $Q(x_k)$ is the probability of the kth microhabitat fixing in the absence of mutation, given by

$$Q(x_k) = \lambda_k(0)\tau_{\min}, \tag{3.11}$$

and $\tau_{\min}$ is the minimum of M exponential processes, given by

$$\tau_{\min} = \left(\sum_{k=0}^{M-1} \lambda_i(0)\right)^{-1}, \tag{3.12}$$

where $\lambda_i(n_{\text{fix}})$ is the exponential rate for microhabitat i with selection s_i and n_{fix} fixated neighbors (Equation 2.5) and $\mathcal{P}(s_i)$ is the fixation probability of a single mutant in a microhabitat with selection s_i (Equation 2.8). We also defined $\tau_{\max}^{\beta}(x_k)$, which is the maximum expected time for the remaining two microhabitats to reach fixation assuming microhabitat x_k has fixated. With $M = 3$, this term was simple enough that it could be written out completely for a given fixated microhabitat. For instance, if microhabitat x_2 has already reached fixation, this term with $M = 3$ is given by

$$\tau_{\max}^{\beta}(x_2) = \frac{1}{\lambda_0(0) + \lambda_1(1)} + \frac{\lambda_0(0)}{\lambda_0(0) + \lambda_1(1)}\lambda_1(2)^{-1} + \frac{\lambda_1(1)}{\lambda_0(0) + \lambda_1(1)}\lambda_0(1)^{-1}, \tag{3.13}$$

where the first term is the expected minimum of two independent exponential processes occurring at weights $\lambda_0(0)$ and $\lambda_1(1)$. The second term in Equation (3.13) is the product of the probability that microhabitat x_0 first reaches fixation and the rate at which microhabitat x_1 would then fixate. The third term is analogous but accounting for the opposite order of fixation. We now generalize this expression for $M > 3$. We will generalize the original calculation assuming no fitness cost, although adding a fitness cost would be straightforward, as we have found that the $M = 3$ approximation produces qualitatively similar results to the full MFPT calculation even upon adding a fitness cost. First, we also generalize the existing notation to accommodate multiple fixated microhabitats. If $\{x_i\}$ denotes the set of fixated microhabitats, we can define $\tau_{\min}(\{x_i\})$ to be the expected minimum time for the remaining $M - |\{x_i\}|$ exponential processes to reach fixation.

This minimum is given by

$$\tau_{\min}(\{x_i\}) = \left[\sum_{j \notin \{i\}} \lambda_j \left(\sum_{k \in \{i\}} \delta_{j,k+1} + \delta_{j,k-1} \right) \right]^{-1}, \tag{3.14}$$

where $\delta_{m,n}$ is the Kronecker delta function. We also generalize

$$Q(x_j | \{x_i\}) = \lambda_j \left(\sum_{k \in \{i\}} \delta_{j,k+1} + \delta_{j,k-1} \right) \tau_{min}(\{x_i\}), \tag{3.15}$$

which defines the probability that microhabitat x_j is the first to fixate among unfixed microhabitats given that microhabitats $\{x_i\}$ have already reached fixation. With these definitions in place, $\tau_{\max}^\beta$ can now be written recursively as

$$\tau_{\max}^\beta(\{x_i\}) = \tau_{\min}(\{x_i\}) + \sum_{j \notin \{i\}} Q(x_j | \{x_i\}) \tau_{\max}^\beta(\{x_i\} \cup x_j). \tag{3.16}$$

Solving for the fixation times within this approximation scheme would require evaluating $\hat{\tau}_f = \tau_{\max}^\beta(\{\})$. However, this recursive definition is still not amenable to a closed-form solution because $\hat{\tau}_f$ has M terms of the form $\tau_{\max}^\beta(\{x_{i_1}\})$, with a single fixed microhabitat x_{i_1}, each of which has $M-1$ terms of the form $\tau_{\max}^\beta(\{x_{i_1}, x_{i_2}\})$ with both microhabitats x_{i_1} and x_{i_2} fixated. As we see, this recursion leads to a combinatorial growth in the number of terms corresponding to calculating the contribution to the fixation time estimate of every possible trajectory (i.e., every unique sequential ordering of microhabitats to fixate). Consequently, no closed-form solution to this equation can be written down. More details can be found in Appendix C.1.

Instead, we first use this recursive definition to bound the fixation time by the expected fastest and slowest fixation trajectory. Calculating the fastest or slowest trajectory greatly simplifies the expression for $\tau_{\max}^\beta$ since it is equivalent to choosing the smallest or largest exponential rate to determine the next microhabitat to fixate as opposed to considering every possible order. To form the lower bound, we assume that the microhabitat with the smallest exponential rate is the next to reach fixation deterministically. That is, we assume

$$\hat{\tau}_{\max}^\beta(\{x_i\}) = \tau_{\min}(\{x_i\}) + \tau_{\max}^\beta(\{x_i\} \cup x_{i_n}) \tag{3.17}$$

where x_{i_n} is the microhabitat most likely to reach fixation, given by

$$x_{i_n} = \operatorname{argmin}_{j \notin \{i\}} Q(x_j | \{x_i\}). \tag{3.18}$$

36
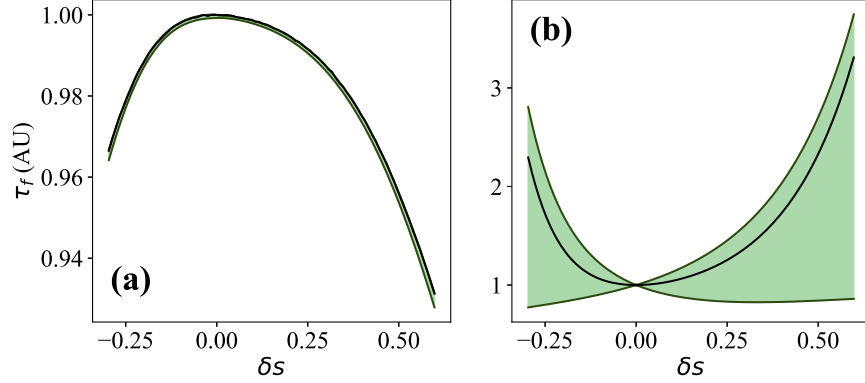
Figure 3.6: (a) The upper and lower bounds tightly constrain the fixation time across different landscapes with $\mu = 10^{-6}$, $\beta = 0.01$, $\langle s \rangle = 1/6$, and $M = 5$, when the fixation times are dominated by the arrival time of the first mutant. (b) The upper and lower bounds do not in general tightly constrain the fixation times with $\mu = 0.01$ and $\beta = 10^{-6}$, although the bounds tightly constrain fixation in landscapes that can be treated as perturbations of the homogeneous landscape.

This microhabitat is assumed to be the next to reach fixation deterministically, so that $Q(x_j|\{x_i\}) = \delta_{i_n,j}$. Upon using this in our recursion relation, there is only a single term at each depth when calculating the lower bound. A similar approximation can be performed to calculate the upper bound the fixation time. Because the upper and lower bounds are calculated using the generalized approximation for fixation rather than the MFPT calculations, quantitative agreement is only expected in a regime in which the original approximation is valid, i.e., when we have $\mu \ll \beta$ or $\beta \ll \mu$. More details on the generalized recursion relation and the procedure for calculating bounds can be found in Appendix C.2.

To evaluate the validity of the bounds in systems with $M > 3$, we first extended the MFPT calculation to a system with $M = 5$. The explicit transition rates with $M = 5$ can be found in Appendix C.3. In Figure 3.6(a) the bounds are plotted along with the exact value that would be obtained directly from the approximation for $\mu = 10^{-6}$ and $\beta = 0.01$. Note that in this limit, the fixation time is dominated by the arrival of the first mutant so that $\tau_f \simeq \tau_{\min}$. Consequently, the upper and lower bounds tightly constrain the fixation time for all possible landscapes characterized by δs. The bounds are shown for the limiting case $\beta \ll \mu$ in Figure 3.6(b), which was generated using $\mu = 0.01$ and $\beta = 10^{-6}$. The different trajectories leading to fixation now exhibit a much wider range of fixation times, which are no longer simply dominated by the arrival of the first mutant. However, the bounds still tightly constrain the fixation time in the neighborhood near $\delta s = 0$. Intuitively, when all of the microhabitats have very similar selection pressures, the order in which individual microhabitats achieve fixation is less important in determining the fixation time of the whole system.

37
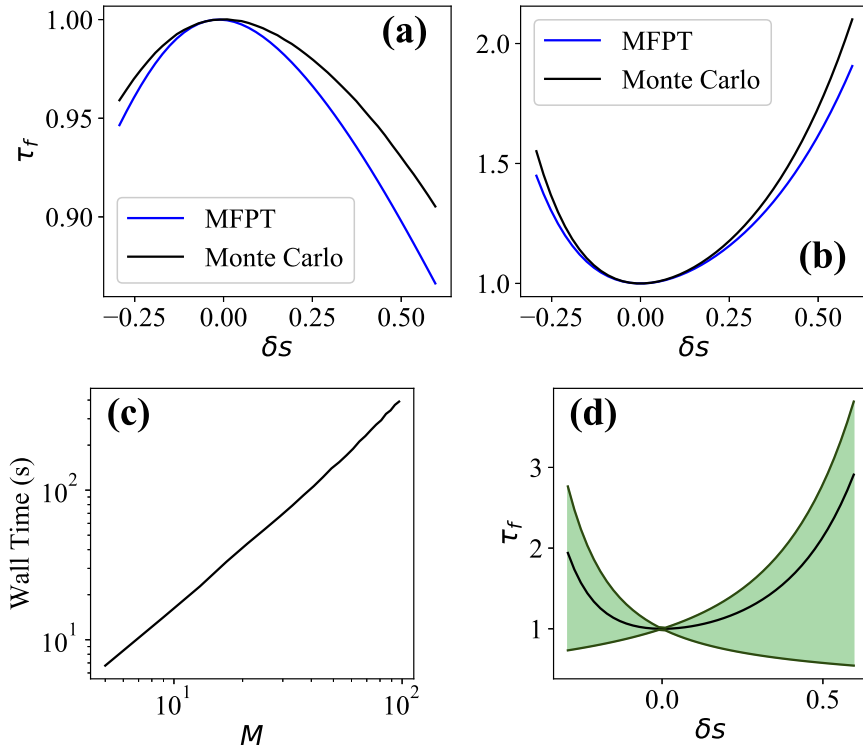
Figure 3.7: (a) The fixation times estimated using Monte Carlo capture the qualitative trend of the full MFPT calculation with $M = 5$ using $N = 15$, $\mu = 10^{-6}$, and $\beta = 0.01$. (b) In the opposite regime with $\mu = 0.01$ and $\beta = 10^{-6}$, the Monte Carlo estimation again reasonably matches the exact fixation times. (c) The wall time for generating 10^8 Monte Carlo trajectories as measured on a standard desktop computer scales according to a power law, allowing fixation times to be calculated in much larger systems than would be feasible than by exactly solving the MFPT equations. (d) Example of estimating spatial fixation times in a larger system using $M = 11$ along with the analytical bounds with $\mu = 10^{-2}$, and $\beta = 10^{-7}$. All fixation times calculated with $\langle s \rangle = 1/6$.

The bounds can be calculated for systems with an arbitrary number of microhabitats M, but we need a method capable of verifying the results. It is impractical to continue to extend the MFPT calculation for $M > 5$ since the number of equations that must be solved to calculate the fixation time scales exponentially with M. Instead, we return to our generalized fixation approximation (Equation 3.16) and use this expression to efficiently Monte Carlo sample individual trajectories towards fixation. To calculate the fixation time from a trajectory, the time required for each microhabitat to fixate is generated by sampling an exponential distribution with rate λ_i for microhabitat x_i. Initially, the microhabitat with the fastest fixation time is chosen to fixate in the corresponding amount of time. Exponential samples are then redrawn for the microhabitats neighboring the microhabitat previously chosen to fixate, since the exponential rate for these microhabitats has changed upon adding a fixated neighbor. At this point, the next microhabitat is chosen to fixate according to the random samples, and cumulative trajectory fixation time is updated. This process continues until all microhabitats have reach fixation, at which point the fixation time for the trajectory is recorded. By averaging over a large number of random trajectories, the mean fixation time can be estimated for a given selection landscape. The full details of this Monte Carlo sampling are discussed in detail in Appendix C.1.1. The estimated fixation times obtained using Monte Carlo sampling are compared to those obtained by solving the MFPT equations with $M = 5$ in Figure 3.7(a) and Figure 3.7(b) for $\mu = 10^{-6}$ and $\beta = 0.01$ and $\mu = 0.01$ and $\beta = 10^{-6}$, respectively. We observe qualitative agreement in both limiting cases of our parameter values. Because the Monte Carlo estimate is obtained by sampling the generalized approximation in Equation (3.16), the estimated fixation times agree exactly with those obtained by the explicit approximation of the MFPT calculation but only approximately with the MFPT fixation times. The utility of the Monte Carlo estimated fixation times comes from its ability to explore fixation in much larger systems than would be feasible by solving the MFPT equations. In Figure 3.7(c) the run time of the optimized algorithm on a standard desktop computer is plotted as a function of M. Because the run time scales as a power law, fixation can be calculated for systems with $M \gg 3$. As one example, the fixation times with $M = 11$ are plotted in Figure 3.7(d) along with the analytical bounds with $N = 25$, $\mu = 10^{-2}$, $\beta = 10^{-7}$, and $\langle s \rangle = 1/6$. Note that in this limit the fixation time is not simply given by the minimum of M exponential processes, so the agreement of the bounds is not trivial.

3.3.2   Understanding Fixation Time Scaling Laws

Using the Monte Carlo estimated fixation times, we now wish to explore how fixation times scale with M. Because we have demonstrated that the selection landscape plays an important role in determining the fixation times, we first need a rigorous method of generating the analogue of a given $M = 3$ selection pressure landscape for a given M. To preserve the overall structure of

the underlying landscape, we will restrict ourselves to odd M so that there exists a single center microhabitat with the largest or smallest selection pressure in the system of microhabitats. The selection pressure landscape is again assumed to be symmetric about this center microhabitat, and the selection landscape is assumed to be strictly monotonic between the leftmost microhabitat and the center microhabitat. For a general landscape, we require $\lceil M/2 \rceil$ individual microhabitat selection pressure values to fully define the selection landscape. With $M = 3$, these two parameters were the mean selection pressure $\langle s \rangle$ and the offset in the center microhabitat δs. In the larger system, we wish to conserve these values for $\langle s \rangle$ and δs. But we will need additional constraints to fully specify a landscape of general M. We choose to add the constraint that $s_0 = s(x_0)^{M=3}$ is also held constant. With this constraint, the $M = 5$ landscape is fully determined. For larger values of M, we add the constraint that the total variance in the difference in selection pressure values between neighboring microhabitats is minimized. This final assumption introduced fully determines the selection landscape with arbitrary M given the $M = 3$ landscape. One can show that the selection pressure in microhabitat x_j with $j < \lceil M/2 \rceil$ obeying these constraints is given by

$$s(x_j) = s_0 + \frac{4j(j+1)}{M^2 - 1}\delta s. \tag{3.19}$$

The derivation can be found in Appendix C.4. As an illustration of the resulting selection landscape, the landscape with $M = 9$ from this approach is shown in Figure 3.8. With a mapping between selection landscapes in $M = 3$ to arbitrary M, we can now examine how fixation times scale with M. This is a challenging question to answer in general without a closed-form solution for the fixation times in $M > 3$. Although fixation times can be Monte Carlo estimated for arbitrary M for each specific choice of parameters, a distilled understanding of the effect of M on the fixation times is not easily gained from the Monte Carlo estimates. Recall that in the $M = 3$ system, an intuitive understanding was possible because the majority of parameter space could be understood, at least qualitatively but often quantitatively, as either the expected minimum or maximum of $M = 3$ random exponential processes. We will again take this approach and focus on understanding how fixation times scale with the number of microhabitats in these two regimes.

We first turn to the limit where the fixation time is the minimum of M random variables. Physically, this limit assumes that $\mu \ll \beta$, so that the fixation time is dominated by the time required for the first mutation event. Therefore, the exponential rates given by Equation (2.5) for all microhabitats have no fixated neighbors, so that the rates reduce to

$$\lambda_i = N\mu \frac{s(x_i)}{1 - (1 - s(x_i))^N}. \tag{3.20}$$

40
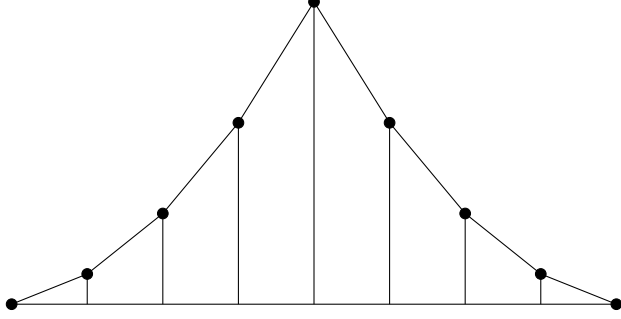
Figure 3.8: The minimal variance interpolated selection landscape with $M = 9$.

Using these exponential rates, we wish to derive an expression for the scaling factor

$$\mathcal{T}_{\min}^{M} \equiv \frac{\tau_f^M}{\tau_f^{M=3}}, \tag{3.21}$$

which relates fixation times with M microhabitats to those calculated in the $M = 3$ system we have studied extensively in the limit where fixation is the minimum of M random variables. After some algebra, one can write down

$$\mathcal{T}_{\min}^{M} = \frac{1}{1 + 2\xi N\mu\tau_f^{M=3}}, \tag{3.22}$$

with

$$\xi = \sum_{i=1}^{\lfloor M/2 \rfloor - 1} \frac{s_0 + \frac{4i(i+1)}{M^2-1}}{1 - \left(1 - s_0 - \frac{4i(i+1)}{M^2-1}\delta s\right)^N}. \tag{3.23}$$

The derivation of this expression as well as confirmation of agreement with the Monte Carlo sampled $\hat{\tau}_f$ can be found in Appendix C.5.1. Note that ξ depends on δs, so that the scaling factor is also a function of the specific selection pressure landscape. For the homogeneous landscape with $s(x_i) = s_0$, the scaling factor reduces to

$$\mathcal{T}_{\min}^{M} = 3M^{-1}, \tag{3.24}$$

independent of $\langle s \rangle$, μ, and N. In Figure 3.9(a) the scaling factor for the homogeneous landscape is plotted in black across two orders of magnitude of M. For each value of M, the range of scaling factors $\mathcal{T}_{\min}^{M}$ observed across all selection pressure landscapes is also plotted in the green region, which is bounded above by the landscape generated with the largest physical value for δs, plotted in the dark green line. The general trend of $\mathcal{T}_{\min}^{M}$ follows our intuition—as more microhabitats are
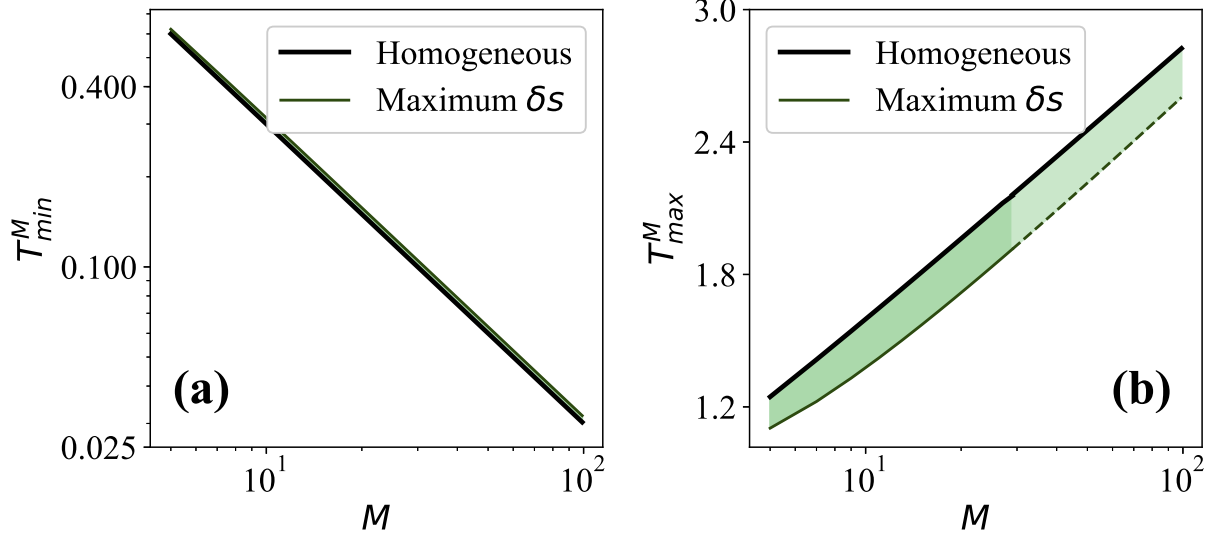
Figure 3.9: (a) The scaling factor relating the fixation times with a given M to those calculated with $M = 3$ in the limit where the fixation times are dominated by the minimum of M exponential processes scales with M according to a power law. The scaling factor for the homogeneous landscape and the landscape with the largest physical value of δs are shown in black and green, respectively, and the shaded region represents the scaling factors obtained with other values of δs. (b) The scaling factor relating the fixation times with a given M to those calculated with $M = 3$ in the limit where the fixation times are dominated by the maximum of M exponential processes scales logarithmically with M. The shaded green region is again the range of different scalings that can be obtained, and the lighter green region is formed by extrapolating the bounds calculated for $M \leq 29$. $\langle s \rangle = 1/6$ for all landscapes.

added to the system, the total number of cells increases and consequently the first mutant appears faster. The scaling factor decreases by more than an order of magnitude across the values for M, and the effect of different selection pressure landscapes has a much smaller impact on the fixation times than the number of microhabitats. It is also worth noting that the fixation time ratio is bounded below by the homogeneous landscape when fixation is the minimum of M exponential processes.

We now turn our attention to the limit where the fixation time is the maximum of M random variables. Physically, this limit corresponds to $\mu \gg \beta$, so that the fixation time of the system is simply the time to fixation in the microhabitat slowest to reach fixation. Interestingly, in the limit $\beta \to 0$, the exponential rates simplify to the same expression in Equation (3.20) even though we are in the opposite regime. We wish to write down an expression for the scaling factor

$$\mathcal{T}^M_{\max} \equiv \frac{\tau^M_f}{\tau^{M=3}_f},$$

(3.25)

which relates fixation times in the $M = 3$ system to those in a system with arbitrary M in the limit where fixation is the maximum of M random variables. The expression for the scaling factor in this limit is much more complicated and is given by

$$\mathcal{T}_{\text{max}}^M = \frac{\sum_{k=1}^{M} (-1)^{k+1} \Sigma_M^{(k)}}{\sum_{k=1}^{3} (-1)^{k+1} \Sigma_3^{(k)}}, \tag{3.26}$$

with

$$\Sigma_M^{(k)} = \sum_{i_1=1}^{M} \sum_{i_2 \neq i_1}^{M} \sum_{i_3 \neq i_2, i_1}^{M} \cdots \sum_{i_k \notin \{i_1, \ldots, i_{k-1}\}}^{M} \frac{1}{\sum_{j=1}^{k} \lambda_j}. \tag{3.27}$$

The derivation of this result as well as the confirmation of its validity can be found in Appendix C.5.2. This scaling factor cannot be further simplified for a general selection landscape, but one can show that with a homogeneous landscape it reduces to

$$\mathcal{T}_{\text{max}}^M = \frac{6}{11} H_M, \tag{3.28}$$

where H_M is the Mth harmonic number. In Figure 3.9(b) the scaling factor for the homogeneous landscape is again plotted in black across two orders of magnitude of M. For each value of M, the range of scaling factors $\mathcal{T}_{\text{max}}^M$ observed across all physical selection pressure landscapes is also plotted in the light green region, bounded below by the landscape with the largest physical value for δs, plotted in the dark green line. Note that expression for $\mathcal{T}_{\text{max}}^M$ is only evaluated up to $M = 29$ due to the number of terms in Equation (3.26) with large values of M. The scaling factor for the landscape generated with the largest physical value of δs is obtained for larger values of M by fitting the explicitly calculated expression for $M \in [5, 29]$. The general trend of $\mathcal{T}_{\text{max}}^M$ also follows our intuition—as more microhabitats are added to the system, the time required for the last one to reach fixation increases. It is also worth noting that the fixation time ratio is bounded above by the homogeneous landscape when fixation is the maximum of M exponential processes. We observe that the number of microhabitats has a much weaker effect on the scaling factor when fixation can be approximated as the maximum of M exponential processes than to the scaling factor when the fixation can be approximated as the minimum of M exponential processes. Additionally, the role of the selection pressure landscape is no longer negligible in determining the scaling factor for a given M—the selection pressure landscape in this regime affects not only $\tau_f^{M=3}$ but also the scaling factor $\mathcal{T}_{\text{max}}^M$.

3.4 Conclusion

Upon both considering imperfect mutations and larger systems, we generalized our results to incorporate observations studying *in vitro* microbial populations. Adding the possibility of imperfect mutants and systems with more microhabitats has not only supported but also strengthened many of the conclusions drawn from the more tightly constrained model. Spatial heterogeneity continues to play an important role in determining the fixation times in a spatially-extended system, and this effect is amplified when the mutations conferring resistance have an associated fitness cost. So much so, in fact, that a fitness cost can prohibit fixation in finite time in the deterministic limit. We also saw the importance of conserving mean selection pressure with a fitness cost, and we used results from the deterministic limit to aid our understanding of the evolutionary dynamics in the finite-sized system. Spatial heterogeneity continued to impact fixation times in larger systems, and we characterized how the fixation times scale as a function of the number of microhabitats in the system. Upon relaxing two of the key assumptions with the earlier model, we have not only confirmed the importance of spatial heterogeneity on the emergence of resistance but also discovered scenarios in which the importance becomes amplified as our toy model better approximates *in vitro* systems. Admittedly, the generalized model is still artificial in many important regards. However, we believe that the results of this generalized model strengthen the case for looking at the effect of spatial heterogeneity on the emergence of resistance in more complex models and even *in vitro* evolution experiments.

CHAPTER 4

Modulating Resistance Emergence in Realistic Microbial Models

4.1 Introduction

The emergence of drug resistance in microbial populations is a complex phenomenon that depends critically on spatially-varying environmental factors. Most studies on drug resistance have focused on characterizing the molecular pathways of resistance or the effects of different treatment regimes on resistance evolution. While most studies assume complete spatial homogeneity, several recent studies indicate that spatial gradients in drug concentration can modulate resistance in both experimental [38] and theoretical models of evolution [32]. However, variation in drug concentration is only one of several factors affecting growth and selection pressure in microbial communities. Furthermore, spatial heterogeneities likely extend beyond simple gradients, raising the question of how resistance evolution is affected by growth profiles with potentially non-monotonic spatial dependencies.

As one example, temperature affects both the native growth rate of bacteria [84, 85] as well as the minimum inhibitory concentration (MIC) of several antibiotics. Although the temperature-dependence on growth rates is a common consideration for microbial experiments, previous work indicates that 30% of bacteria exposed to penicillins had minimum inhibitory concentrations (MICs) at least four times lower at 41 degrees compared to 35 degrees [86]. This suggests that even small, physiologically relevant temperature variations could alter the MIC of common antibiotics as well as the drug-free growth rate of bacteria. Thus, even simple gradients of temperature can potentially generate complex spatial profiles of growth and selection pressure. In this work, we investigate the effects of spatial temperature gradients—and, more generally, spatially non-monotonic growth profiles—on the emergence of resistant mutations.

Additionally, we investigate more generally the effect of spatially heterogeneous growth profiles on the emergence of resistance in spatially-extended systems. Previous theoretical work has demonstrated that the evolutionary dynamics can be significantly affected by different spatial

growth profiles[32], but the system was not studied across different spatial growth profiles conserving mean growth rate. Existing work conserving mean growth has focused on much smaller systems governed by very simple rulesets [81]. In this work, we use a biologically-motivated model of evolutionary dynamics that allows for stochastic mutation, migration, and genetic drift, in addition to spatially-varying population sizes and more complex genotypes than have been studied in the simplified system. In this manner, we can understand how well the results using a toy model of evolution generalize to larger systems with evolution governed by more complicated dynamics.

4.2 Developing Large-Scale Microbial Resistance Model

We consider a growing population of cells in a linear array of L connected microhabitats. Cells migrate to neighboring microhabitats at a rate β. Each cell has a genotype $g \in [1, G]$ that confers some degree of antibiotic resistance, with higher values of g corresponding to a larger MIC to some uniform drug concentration c present in every microhabitat. This choice of genotype representation is motivated by experimental observations that complete resistance can often require multiple sequential mutations [21, 36]. For simplicity, we will first look at the temperature-independent model. Within a given microhabitat i, cells of genotype g replicate at a rate $\phi_g(c)(1 - N_i/K)$, where $\phi_g(c)$ is the growth rate of genotype g, N_i denotes the total population of cells in the ith microhabitat, and K is the carrying capacity of the microhabitat, taken to be a constant across all microhabitats for this work. Upon replication, a cell mutations with probability μ and migrates with probability β. We consider asymmetric mutations, so that a cell of genotype g can mutate to genotype $g + 1$ with probability μ, and symmetric migration, so that cells from microhabitat i migrate to microhabitat $i \pm 1$ with probability $\beta/2$. Lastly, cells die at a fixed rate d, so that at steady-state there still exists a turnover of cells. The units of time are fixed by defining the maximal growth rate to be $\phi_g(0) = 1$.

For all cases, we assume that cells with genotype $g = 1$ are least resistant to the epistatic drug, while cells with $g = G$ have the highest resistance. The growth rate of a bacteria of genotype g is assumed to depend on the MIC $m(g)$ according to

$$\phi_g(c) = \max\{0, 1 - (c/m(g))^2\}, \tag{4.1}$$

as used by [32] and originally inspired by [87]. We assume that the MIC monotonically increases with g and thus has no genotypic fitness valleys. Explicitly, we assume $m(g) = 4^{g-1}$. Thus far, our model closely resembles the work of [32]. We now add in spatial heterogeneity to the system assumed to have uniform drug concentration and no fitness valleys in genotype-space. We will work with the motivating factor of temperature, but there are a number of spatially-dependent factors

that can affect bacteria growth rates. In addition to the MIC temperature dependence motivated by [86], we let the native growth rate also be a function of temperature. For this more general case, we now have

$$\phi_g(c, T) = \max\left\{0, \ \xi(T)\left[1 - \left(\frac{c}{\varphi(T)4^{g-1}}\right)^2\right]\right\}, \tag{4.2}$$

where $\xi(T) \in [0, 1]$ is the normalized growth coefficient and $\varphi(T) \in [0, 1]$ is the normalized MIC coefficient. With arbitrarily complicated forms for $\xi(T)$, $\varphi(T)$, and the spatial temperature dependence, we can generate arbitrarily complicated forms for $\phi_g(c, T)$.

As an explicit example, we can analyze our system with a linearly decreasing profile for $\xi(T)$ and a linearly increasing profile for $\varphi(T)$. These spatial profiles as well as the resulting growth curve $\phi_g(T)$ are shown in Figure 4.1. Even with simple linear profiles for both $\xi(T)$ and $\varphi(T)$, we can produce complicated nonlinear forms for $\phi_g(T)$. To continue with the example, we can simulate the evolution of the system with these functions for $\xi(T)$ and $\varphi(T)$. The model was simulated using a kinetic Monte Carlo algorithm introduced by [32]. This algorithm is not unlike a standard Gillespie algorithm [88] but is more computationally efficient for this particular problem due to the simplicity of the atomic rules and the high fraction of timesteps resulting in a state change. Additional details about our implementation of the algorithm can be found in Appendix D.1. The simulation was run assuming $L = 500$, $K = 100$, and $c = 0.4$, which represent our standard system parameters for all of the results shown. Snapshots of the simulation are shown in Figure 4.2 to illustrate the emergence of mutants with increased resistance. The microhabitats toward the left side of the array have a very small growth rate with $g = 1$, but mutants can grow in these microhabitats. This gives mutants a significant selection advantage over the wild-type cells and leads to mutants first thriving, although not necessarily arriving, in the left-most microhabitats in the evolution snapshots.

Our choice of the functional form and parameterization of $\xi(T)$ and $\varphi(T)$ in this example were arbitrary. To constrain our analysis, we will assume (1) temperature linearly increases along the array of microhabitats and (2) both $\xi(T)$ and $\varphi(T)$ have a single local maximum in the temperature region of interest. As proved in Appendix D.2, the most general curve for ϕ_g obeying these constraints can have at most two local peaks. All forms of ϕ_g with a single peak will produce behavior qualitatively similar to that observed in Figure 4.2, but adding an additional peak allows the possibility of new behavior. For simplicity, we look at the case with symmetric peaks in ϕ_g. For a given displacement Δx and standard deviation σ, the spatial growth profile for the wild-type cells would take the form

$$\phi_1 = \mathcal{N}(0.5 - \Delta x, \sigma) + \mathcal{N}(0.5 + \Delta x, \sigma). \tag{4.3}$$

47

(a) Curves for $\xi(T)$ and $\varphi(T)$ (b) Resulting ϕ_g
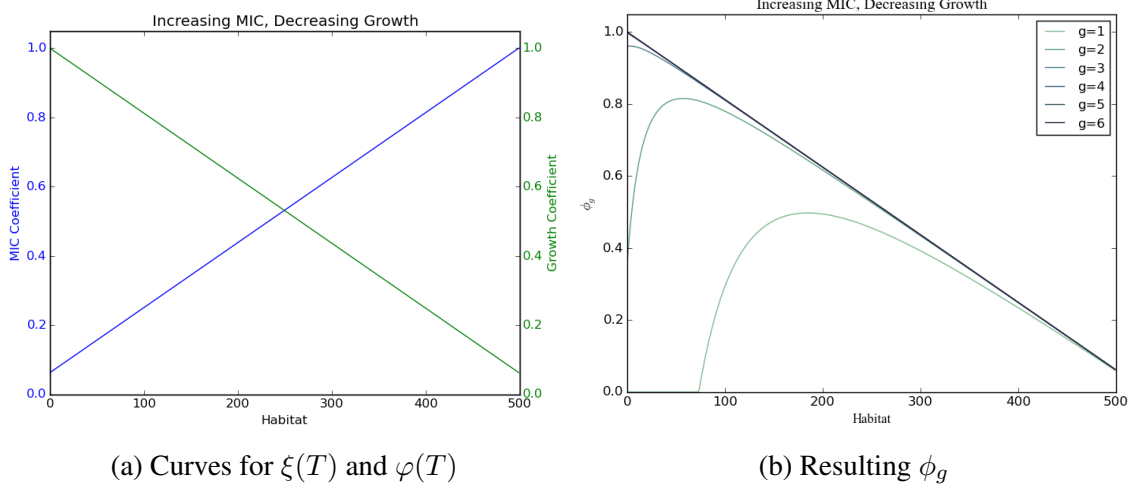
Figure 4.1: (a) Example of linear profiles for both $\xi(T)$ and $\varphi(T)$ resulting in (b) a nonlinear growth curve ϕ_g.

An example of ϕ_g calculated with displacement $\Delta x = 0.35$ and width $\sigma = 0.2$, where we are working in normalized coordinates so that $x \in [0, 1]$, is shown in Figure 4.3. By adjusting the spacing between the means of the two distributions, we can observe how different spatial configurations affect the emergence of mutations. One important consideration is the total spatially integrated ϕ_g—this sets the total amount of growth for a given ϕ_g. To have a meaningful comparison, we must ensure that this integral is conserved as we vary the distance between the peaks. This allows us to isolate the effect of the spatial distribution of growth and comment on the effect of spatial heterogeneity without the compounding factor of mean spatial growth rates. This normalization represents an important constraint that has not been studied in this system.

4.3 Analyzing Fixation

4.3.1 Effect of Spatial Heterogeneity

Initially, we will restrict our analysis to a two-state system ($G = 2$) with perfect mutants, so that ϕ_1 is parameterized by Δx and σ and $\phi_2 = 1$. This is the same assumption we made in the simple toy model of evolution, in which a single mutation separated the sensitive wild-type cells from a fully-resistant strain, but the evolution is now governed by more complex transitions. We will first fix the parameters $\mu = 5 \times 10^{-6}$ and $\beta = 0.1$ to understand the effect of spatial heterogeneity in this system. Using the simple two-state system, we can simulate evolution for a given ϕ_1 and measure T_{50}, defined as the time required for the mutant strain ($g = 2$) to comprise half of the total cells in the population. Each stochastic simulation results in a random trajectory taken to reach

(a) $g = 1$ Steady-State

(b) $g = 2$ Mutant Emerges

(c) $g = 2$ Proliferation

(d) $g = 3$ Proliferation
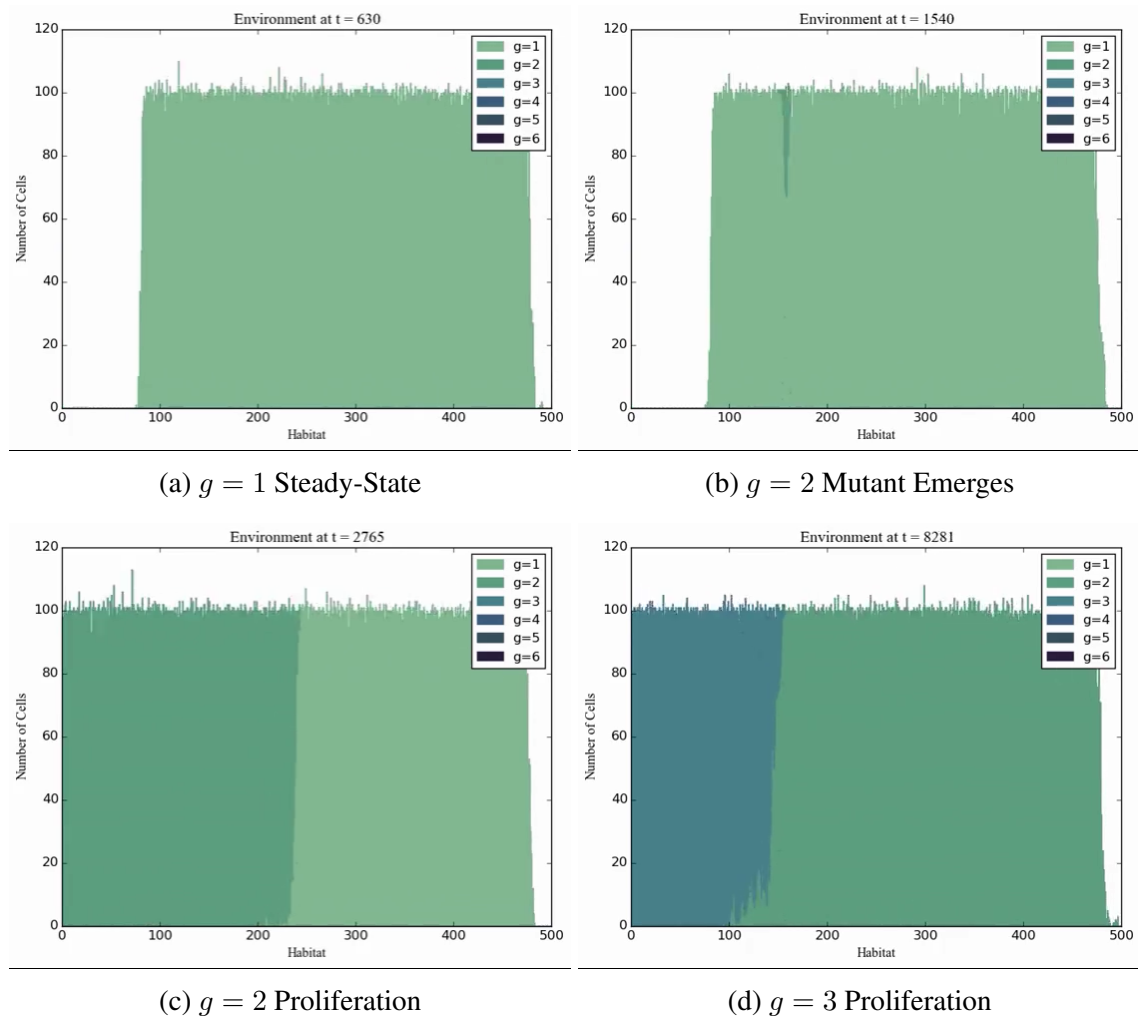
Figure 4.2: Snapshots of the kinetic Monte Carlo simulation with ϕ_g as shown in Figure 4.1b. Note that the asymmetry in ϕ_g manifests as a higher mutation composition in the habitats with a higher growth rate.
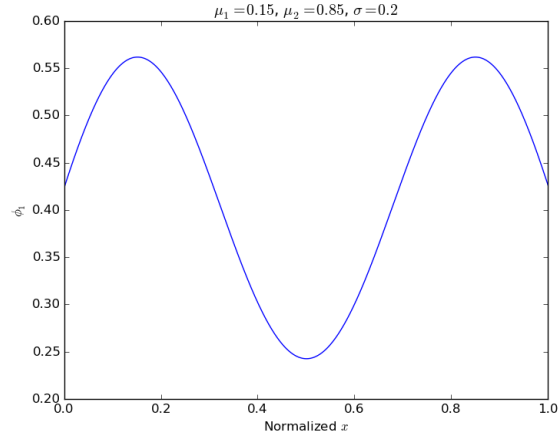
Figure 4.3: Sample curve for ϕ_1 with $\Delta x = 0.35$ and $\sigma = 0.2$ in normalized coordinates.

fixation. By performing many such simulations, we can form estimates for $\langle T_{50} \rangle$ and $\sigma_{T_{50}}$. The raw trajectories of the fraction of the wild-type cells for 560 simulations is shown in Figure 4.4. Also plotted are the calculated mean value of T_{50} as well as its standard deviation.
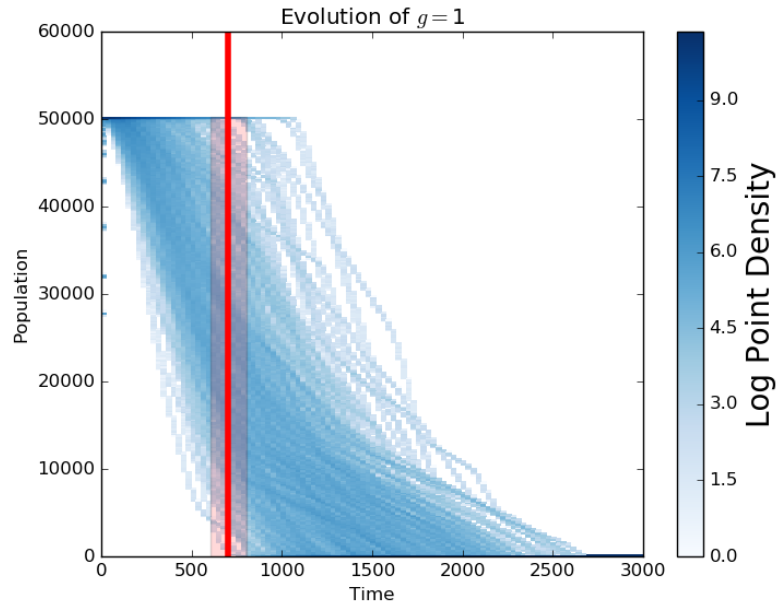
Figure 4.4: Each trajectory represents the path a single simulation of the microhabitat array. By combining 560 simulations, we can calculate the average value for T_{50}, shown as a red vertical line, as well as its standard deviation, shown as a pink region about the mean.

These trajectories can be simulated for different values of Δx in the wild-type growth profile to calculate these first two fixation time distribution moments for different landscapes. As the separa-

50

(a) Spatial Fixation Times   (b) Binning ϕ_1
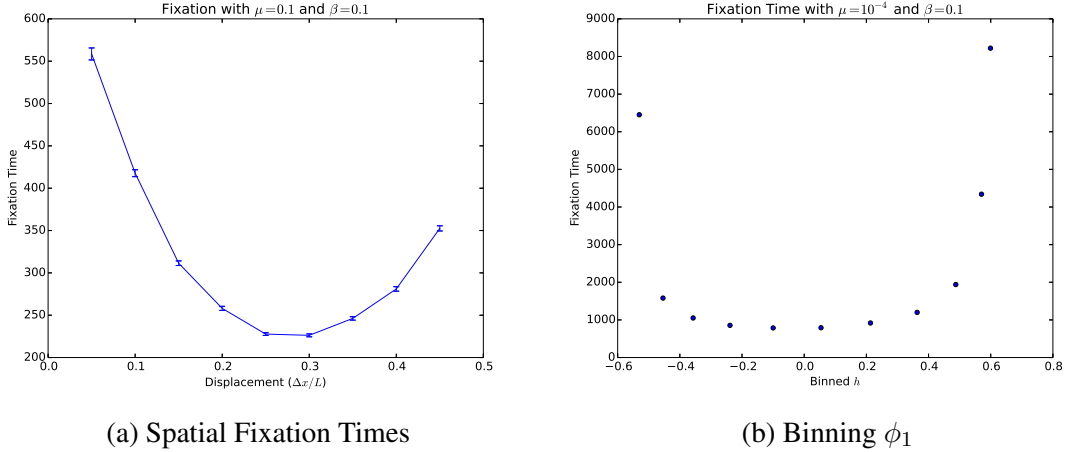
Figure 4.5: (a) Different spatial landscapes, formed by changing Δx, produce different fixation times despite conserving $\langle \phi_1 \rangle$. (b) Binning ϕ_1 allows us to roughly characterize each spatially-extended landscape by a single value of h.

tion Δx between the two Gaussians in ϕ_1 is changed, the spatial growth profiles are normalized to conserve $\langle \phi_1 \rangle_x$. As one example of this procedure, the spatial fixation times extracted for a system using $\mu = 10^{-4}$, $\beta = 0.1$, and $\sigma = 0.2$ are shown in Figure 4.5a. We can observe that different landscapes produce different fixation times, even though $\langle \phi_1 \rangle$ is conserved across all landscapes. We can also relate the ϕ_1 in the spatially-extended system with $L = 500$ to the simple fitness landscape parameterized by B and h in the $M = 3$ system. For a given growth curve ϕ_1, the microhabitats can be grouped into three bins, over which ϕ_1 is averaged, to calculate the analogous values for B and h. The fixation times as a function of h are shown in Figure 4.5b.

We first look at how the fixation times across different landscapes depend on the spatial growth profile characterization. We fix $\mu = 5 \times 10^{-6}$ and $\beta = 0.1$ and vary σ, which controls the width of the Gaussians in ϕ_1. Additionally, we compare these fixation times to those obtained with a spatially homogeneous growth profile with $\phi_1(x_i) = \langle \phi_1 \rangle_x$. The first two moments calculated with $\sigma = 0.2$ from 560 stochastic simulations are plotted in Figure 4.6. We see that the fixation times are again affected by the choice of Δx. And interestingly, the fixation times can be either faster than the spatially homogeneous times (small Δx) or slower (large Δx). This behavior is somewhat intuitive—when there is no displacement, we have a single Gaussian curve and thus a single peak. There are a relatively small number of cells with significant selection pressure, so increasing the spacing between the Gaussians allows a larger number of cells to experience a significant selection pressure upon mutating. As the two Gaussians move apart, migration of cells between the two peaks is reduced. Since we have a constant death rate of $d = 0.1$, habitats with $\phi_1 < d$ will be ineffective at allowing mutations to migrate. As Δx increases, the two peaks become more

51

independent. If $\Delta x/L$ is significantly larger than σ, we therefore require two separate mutations before the number of mutants can grow to 50% of the population.
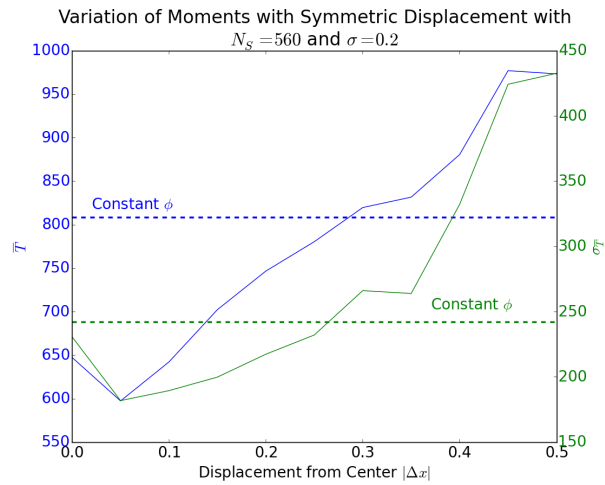
Figure 4.6: As the displacement of the mean of the symmetric Gaussians is increased, both the average value for T as well as its standard deviation decrease before increasing. Interestingly, both the average as well as the standard deviation for the scenario with a spatially varying ϕ_1 can either increase or decrease relative to the scenario with a spatially homogeneous ϕ_1, shown as a dashed line.

Similar results were compiled for Gaussian peaks with different widths σ and $\sigma = 0.3$. With the Gaussians parameterized by $\sigma = 0.1$ (Figure 4.7), we see the same type of behavior and the reasoning above still holds. Interestingly, though, the spatially heterogeneous ϕ_1 curves have strictly larger values for the first and second moments of T_{50} compared to a spatially heterogeneous profile. However, once the standard deviation of the Gaussians is increased to $\sigma = 0.3$ (Figure 4.8), we no longer observe minima in the first and second moments of T_{50} at a spatially displaced ϕ_1. While the spatially heterogeneous ϕ_1 can again produce moments that are either higher or lower than the spatially homogeneous case, increasing the separation between the peaks always increases the moments. As a brief note, we can only compare the trends between the different Gaussian standard deviations as the area of ϕ_1 is only conserved for a given standard deviation (which explains why the constant ϕ_1 values are different in the three plots).

These results suggest the importance of not just Δx, analogous to h (or δs) from the toy model, but also the importance of σ, analogous to $\langle s \rangle$, in determining the specific effect of spatial heterogeneity in this more realistic model.We have seen that for a fixed set of system parameters μ and β, different parameterizations of ϕ_1, characterized by σ, have led to a system where fixation is strictly slowed slowed by spatial heterogeneity (Figure 4.7), generally sped by spatial heterogeneity (Figure 4.8), and also is dependent on the specific distribution of spatial heterogeneity (Figure 4.6).
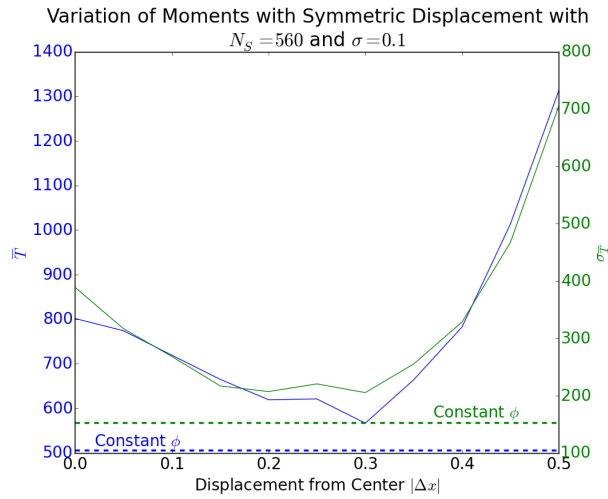
Figure 4.7: When the standard deviation of the Gaussians is decreased to $\sigma = 0.1$, the trends observed with $\sigma = 0.2$ are more pronounced.
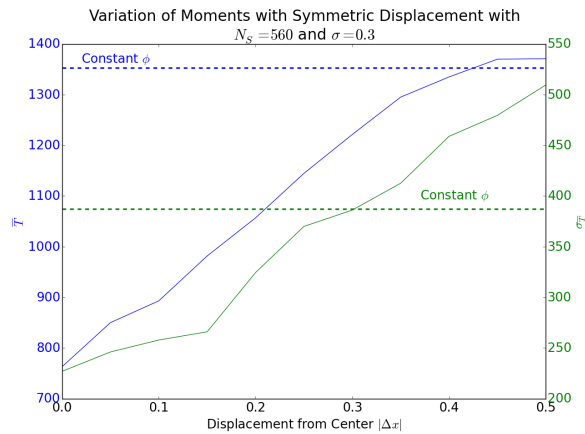
Figure 4.8: When the standard deviation is increased to $\sigma = 0.3$, however, the trend becomes monotonic.

(a) $\mu = 0.1$, $\beta = 0.1$ (b) $\mu = 10^{-6}$, $\beta = 0.1$
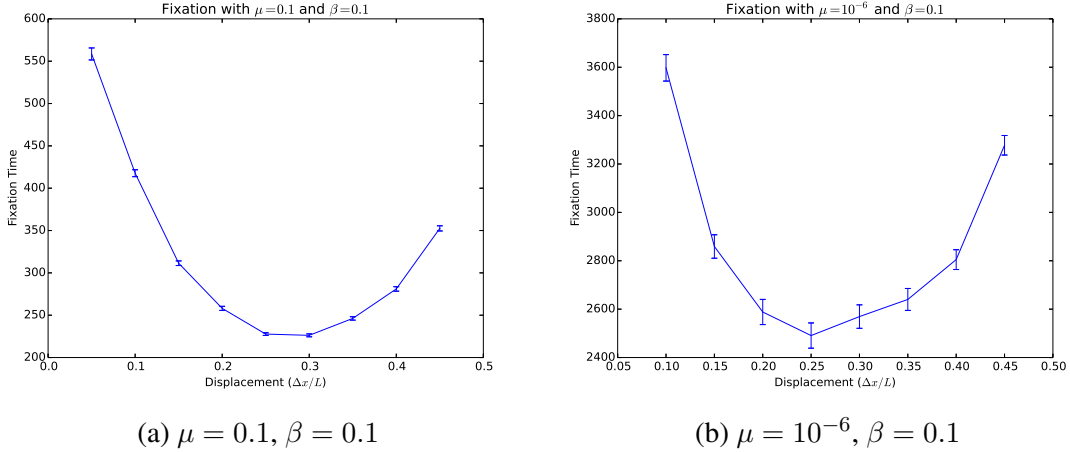
Figure 4.9: The same qualitative spatial trend in fixation time is found across a large range of parameters.

4.3.2 Effect of System Parameters

We observed these same three regimes in the toy model of evolution. For this specific example, however, the different regimes are achieved by fixing the system parameters (μ and β) and varying the parameterization of ϕ_1, suggesting that the reasoning behind the manifestation of spatial heterogeneity may be different in this model.

We first look at the fixation times calculated in this model for different system parameters. In general, we observe that the system parameters have a much smaller effect on the fixation times compared to the parameterization of ϕ_1. This is illustrated in Figure 4.9, which compares the T_{50} fixation times obtained with $\mu = 0.1$ and $\beta = 0.1$ to those obtained with $\mu = 10^{-6}$ and $\beta = 0.1$, assuming $\sigma = 0.2$. Note that although the overall scale of the fixation times changes with the parameters, the overall spatial trend is qualitatively very similar in both of these two cases, even though μ has changed by five orders of magnitude.

Compared to the smaller system, the range of qualitative behavior across parameter values is significantly reduced. This is because due to the effective mutation and migration rates differing between the two systems, even if μ and β are fixed. In general, we require $M - 1$ indepedent migration events for a mutant in the left-most microhabitat to reach the right-most microhabitat. For this reason, the spatial diffusion of mutants through migration becomes more suppressed as the number of microhabitats increases for a fixed migration rate. Additionally, this model considers migration as a behavior distinct from replication, unlike the toy model which only permitted migration upon replication. This further differentiates migration in the two different frameworks.

With the system sizes that we study using this kinetic Monte Carlo simulation, we would require an extremely small value of μ to probe the regime in which the fixation times are dominated by
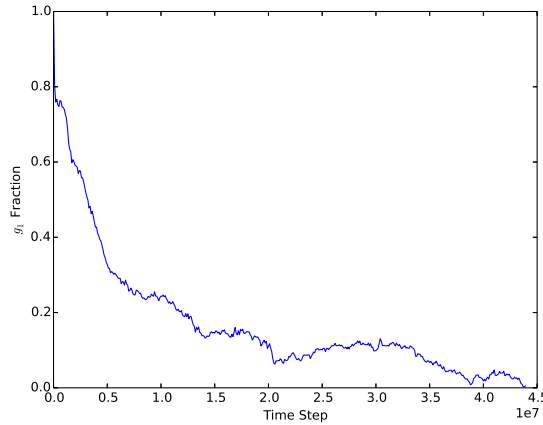
Figure 4.10: In the raw trajectory of the fraction of wild-type cells g_1, the majority of the fixation time is spent with $g_1 < 0.2$. Parameters: $\beta = 0.1$, $\mu = 10^{-4}$, $L = 500$, $K = 5$, $\Delta x = 0$, and $\sigma = 0.2$.

the arrival of the first mutant. For such small mutation rates, this simulation is very inefficient, since many timesteps without mutation will be simulated. Furthermore, the fixation time in this regime is easily approximated with great accuracy by modeling *de novo* mutations as exponential processes (Appendix A.7), so there is little reason to simulate the system when the fixation time is the minimum of M independent fixation events. We can confirm that this is the case in our system by examining the proliferation timescale, an idea introduced with the toy model. As an extreme example of the significance of this timescale, we look at a single fixation trajectory for a two-state system with $L = 500$ microhabitats and $K = 5$ with system parameters $\beta = 0.1$ and $\mu = 10^{-4}$ shown in Figure 4.10. Despite $\mu \ll \beta$ and having a very small carrying capacity, the time required for the wild-type ($g = 1$) genotype fraction to be reduced to 20% of the population is much less than the time required for this small wild-type subpopulation to become extinct. This confirms that the proliferation timescale is rarely negligible in this larger system, and, as a result, the mutation rate and migration rate play a reduced role in determining the fixation times, although the overall timescale is still determined by μ. For this large spatially-extended system, we instead observe that the specific parameterization of ϕ as well as the number of microhabitats L and population size N become much more important factors.

4.3.3 Effect of Spatial Monotonicity

The importance of the parameterization of double-peaked wild-type growth profile naturally motivates studying alternative expressions for ϕ_1. To construct a monotonic growth profile, we can

55
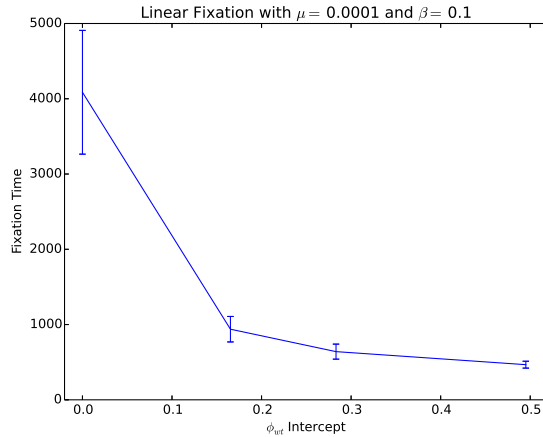
Figure 4.11: Increasing the y-intercept of the linear ϕ_{wt} leads to a faster fixation time. Increasing the intercept reduces the number of microhabitats with small selection pressure.

assume a simple linear relationship given by

$$\phi_1(x) = b + \alpha x, \tag{4.4}$$

where b is the growth at $x = 0$ and α is the spatial slope of the growth rate. We can vary b and determine α by imposing conservation of $\langle \phi_1 \rangle_x$. Using $\mu = 10^{-4}$ and $\beta = 0.1$, we find in Figure 4.11 that the fixation times decrease as the system becomes more homogeneous spatially.

In general, we can understand the role of heterogeneity on the fixation times by comparing the range of calculated fixation times using both the non-monotonic Gaussian parameterization of ϕ_1 and the linear parameterization of ϕ_1. For the simple two-state system, we see in Figure 4.12a that both parameterizations lead to a similar range of possible fixation times. However, if we return to the more complicated model with $G > 2$, we observe qualitatively different results if we define the fixation time T_{50} as the time required for genotype $g = G$ to constitute 50% of the population. With $G = 6$ and the MIC given by $m(g) = 4^{g-1}$, we find that the fixation times are much slower in the spatially non-monotonic system (Figure 4.12b).

This result is not unexpected given what we have learned about evolution in this spatially-extended system. We have seen that even in the two-state sytem, the fixation times are dominated by τ_2, the proliferation timescale. Depending on the specific form of ϕ_g, there can exist many microhabitats in which cells cannot thrive until a certain genotype is reached. As an explicit example, we return to the left-most microhabitats for the ϕ_g plotted in Figure 4.1b. Even in less extreme cases, mutants conferring additional resistance are most likely to thrive in microhabitats in which they offer a significant selection advantage over the wild-type cells (Figure 4.2). With a simple linear gradient in ϕ_g, sequential mutants most often arise in microhabitats near this boundary. Evo-

(a) $G = 2$ System   (b) $G = 6$ System
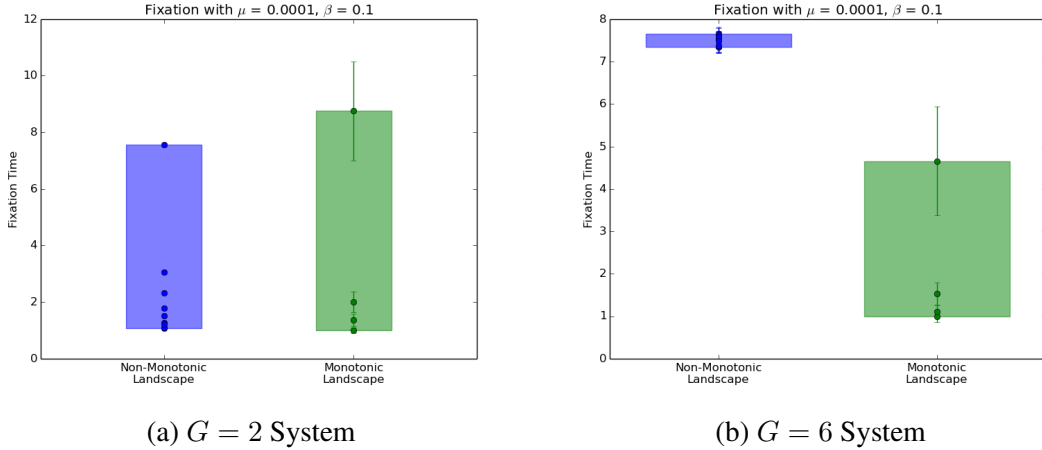
Figure 4.12: The monotonic and non-monotonic landscapes produce a similar range of fixation times in the two-state system, but the monotonic landscape reaches fixation significantly faster in the multi-state system with $G = 6$.

lution in these systems is "directed," in the sense that there exists a natural direction for sequential mutations to appear. This intuition has been confirmed rigorously by previous work studying this system with monotonic growth profiles [32].

4.4 Conclusion

Although the model of evolution that we developed remains a drastic simplification of the evolution process of *in vitro* microbial communities, it does allow us to better understand the emergence of resistance in spatially-extended systems. By moving beyond semi-analytical methods, we were able to relax several key assumptions in the toy model of evolution to better understand evolution under more complex dynamics. First, this model supports microhabitats in which the wild-type cells cannot survive, requiring one over even several mutations before they can be inhabited. Second, the each microhabitat has a constant carrying capacity but not population size. Third, the spatially-extended nature of the system allows for a less trivial consideration of the importance of different spatial growth profile distributions with regards to monotonicity. And fourth, it allows us to probe the effect of multiple advantageous mutations before complete resistance is achieved.

With this more general model, we were able to observe some of the same ideas previously observed in the toy model. Principally, the fixation times are still affected by the spatial distribution of growth, even while the mean growth rate is conserved. We also saw evidence of the same three regions: (1) heterogeneity always speeds fixation, (2) heterogeneity always slows fixation, and (3) heterogeneity can speed or slow fixation, depending on the spatial distribution. However,

the reasoning behind these different regimes was very different, since the proliferation timescale dominates the time to fixation in the spatially-extended system. We also saw how the effect of spatial heterogeneity can be amplified by both the specific parameterization of the growth profiles as well as the monotonicity of those profiles. This model builds upon our previous work to allow us to better understand the different mechanisms by which the environment can shape the emergence of evolution in microbial populations and reinforces the importance of spatial heterogenity on evolution dynamics. In the future, the results obtained with this more general model may motivate future spatial drug dosing strategies to control the emergence of drug resistance in spatially-extended systems.

CHAPTER 5

Measuring Resistance Emergence in Turbidostat Experiments

5.1 Introduction

Ultimately, we wish to understand how resistance develops not in simplified models of evolution but in *in vitro* microbial populations. Even in batch culture experiments performed under carefully controlled conditions, evolutionary dynamics are greatly complicated by the continual environmental changes as cells replicate, consume nutrients, and excrete waste [89], and these environmental changes can have profound effects on cellular physiology [90, 91, 92]. To mitigate these concerns, continuous culture devices have been used to create quasi-static environments. To culture cells in a quasi-static environment, inoculated growth media is continually diluted with fresh media. With proper tuning, these continuous culture devices dilute cell density and waste products at the same rate that they are being produced to create an unchanging environment [93, 94]. We are particularly interested in evolution experiments performed using turbidostats, a class of continuous culture devices that measure the cell density and apply feedback upon it. Because the dilutions are determined by the cell density measurements, the population sizes can be carefully controlled while the cells are growing without nutrient limitation, i.e., in exponential phase. A number of readily available turbidostat platforms exist to promote experiment reproducibility [95, 96].

Even in conceptually simple evolution experiments, however, resistance can be challenging to measure and quantify. And few experiments involving *in vitro* bacteria are easy to perform in practice due to the careful practices required to avoid contamination and generate reproducible results [97]. Reproduciblity even in well-controlled experiments is always a concern, as evolution has a large degree of variability due to stochasticity in both the emergence of *de novo* mutations and the genetic drift of existing mutant populations [98]. A related challenge is meaningful parameter estimation when the underlying evolution data has significant irreducible noise [99, 100, 101].

We are interested in studying the evolution dynamics of a population of microbes in exponential phase. As wild-type cells exposed to some bacteriostatic agent evolve, mutants can arise that

eventually dominate the population due to their higher fitness values than the sensitive, wild-type cells. With large microbial populations, the population is rarely truly composed of a single genotype [102]. Rather, many genotypes exist at very small population fractions (10^{-8}–10^{-6}), creating a mixed population with genotypic diversity. Due to this genotypic impurity, populations of certain bacteria may already contain the mutants resistant to the drug present [103]. Instead of relying on *de novo* mutation events, mutant fixation can be achieved from existing mutant cells slowly taking over the population through their selective advantage. In this manner, the genetic diversity permits evolution experiments that have reproducible evolution curves, since they are not relying on rare, Poisson-like *de novo* mutation events.

These microbial evolution experiments are often studied using a turbidostat or similar instrument. A turbidostat typically uses optical density (OD) measurements to determine the dilutions required to keep turbidity, and therefore population size, quasi-static. In practice, the turbidity is not strictly held constant but allowed to oscillate between a lower threshold and upper threshold. When the measured OD passes the specified upper threshold, a dilution is performed to "reset" the population back to the specified minimum optical density threshold. The lower and upper thresholds have some non-zero difference to prevent the dilutions to be determined from noise on the optical density measurements, so that the population size is only quasi-static. To understand evolution under these quasi-static conditions, we require a careful procedure for estimating parameters from data recorded during turbidostat operation.

With a careful parameter estimation procedure, we can study evolution in a turbidostat and hope to identify subtle shifts in the fixation dynamics introduced from spatial heterogeneity. The toy model of evolution is a natural candidate to study experimentally due to the ease of translating the model assumptions into turbidostat conditions. For example, the Moran-like evolution assumption requires constant population sizes, which are naturally studied using a turbidostat. The toy model also assumed three microhabitats evolving concurrently, which one can easily mirror experimentally using three vials in a turbidostat. Additionally, turbidostat experiment already contains many pumps between microhabitats for pumping in fresh media and pumping out media to perform dilutions. These pumps can be retrofitted to perform migration between the subpopulations growing in different vials. In this work, we lay the groundwork for the type of dynamics we expect to be recorded during a turbidostat experiment and the subsequent quantification of the evolution from the recorded data using an advanced Bayesian parameter estimation pipeline.

5.2 Modeling Idealized Turbidostat

5.2.1 Perfect Turbidostat

For bacteria growing in exponential phase in a turbidostat, we can model the population as fluctuating in time between some value $N_{\min}$, determined from the lower optical density threshold, and $N_{\max}$, determined by the upper optical density threshold. We treat the population as a collection of independently evolving subpopulations, each of which contains cells of a single genotype, growing deterministically in exponential phase, growing unabated until the total population size reaches $N_{\max}$. The different genotypes can be introduced into the population either through *de novo* mutation or from existing in small numbers in the initial bacteria stock. Due to the high degree of reproducibility observed in many turbidostat experiments, there is some evidence that the mutants that dominate the population largely are introduced during inoculation. For this reason, we neglect mutation events in our current evolution model. For an arbitrary time t, the population can be decomposed into the subpopulations of the m different genotypes:

$$N(t) = \sum_{i=1}^{m} n_i(t), \tag{5.1}$$

where each subpopulation n_i grows exponentially with rate k_i determined by its fitness value. If we look at the system some amount of time τ following the most recent turbidostat dilution, assuming to have occurred at time t_0, we can write

$$n_i(t_0 + \tau) = n_i(t_0)e^{k_i \tau}, \tag{5.2}$$

provided that $N(t_0 + \tau) < N_{\max}$. If we look at the system a time T after t_0, where T is the time required for the population to grow from $N_{\min}$ to $N_{\max}$, immediately preceding the next dilution we have

$$N_{\max} = \sum_{i=1}^{m} n_i(t_0 + T)e^{k_i T}, \tag{5.3}$$

while a short amount of time δt later, during which the dilution is performed, we have

$$N_{\min} = \sum_{i=1}^{m} n_i(t_0 + T + \delta t). \tag{5.4}$$

If we assume a well-mixed population, so that all genotypes are removed proportional to their population size, we can calculate the change in population for each genotype, given by Δn_i:

$$\Delta n_i = \frac{n_i(t_0 + T)}{\sum_{j=1}^m n_j(t_0 + T)} \Delta N, \tag{5.5}$$

where $\Delta N = N_{\max} - N_{\min}$. This provides a consistency condition on the genotype subpopulations by enforcing $n_i(t_0 + T + \delta t) = n_0(t_0 + T) - \Delta n_i$.

For simplicity, we will focus on the case with two subpopulations: a wild-type population constituting the majority of the inoculated cells and growing with rate k_0 as well as a mutant population growing with rate $k_1 > k_0$. An illustration of the turbidostat dynamics are shown in Figure 5.1, generated using $n_1(t = 0) = 100$ and population thresholds $N_{\min} = 1000$ and $N_{\max} = 1500$.
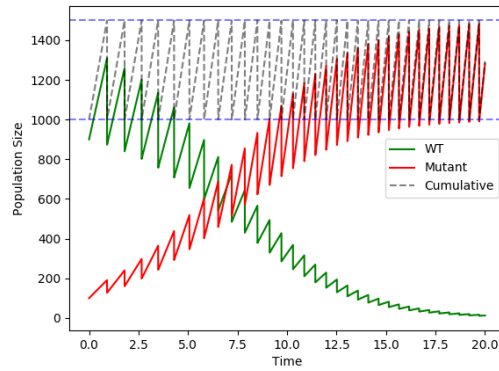
Figure 5.1: The deterministic turbidostat evolution can be easily simulated to model to proliferation of the faster-growing mutant cells (red) over the wild-type cells (green). The turbidostat dilutions are determined by the cumulative number of cells, shown in gray. Turbidostat simulated using $k_0 = 0.15$, $k_1 = 0.3$, $N_{\min} = 1000$, $N_{\max} = 1500$, and $n_1(t = 0) = 100$.

Instead of continuing to work with the population sizes, it is more natural to work with the mutant fraction

$$z(t) \equiv \frac{n_1(t)}{n_0(t) + n_1(t)}, \tag{5.6}$$

which has the major benefit of continuity across dilutions. If we use $z_0 \equiv z(t = 0)$ to denote the initial mutant inoculation fraction, we can express our subpopulation sizes immediately preceding

the first dilution event, assumed to occur at time T_0, as

$$n_0(T_0) = (1 - z_0)N_{\text{min}}e^{k_0 T_0}, \tag{5.7}$$

$$n_1(T_0) = z_0 N_{\text{min}}e^{k_1 T_0}, \tag{5.8}$$

so that the number of cells removed from each population during the next dilution will be given by

$$\Delta n_0(T_0) = \frac{(1 - z_0)e^{k_0 T_0}}{(1 - z_0)e^{k_0 T_0} + z_0 e^{k_1 T_0}}\Delta N, \tag{5.9}$$

$$\Delta n_1(T_0) = \frac{z_0 e^{k_1 T_0}}{(1 - z_0)e^{k_0 T_0} + z_0 e^{k_1 T_0}}\Delta N, \tag{5.10}$$

and the mutant fraction will be given by

$$z(T_0) = \frac{z_0 e^{k_1 T_0}}{(1 - z_0)e^{k_0 T_0} + z_0 e^{k_1 T_0}}. \tag{5.11}$$

This expression can be shown to generalize to

$$z(t) = \frac{z_0 e^{k_1 t}}{(1 - z_0)e^{k_0 t} + z_0 e^{k_1 t}}, \tag{5.12}$$

where t is simply the total runtime of the evolution experiment, with no reference made to the times at which dilutions occur. This closed-form solution is evaluated using $k_0 = 0.15$ and $k_1 = 0.3$ in Figure 5.2. Note that this derived expression suggests that the mutant fraction evolution should be independent of the optical density thresholds chosen, and this is indeed what is observed comparing the $z(t)$ calculated using the closed-form expression to what is obtained from extracting $z(t)$ from the subpopulations $n_0(t)$ and $n_1(t)$ obtained from simulating the idealized turbidostat.

5.2.2 Adding Optical Density Measurement Noise

In the very idealized model of turbidostat, the population size fluctuations affect the subpopulations $n_0(t)$ and $n_1(t)$ but have no effect on the mutant fraction $z(t)$. We now generalize the model to allow for sources of experimental noise. Since the experiment uses a macroscopic quantity of cells, we can safely neglect stochastic effects introduced from the random single-cell dynamics. And if we assume that the timescale for the initial mutant subpopulation is faster than the timescale for *de novo* mutations, these can also be neglected. Therefore, the main source of noise is introduced from experimentally measuring the optical density values. Because there is noise associated with each OD measurement, the logic used by the turbidostat to determine when a dilution is required also needs to be more complex to add robustness to outliers. To be precise, we assume that the
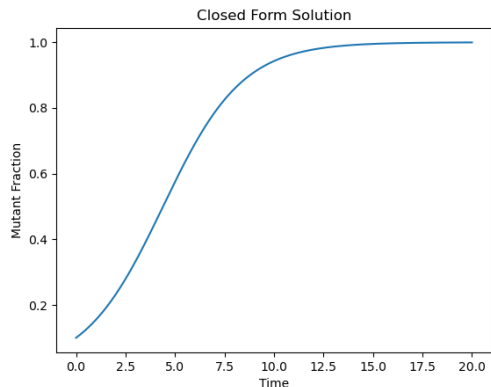
Figure 5.2: Evaluating the closed-form expression for $z(t)$ reveals a continuous evolution curve using $k_0 = 0.15$, $k_1 = 0.3$, and $z_0 = 0.1$.

optical density measurements obey

$$OD_{\text{measured}} \sim \mathcal{N}(OD_{\text{true}}, \sigma_{OD}), \tag{5.13}$$

where the true optical density values continue to evolve exponentially. For a series of optical density measurements, the mean or median optical density value calculated over some specified window is used to determine when a dilution is needed. Unless otherwise specified, we will be assuming a window length of 8 timepoints, as this was used exclusively by the turbidostat generating the recorded dataset. Using $OD_{\min} = 0.43$ and $OD_{\max} = 0.6$ with $\sigma_{OD} = 0.01$, we produce the measured optical density time series in Figure 5.3, again using $k_0 = 0.15$, $k_1 = 0.3$, and $z_0 = 0.1$. It is worth noting that the true population fractions are also affected by the optical density noise, since the dilutions can only be applied using knowledge of the measured values.

We now return to the ultimate quantity of interest, the mutant fraction $z(t)$, which can trivially be calculated given complete knowledge about the subpopulations $n_0(t)$ and $n_1(t)$. To understand the effect of noise on the mutant fraction, 10^3 complete evolution experiments were simulated using both $\Delta OD = 0.03$ and $\Delta OD = 0.17$, where both sets of thresholds were centered about $\langle OD \rangle = 0.515$. The mean $z(t)$ and the single standard deviation range are plotted using both the measured and true subpopulation sizes in Figure 5.4. Note that there is no significant shift introduced by the finite optical density bounds, and certainly no measurable shift, upon adding sensor noise to the turbidostat. This very simple model gives us confidence that careful parameter estimates are possible despite OD measurement noise.
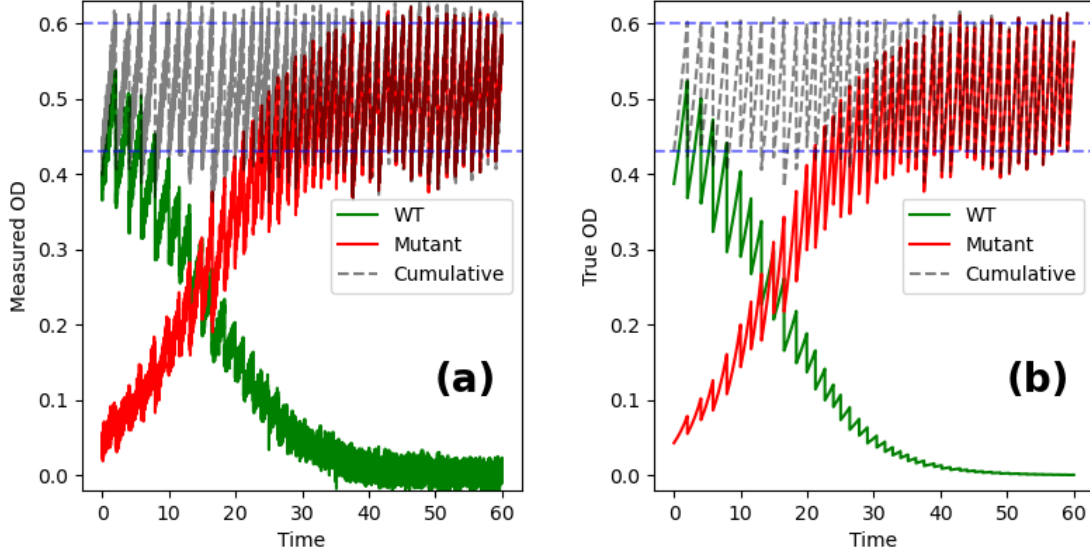
Figure 5.3: The measured and true population sizes evolve similarly upon introducing noise into the system. Parameters: $k_0 = 0.15$, $k_1 = 0.3$, and $z_0 = 0.1$.

5.3 Bayesian Parameter Estimation

To understand the experimental data generated during turbidostat evolution, we require a statistical inference method to estimate the parameters governing evolution. These estimates need to be robust to both irreducible noise as well as outlier values in the optical density measurements. These estimates also need to account for correlations that exist between parameters to accurately quantify uncertainty associated with our estimation. And finally, our statistical inference pipeline needs to be data-efficient to work with small datasets, since the evolution experiments are both time and resource intensive.

5.3.1 Turbidostat Data Sources

During the operation of a turbidostat, there are typically two different observables recorded, both of which are related to its operation. The OD values are recorded at regular intervals on the order of 1 Hz. The turbidostat uses some specified logic function, ideally robust to outliers in the measured OD values, to determine when a dilution is needed. The times at which these dilutions is performed are also recorded. We therefore have two sources of observable variables, causally and mathematically related, that inform us of the dynamics occurring on two different time scales. The OD measurements represent our fine scale knowledge of the evolution, and they completely determine the dilution measurements, the coarse-scale observable of the evolution, provided that the turbidostat logic function is known. To illustrate these two different measurements, both time

(a) Measured Mutant Fraction   (b) True Mutant Fraction
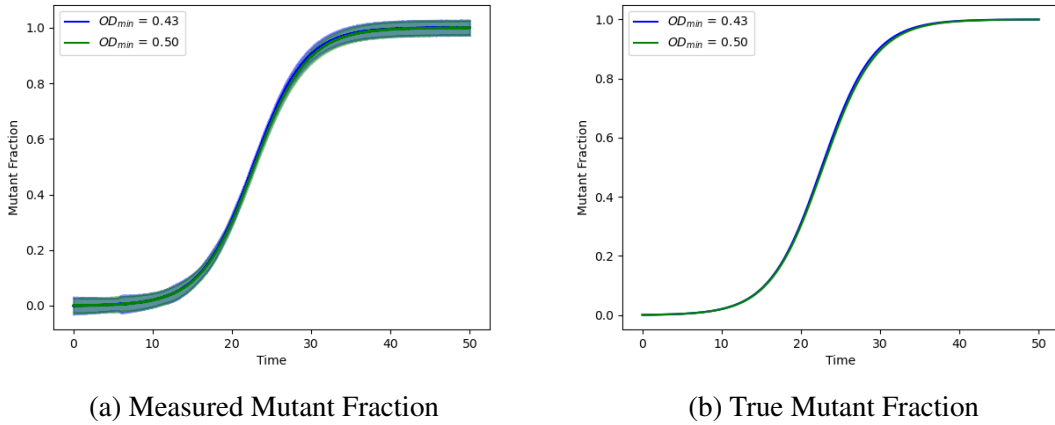
Figure 5.4: Simulating 10^3 turbidostat evolution experiments with OD sensor noise. Parameters: $k_0 = 0.15$, $k_1 = 0.3$, and $z_0 = 0.1$.

series are plotted for a portion of a sample evolution experiment in Figure 5.5. The subset of the OD time series in Figure 5.5a shows the recorded OD values, in addition to the specified OD thresholds (dotted horizontal lines) and times at which a dilution is performed by the turbidostat (solid vertical lines). The times at which the dilutions are performed as well as the size of the dilution (measured as a function of dilution pump duration) for the full experiment are shown in Figure 5.5b. Note that this data and all other experimental data shown comes from *E. faecalis* turbidostat experiments performed by Anh Huynh. I have no ownership of the data, which Anh has generously allowed me to use to illustrate turbidostat operation and test our parameter estimation methods.

The OD time series measurements are discrete, with measurements recorded every Δt. The times at which dilutions are performed can be represented as a sparse time series, with the majority of time points not requiring a dilution. The sparsity of the dilution data makes interpretation more challenging, so it is useful to apply a Gaussian kernel to the dilution data to get a continuous function representing the mean pump rate. From this smoothed dilution data, the exponential growth rates k_0 and k_1 can be extracted, and these values can be used to evaluate our expression for $z(t)$, as illustrated in Figure 5.6. Note that because we have a closed-form expression for $z(t)$, we do not need to include this as a separate parameter. Rather, k_0 is estimated from the portion of the data in which the population has no significant mutant fraction and k_1 is estimated from the portion of the data in which the population has no significant wild-type fraction. We observe remarkable agreement between the measured pump rates upon smoothing and the evaluated $z(t)$.

(a) OD Time Series

(b) Dilution Time Series
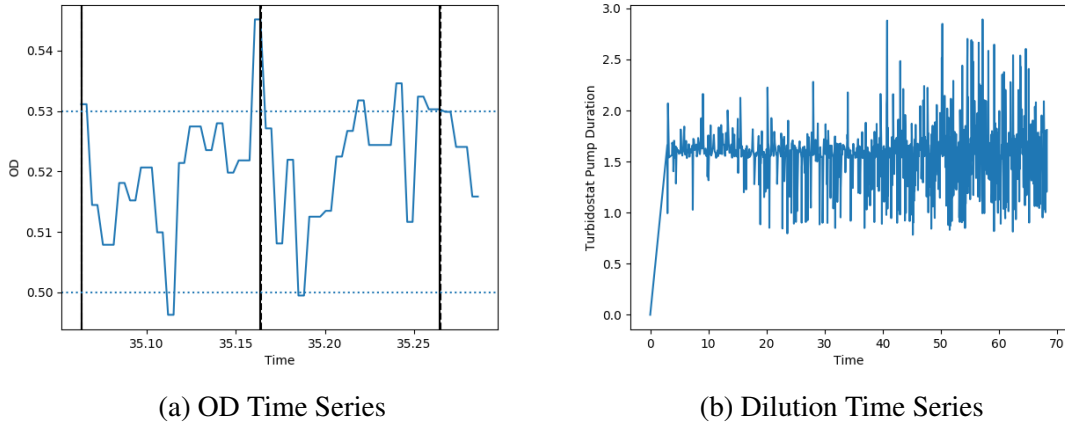
Figure 5.5: (a) Experimentally recorded time series of the OD measurements, with the dotted horizontal lines representing the specified OD thresholds and the solid vertical lines denoting times at which dilutions were performed. (b) Experimentally recorded durations of performed dilutions as a function of experiment time. Data recorded by Anh Huynh.

5.3.2 Modeling Turbidostat Growth

Although this procedure for extracting growth rates and estimating $z(t)$ is very straightforward, it is not without problems. First, the parameter estimates are affected by the Gaussian kernel applied rather than determined exclusively from measured quantities. Second, the dilution events become much less common as the optical density bounds are widened. With $\Delta OD = 0.17$, for instance, the time between dilutions is roughly one hour. Third, most of the experimental observations are discarded when the dilutions are used rather than the optical densities measurements. And lastly, there is no good way of combining estimates formed from different datasets. Due to the large amount of variation between individual experiments, turbidostat evolutions are often performed in replicates. Estimating the parameters across multiple replicates becomes more challenging. The naive approach of forming aggregate parameter estimates from the parameters extracted from the individual evolution experiments fails to account for the different quantities that do and do not vary across experiments and their respective correlations.

To remedy these issues, I developed a statistical analysis pipeline to precisely estimate the posterior distributions of the different parameters of interest. Our chief parameters of interest are the growth rates k_0 and k_1, since only these two parameters are needed to evaluate the closed form solution for $z(t)$. However, there are many nuisance parameters that require estimation from the time series data. We have already looked at one such parameter in the OD sensor noise σ_{OD}. Additionally, we observe in the experimental turbidostat data that the measured optical density values take some small amount of time following a dilution to return to $OD_{\min}$. This finite pump

67
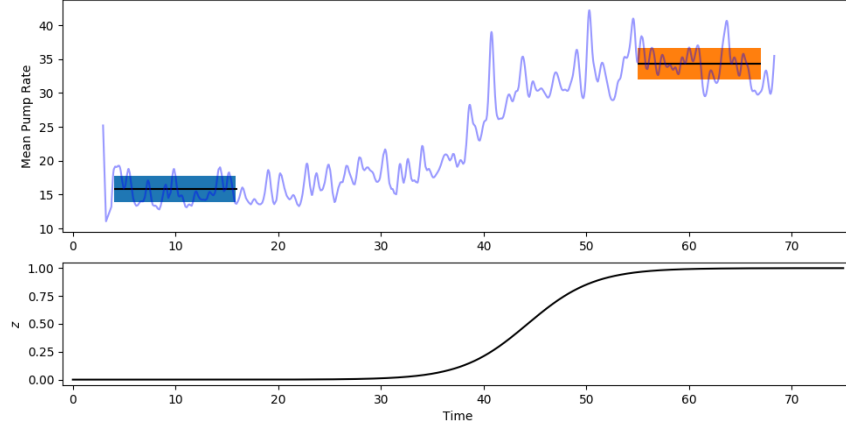
Figure 5.6: Applying a Gaussian kernel to the dilution pumps events produces a continuous curve (top panel), from which the exponential growth rates k_0 (mean and standard deviation shown in blue) and k_1 (mean and standard deviation shown in orange) can be estimated. Using these estimated growth rates, the expression for the mutant fraction $z(t)$ can be evaluated (bottom panel).

rate introduces a relaxation time $\tau \sim \mathcal{N}(\langle \tau \rangle, \sigma_\tau)$, illustrated in Figure 5.7, and adds two additional parameters. These five parameters should be the same across all turbidostat experiments, provided that the drug concentration is unchanged across experiments. We also add an estimate of $OD_{\min}$ for each vial. Although this is specified by the experimental set-up, there is no guarantee that this value is what is reached following a dilution. For instance, in some experiments where the pump needles were not sufficiently low enough in the vials, the pump cannot physically remove enough media to reduce the optical density to the desired value. By specifying this as an additional parameter for each experiment, our estimates of the parameters of interest are safeguarded against these types of experimental errors.

5.3.3 Point Parameter Estimation

For the ith individual experiment, we form a vector containing the parameter estimates given by

$$\theta^{(i)} = \begin{pmatrix} k_0^{(i)} \\ k_1^{(i)} \\ \sigma_{OD}^{(i)} \\ \langle \tau \rangle^{(i)} \\ \sigma_\tau^{(i)} \\ OD_{\min}^{(i)} \end{pmatrix}. \tag{5.14}$$

With multiple experiments performed using the same drug concentration, however, we expect all of these estimates except that of $OD_{\min}$ to be the same across all of the experiments. We form
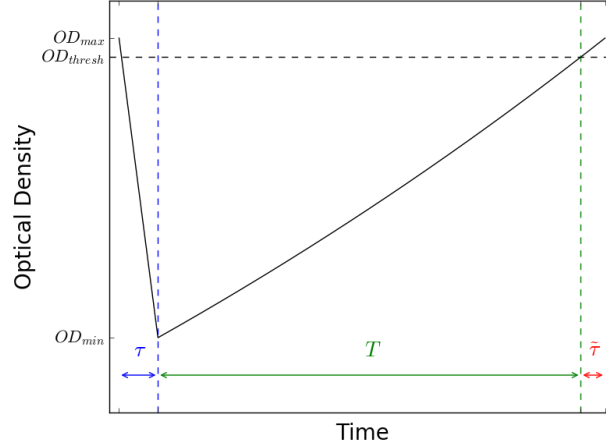
Figure 5.7: In the absence of measurement noise, the ideal period of the turbidostat time series OD measurements has a relaxation time τ, a period time T, and a lag $\tilde{\tau}$ introduced by the turbidostat logic function.

an estimated parameter vector θ^{initial} by aggregating these individual estimates according to some function f^{agg}:

$$\theta^{\text{initial}} = f^{\text{agg}}\big(\theta^{(1)}, \ldots, \theta^{(N)}\big), \tag{5.15}$$

which combines the independent estimates of the shared parameters. To evaluate the quality of a given parameter vector, I created an end-to-end simulation of the turbidostat dynamics for a single population. For a given parameter vector, the OD time series can be simulated for any desired number of periods. Using an end-to-end turbidostat simulation provides several benefit over alternative approaches. First, it allows to work with the OD measurements rather than the period measurements, which are performed much less frequently and determined entirely from the OD time series. Second, it allows us to directly use arbitrary logic functions used in the experimental apparatus, even if they do not have easily expressed derivatives. And finally, it allows us to obtain more synthetic data than we could ever hope to record experimentally. Since the evolution experiments are very slow and time-consuming, we need to be able to maximize what we can learn from each individual experiment.

Utilizing the end-to-end turbidostat simulation, the resulting distribution of OD values can then be compared to the empirical distribution of measured values, as shown in Figure 5.8, and the difference between the two distributions can be quantified using Kullback-Leibler (KL) divergence
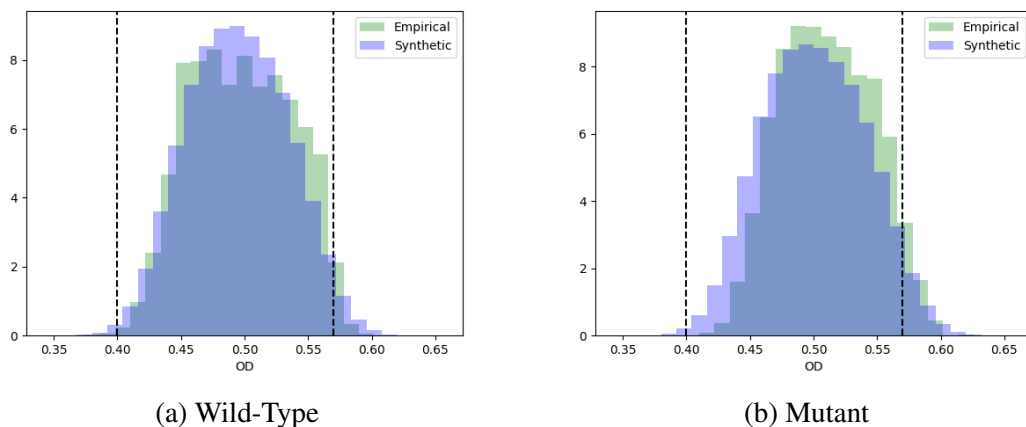
(a) Wild-Type   (b) Mutant

Figure 5.8: Forming histograms of the OD time series generated from (a) wild-type and (b) mutant cells using both empirical data and synthetically generated data.

[104], defined as

$$D_{\mathrm{KL}}(P \parallel Q) = \int_{-\infty}^{\infty} p(x) \log \left(\frac{p(x)}{q(x)} \right) dx \tag{5.16}$$

for continuous random variables P and Q with probability densities p and q, respectively. By convention, P represents the observed probability density and Q represents the model approximation of P. To be explicit, we can write Q_θ to remind ourselves that the modeled probability density is generated from a model parameterized by θ. Starting with our initial parameter vector, the parameter vector resulting in the smallest KL divergence between the empirical OD distribution is found using simulated annealing [105]. This global optimization produces θ^{SA}, the single "optimal" parameter vector found to best recreate the observed data.

5.3.4 Markov Chain Monte Carlo

Our understanding of the system parameters is not yet complete, however, since the simulated annealing approach does not readily lead to quantification of uncertainty on the parameter estimates or relative improvement for alternative parameter values. For these reasons, we employ a Markov Chain Monte Carlo (MCMC) approach to estimate the posterior distributions of each parameter [106]. With conventional MCMC using the Metropolis Hastings algorithm [107], a candidate parameter vector θ' is proposed according to some kernel $q(\theta \rightarrow \theta')$. The acceptance ratio h is

then calculated according to

$$h(\theta, \theta') = \min\left(1, \frac{P(\theta'|\mathcal{D})}{P(\theta|\mathcal{D})} \frac{q(\theta' \to \theta)}{q(\theta \to \theta')}\right), \tag{5.17}$$

where $P(\theta|\mathcal{D})$ is the posterior probability of parameters θ given the observed data $\mathcal{D}$. The proposed vector θ' is accepted with probability h and rejected with probability $1 - h$. The posterior distribution is often expressed using Bayes's rule as

$$P(\theta|\mathcal{D}) \propto P(\mathcal{D}|\theta)P(\theta), \tag{5.18}$$

where we have written the posterior distribution $P(\theta|\mathcal{D})$ as the product of the likelihood $P(\mathcal{D}|\theta)$ and prior $P(\theta)$. Using the end-to-end simulation of the turbidostat, we can readily generate synthetic data $\mathcal{D}_\theta^{\mathrm{syn}}$ for a given choice of parameter vector θ. However, we cannot map this a probability distribution. If we had an expression for the likelihood function, we could simply maximize it to calculate the maximum likelihood values (or alternatively, the maximum *a posteriori* estimates if we wanted to incorporate the prior distributions). Without a likelihood function, we cannot apply conventional MCMC to our problem. Instead, we use a variant that does not require likelihood functions to converge to the stationary distributions for the posteriors [108]. This algorithm works as follows:

1. If we start at a set of parameters θ, we propose a new sets of parameters θ' according to some kernel $q(\theta \to \theta')$.

2. We generate synthetic data $\mathcal{D}_{\theta'}^{\mathrm{syn}}$ using our model $\mathcal{M}$, which generates synthetic time series for a given choice of parameters θ'.

3. We can use some distance function ρ and some threshold ϵ such that if $\rho(\mathcal{D}_{\theta'}^{\mathrm{syn}}, \mathcal{D}) \leq \epsilon$, then we continue to the next step. Otherwise, we return to the first step to propose a new choice of parameters

4. Calculate

$$h(\theta, \theta') = \min\left(1, \frac{\pi(\theta')q(\theta' \to \theta)}{\pi(\theta)q(\theta \to \theta')}\right) \tag{5.19}$$

5. Accept θ' with probability h; otherwise stay at θ. Then repeat.

This procedure relies on using our end-to-end simulation to generate synthetic data, whose agreement with the empirical data can again be quantified using KL divergence. In our case, the threshold ϵ is determined by using the KL divergence resulting from θ^{SA}.

Note that this approach is very computationally expensive. Evaluating a single parameter vector requires simulating N_{period} periods for each of the N_{vial} independent evolution experiments. We have typically used $N_{\text{period}} = 5 \times 10^3$ to ensure that the resulting OD distribution has KL divergence values that do not differ significantly between runs due to the stochasticity present within each simulated period.

It is briefly worth explaining in more detail why we choose to work with the recorded optical density values rather than the recorded period times between dilution events. The period times are the much more natural observable quantity due to their ease of interpretation. In addition, the period times have much nicer theoretical properties than the OD time series. To understand the OD values, we need to simulate a large number of periods and compare the synthetic distribution to the empirical distribution using KL divergence. The period times, on the other hand, can be expressed as a function of the parameters of interest according to

$$T_i = \frac{1}{k_i} \log(OD_{\text{max}}/OD_{\text{min}}),\tag{5.20}$$

where T_i represents the period time of a population with genotype i growing at rate k_i. For mixed populations, one can show that the mutant fraction is given by

$$z(t) = \frac{e^{k_0 T_0} - e^{k_0 T(t)}}{e^{k_1 T(t)} - e^{k_0 T(t)}},\tag{5.21}$$

where $T(t)$ is the period time measured at time t. All of the analysis and parameter estimation pipeline can be modified to work with the period times to take advantage of these theoretical properties, but we have found that in practice the estimates are worse in quality. The sparsity of the period data is not enough to produce confident parameter estimates, despite the nice properties associated with the period data. For this reason, we have used the approach described above with the OD measurements.

5.4 Analyzing Synthetic Dataset

To illustrate the full statistical pipeline, we can start by looking at a synthetic dataset that we have generated. This allows us to understand the effect of the different components in our statistical pipeline and also evaluate the quality of the parameter estimates. Our synthetic dataset consists of the twenty hours of a simulated OD time series from ten separate vials. Each vial has a unique value for OD_{min} to be estimated, as well as the five parameters shared across all vials (k_0, k_1, σ_{OD}, $\langle \tau \rangle$, and σ_τ), where the relaxation time τ as well as the individual OD measurements are assumed to follow a Gaussian distribution. We then take this synthetic data as the empirical data whose

parameter values we wish to estimate without knowledge of the generating values.

For reference, we can compute the KL divergence between the recorded synthetic data and from simulating new data using the same generating parameters. Although these two distributions should be identical in a noiseless system, stochasticity causes the KL divergence to be a non-zero value, in our case $D_{\mathrm{KL}}(P \parallel Q_\theta^{\mathrm{generating}}) = 1.9 \times 10^{-3}$. We can then look at the different methods of estimating θ from our synthetic data and compare the resulting KL divergence values to evaluate the quality of the estimated parameters. First, we start with some rough estimate of the parameter values, obtained by randomly perturbing the generating parameter vector. This rough estimate is intended to reflect the relatively large degree of uncertainty that we would obtain by aggregating individual vial estimates. The initial parameter vector θ^{agg} produced a KL divergence roughly 26 times that of the generating parameter vector. We take θ^{agg} as the initial condition for our simulated annealing algorithm, which performs a global optimization to arrive at θ^{SA}, which has a KL divergence of 1.5 times that of the generating parameter vector. Finally, we take θ^{SA} as a promising starting point in parameter space to run our MCMC algorithm. The MCMC algorithm is initialized with a random perturbation of θ^{SA}, and new parameter vectors are accepted according to their broad prior distributions if they satisfy $\epsilon = 2D_{\mathrm{KL}}(P \parallel Q_\theta^{\mathrm{SA}})$. The mean value of the posterior distribution is used to form θ^{MCMC}, which produces a KL divergence equal to 0.98 times that of the generating parameter, suggesting that the final parameter vector explains the synthetic data approximately as well as the generating parameters, despite the relatively loose ϵ tolerance used. The results of this sequential pipeline are summarized in Table 5.1. The estimated posterior distributions formed from the MCMC algorithm are shown in Figure 5.9, which also illustrate the point estimates θ^{agg} and θ^{SA} as well as $\theta^{\mathrm{generating}}$. We can see visually that the mean of the posterior is close to the true generating value for all of the parameters, and the MCMC algorithm improves upon the simulated annealing value for almost every parameter.

θ Estimation Method	$D_{\mathrm{KL}}(P \parallel Q_\theta)/D_{\mathrm{KL}}(P \parallel Q_{\theta^{\mathrm{generating}}})$
Aggregate	26
Simulated Annealing	1.5
MCMC	0.98

Table 5.1: The different sequential components in our statistical pipeline continue to refine the mean KL divergence values, with the parameter vector estimated using MCMC producing a comparable KL divergence to the generating parameter values.

(a) σ_{OD} Posterior　　(b) k_0 Posterior　　(c) k_1 Posterior　　(d) τ Posterior

(e) σ_τ Posterior　　(f) $OD_{\min}^0$ Posterior　　(g) $OD_{\min}^1$ Posterior　　(h) $OD_{\min}^2$ Posterior

(i) $OD_{\min}^3$ Posterior　　(j) $OD_{\min}^4$ Posterior　　(k) $OD_{\min}^5$ Posterior　　(l) $OD_{\min}^6$ Posterior

(m) $OD_{\min}^7$ Posterior　　(n) $OD_{\min}^8$ Posterior　　(o) $OD_{\min}^9$ Posterior
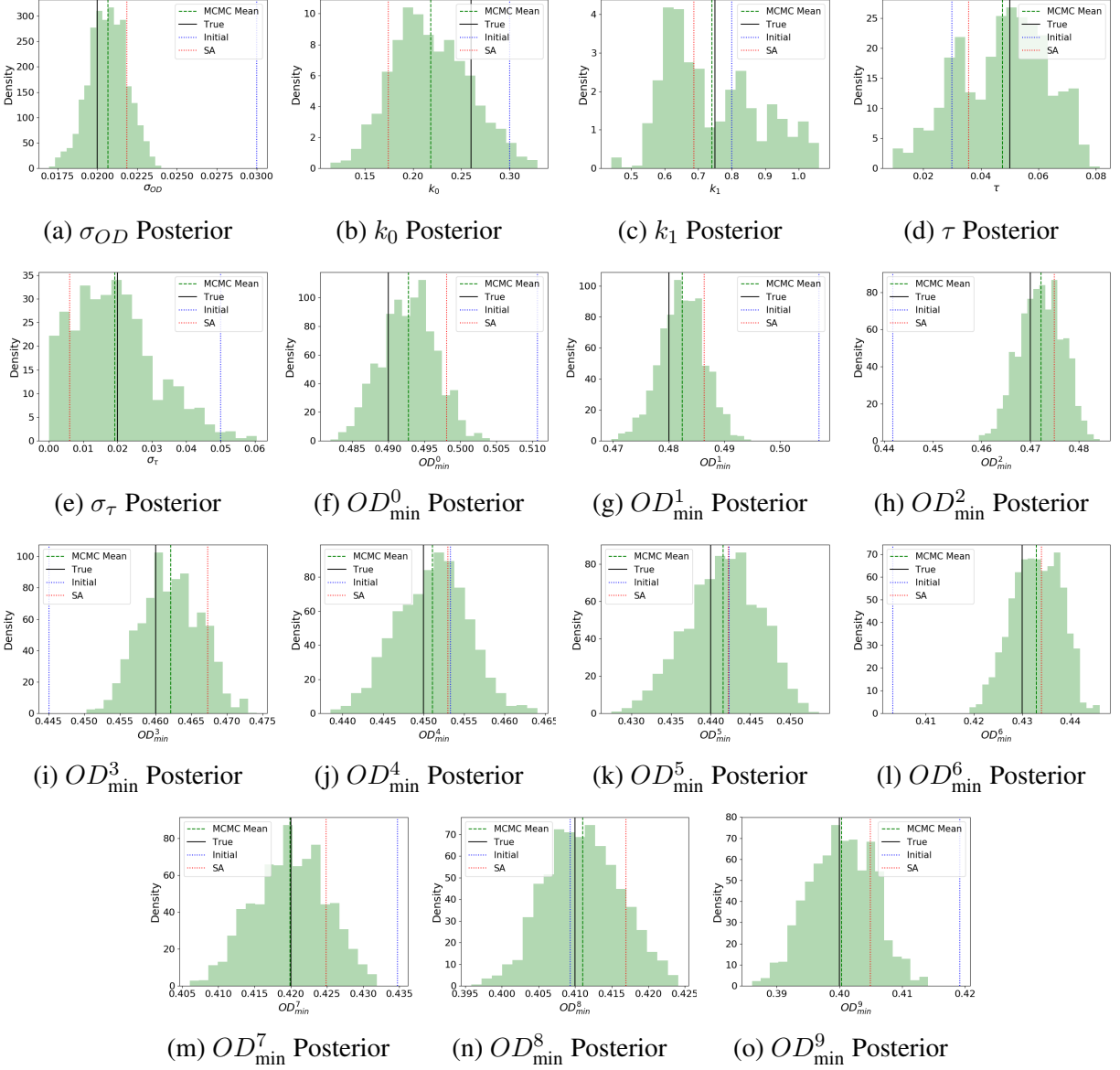
Figure 5.9: Full estimation illustration of the posterior distributions produced from the MCMC algorithm with the various point estimates and the generating parameter values for reference

5.5 Analyzing Experimental Data

With the parameter estimation pipeline validated, we can apply our algorithm to evolution data collected in turbidostat experiments. One dataset highlighted here consists of a sixteen evolution experiments performed by Anh in which the sensitive population was inoculated with a mutant fraction $z_0 = 10^{-3}$. The evolution was performed in the presence of doxycycline at a concentration of 800 ng/mL for several different values of ΔOD. Just as with the synthetic dataset, rough parameter estimates were formed from the appropriate portions of the evolution data, and these rough estimates were refined using the statistical inference pipeline. The resulting posterior distributions for the five shared parameters are shown in Figure 5.10. Despite considerable noise in the raw data, the pipeline was able to produce reasonable parameter estimates for all of the parameters, including the individual $OD_{\min}$ estimates not shown due to the large number of parameters.

Having access to these posterior distributions allows us to perform types of analysis not previously possible. As one example, we can estimate the mutant fraction $z(t)$ for any particular vial using the measured time between pump dilutions in conjunction with random samples drawn from the posterior distributions. For each recorded time between dilutions, 10^6 parameters are sampled from their respective posteriors to calculate the resulting mutant fraction associated with that particular dilution event using Equation 5.21. By leveraging our posteriors, we can better estimate $z(t)$ and also quantify our uncertainty associated with each measurement. The estimated mutant fraction $z(t)$ for a representative vial is shown in Figure 5.11, where the estimate is shown in black and the standard deviation is shaded in purple. From this profile, we can quantify metrics such as T_{50}, defined as the time required to reach $z = 0.5$, to look at different environmental factors that could introduce a shift in the observed fixation times.

As a comparison, the parameter estimation procedure can be repeated for a different set of experimental data recorded by Anh. The data consists of evolution from sixteen independent experiments, where again the populations are inoculated with $z_0 = 10^{-3}$. The only change is that these experiments were performed using 400 ng/mL of doxycycline. We would expect that this reduced concentration should lead to a larger value for k_0 and k_1, since both the wild-type and mutants should grow more quickly. However, the distributions for σ_{OD}, τ, and σ_τ would not be expected to change significantly with the drug concentration. Using the statistical pipeline, the posterior distributions were generated for each parameter, and the five shared parameter posterior distributions are plotted in Figure 5.12. Indeed, we observe a shift in the distributions for k_0 and k_1, while the other posteriors have undergone little change.
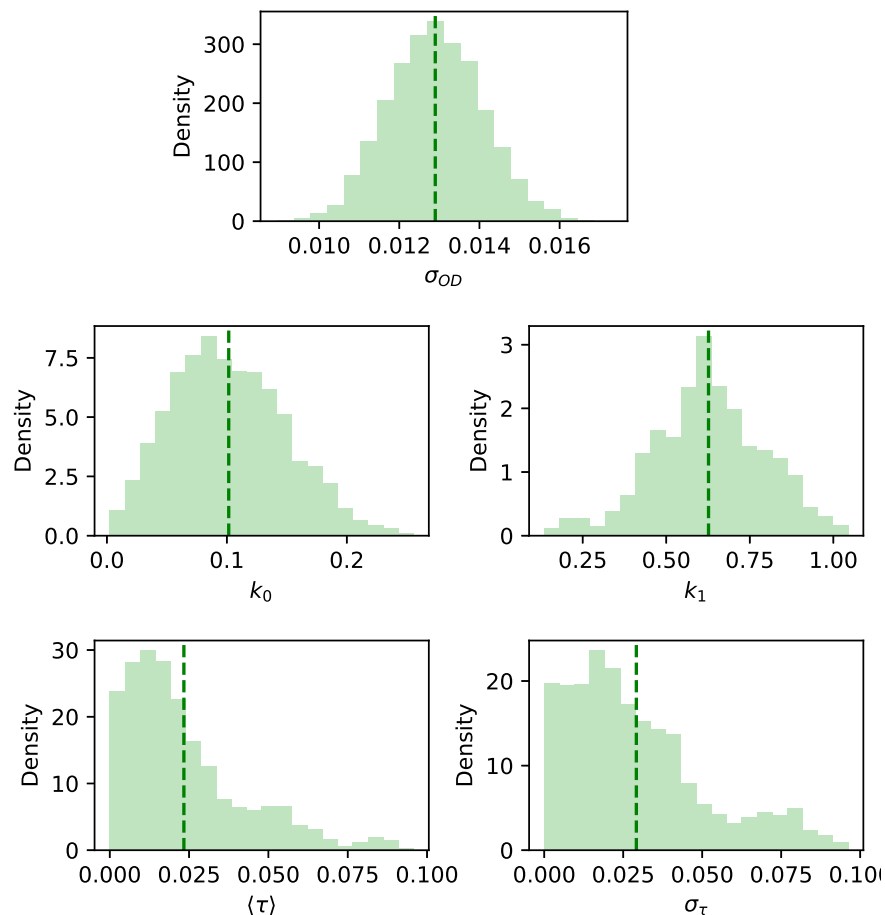
Figure 5.10: Posterior distributions of the parameters common to all 16 experiments performed in 800 ng/mL of doxycycline.

5.6   Future Work

The synthetic dataset and simple comparison of the parameter estimates produced from experimental data recorded using different drug distributions demonstrates the utility of the statistical pipeline that we have developed. Ultimately, we want to use this pipeline to carefully study more complicated evolution experiments that incorporate spatial structure and spatial heterogeneity. The end goal is to test predictions made using the toy model of evolution as well as the kinetic Monte Carlo system to observe the effects of heterogeneity in *in vitro* populations. Future work will be directed towards using and adapting the pipeline to perform parameter estimation for spatially-extended evolution experiments performed using the turbidostat.
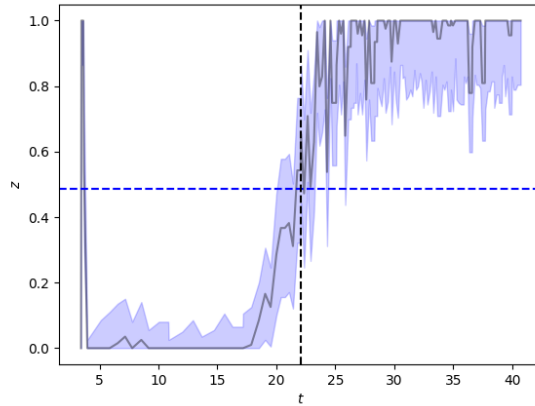
Figure 5.11: The mutant fraction can be estimated by combining the dilution events with samples drawn from the posterior distributions. This allows the quantification of metrics such as T_{50} to examine shifts in the fixation times.

5.7 Conclusion

We developed a robust statistical pipeline that was able to estimate the parameters governing the evolution of microbes in turbidostat experiments and quantify the uncertainty associated with the parameter estimates. By using an end-to-end simulation of the turbidostat dynamics, we were able to find a set of parameters that produced an OD distribution that was significantly more similar to the empirically observed distributions than competing methods for a synthetic dataset. The pipeline also produced reasonable estimates for parameters from data obtained during turbidostat evolution experiments, although more work will be required to validate the estimated values. Moving forward, we will build upon this pipeline to allow us to quantify the evolution of resistance in spatially-extended systems. Ultimately, we wish to estimate parameters of spatially-extended systems evolving under spatial heterogeneity. Through careful parameter and uncertainty quantification, we hope to confirm predictions made by the previously developed models regarding the effect of spatial heterogeneity on the emergence of resistance in *in vitro* microbial populations.
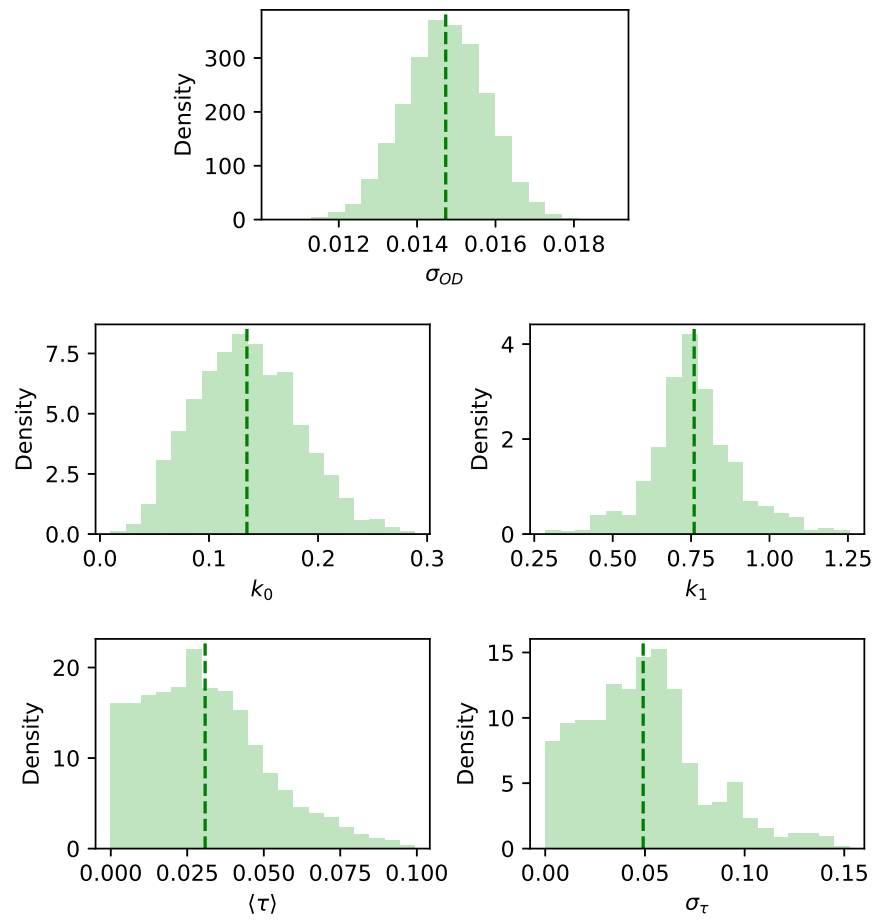
Figure 5.12: Posterior distributions of the parameters common to all 16 experiments performed in 400 ng/mL of doxycycline.

CHAPTER 6

Conclusion

In this work, we have established environmental spatial heterogeneity as a critical factor on evolutionary dynamics in microbial populations and have characterized the impact of spatial heterogeneity on the time to resistance. Spatial heterogeneity has already been identified by several experimental and theoretical studies as a factor capable of significantly altering the evolutionary dynamics leading to antibiotic resistance [31, 32, 33, 34, 35, 36, 37, 38, 29]. These studies have established that the time to reach fixation can vary, often drastically, in systems under different selection pressure landscapes induced from simple spatial drug gradients. However, the impact of spatial heterogeneity on fixation has not been completely understood from this early work, motivating a more careful study of spatial heterogeneity.

To study the effects of spatial heterogeneity on the impact of antibiotic resistance in microbial populations, we have developed simplified models of evolution in spatially-extended systems. We first created a toy model assuming Moran-like evolutionary dynamics sufficiently simple to allow for semi-analytic calculations for the mean fixation time. We found that different distributions of selection pressure in small systems can still significantly impact the time to fixation, and that the effect of spatial heterogeneity depends critically on the other parameters of the system. Upon generalizing the initial evolution model to account for imperfect mutants and larger systems, we observed that not only does spatial heterogeneity continue to affect the fixation times but that this effect is often amplified in these more general systems. We then developed a more complex evolution model supporting much larger systems that was simulated using a kinetic Monte Carlo algorithm and found that spatial heterogeneity again plays an important role in determining the fixation time, despite conserving the spatially-averaged growth across different landscapes. Finally, we developed a statistical pipeline for robust parameter estimation from evolution data recorded from populations growing in turbidostats. This statistical pipeline will allow testing the various hypotheses posited by the developed models.

Using these models that we have created, we have focused on better understanding the role spatial heterogeneity plays in the evolutionary dynamics of microbial populations in three key areas.

First, we wanted to understand the effect of spatial heterogeneity without additional confounding properties that vary across different spatial profiles. In studying spatial selection profiles, for instance, this would constrain all landscapes to conserve spatially-averaged selection pressure to isolate the effect of spatial distribution of selection pressure. Physically, the selection pressure landscape can be imagined as the result of some drug present. A natural question to ask is whether different spatial drug dosing schemes still affect the time to resistance when a fixed amount of drug is distributed throughout the system. We have found that the spatial distribution of selection pressure critically impacts the evolutionary dynamics and time to resistance in all of the models that we have developed. We have seen how the specific effect of heterogeneity can depend on both the parameters of the spatially-extended system (namely, the mutation and migration rates) as well as the specific parameterization of the selection pressure landscapes. We have characterized these relationships and developed simple approximations allowing us to intuitively understand the impacts of heterogeneity in certain limiting cases.

Our second goal was to explore spatial selection pressure landscapes more complex than simple linear drug gradients. With simple linear drug gradients, there exists a single direction of directed selection pressure—as the drug concentration increases, the mutants become increasingly advantageous. And importantly, additional mutations conferring resistance would also benefit from arising in the same natural direction created by the drug gradient. However, drug concentration is just one environmental factor affecting the selection pressure landscape, which is also impacted by temperature, pH, nutrient distribution, carrying capacity, and other factors. Collectively, simple spatial distributions of these factors naturally lead to non-monotonic profiles of spatial selection pressure that have no single directed selection pressure direction. Our results have focused on studying evolution under these non-monotonic selection profiles. We have seen that while the effect of spatial heterogeneity is important for affecting the emergence of resistance under non-monotonic selection profiles, the non-monotonicity is less important for systems with few microhabitats. With larger systems approximating a continuous spatially-extended system, however, the monotonicity of the selection profile becomes very important upon considering multiple beneficial mutations conferring resistance.

Finally, our third goal was to leverage our knowledge of evolutionary dynamics to use spatial heterogeneity to slow down the emergence of resistance. Up until this point, much of the existing experimental and theoretical work with simple drug gradients has illustrated the ability of heterogeneity to lead to significantly faster resistance. Spatial selection pressure landscapes represent an exciting method of slowing resistance because they are an external, tunable parameter of the system unlike the genotype landscape of the cell, which is inherent to the cell and difficult to modulate. While we have observed that heterogeneity can lead to faster resistance, as observed in the existing literature, we have also uncovered many scenarios in which certain selection pressure landscapes

lead to slower resistance than the corresponding spatially homogeneous landscape conserving the same mean selection pressure. Such scenarios were observed in the models of evolution in both the small and large systems. We also saw how initial mutant subpopulations can still benefit from spatially-optimized selection profiles. The existence of spatial drug dosing strategies capable of slowing the resistance independent of specific drug mechanisms or genotypic landscapes is an exciting prospect for future strategies to slow the emergence of antibiotic resistance.

We crafted our models with the goal of capturing fundamental relationships between evolutionary dynamics and resistance emergence in a manner independent of specific considerations such as drug mechanisms, gene regulatory networks, and molecular biology. The independence of these important considerations allows for the possibility of models capturing more general relationships applying to evolution in a wide range of spatially-structured populations. However, a corrollary of these assumptions is that additional work is required to understand our findings across a range of different evolution models and ultimately in *in vitro* and even *in vivo* systems. To maximize the generality of our findings, we have focused on gaining an intuition behind the impact of heterogeneity on the evolutionary dynamics to form a more general set of principles that may apply to more general systems. By concentrating on these general intuitive principles and associated testable hypotheses, we can hope to refine how we think about evolution in more complex systems. The statistical pipeline we developed for robust parameter estimation is one preliminary step towards testing the model predictions in systems that do not strictly obey all of the assumptions in the simplified models. Additionally, much work remains to fully understand the emergence of antibiotic resistance, even through the specific lens of spatial heterogeneity, because it is the culmination of a fundamentally stochastic process governed by a complex interplay of many genomic, mechanistic, and ecological considerations. Each of these facets is at least somewhat understood in isolation, but future work is required to understand resistance emergence when all of these considerations shape the evolutionary dynamics.

Although future work is needed to more holistically understand spatial heterogeneity, in particular under different evolution paradigms, and to connect this work to existing studies focusing more on genomic and mechanistic considerations, we believe that this work helps further establish spatial heterogeneity as an important facet for a mature understanding of microbial evolution. We believe that utilizing spatial heterogeneity to create spatial selection landscapes can also be a powerful tool in the fight to slow the emergence of antibiotic resistance. Because spatial selection profiles can be modulate externally through a number of factors, spatially optimized drug dosing strategies could be implemented that guarantee slower resistance in expectation. Such spatial drug dosing strategies could be optimized for specific infection subpopulations, bacteria motility, and mutation rates between various genotypes to create prescriptive approaches tailored to specific infections and systems.

APPENDIX A

Toy Model of Evolution

A.1   Backward Master Equation Operator

The master equation operator Ω is populated from Markovian transition rates that come from the dynamics allowed in our system within the Moran framework. We denote $T^+[x_i]$ as a transition that increases the number of mutants in microhabitat x_i. Explicitly, our three transition rates increasing the number of mutants from an initial state (j, k, ℓ) are

$$
T^+[x_0](j, k, \ell) = \underbrace{(N - j)}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_0} \underbrace{\frac{(N - j)r_0(x_0)}{(N - j)r_0(x_0) + jr^*(x_0)}}_{\text{Pick WT}} \underbrace{\mu}_{\text{WT mutates}}
$$

$$
+ \underbrace{(N - j)}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_0} \underbrace{\frac{jr^*(x_0)}{(N - j)r_0(x_0) + jr^*(x_0)}}_{\text{Pick mutant}} \underbrace{(1 - \beta)}_{\text{No migration}}
$$

$$
+ \underbrace{(N - j)}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_1} \underbrace{\frac{kr^*(x_1)}{(N - k)r_0(x_1) + kr^*(x_1)}}_{\text{Pick right mutant}} \underbrace{\beta}_{\text{Migration}} \quad \text{(A.1)}
$$

$$T^+[x_1](j,k,\ell) = \underbrace{(N-k)}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_1} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1) + kr^*(x_1)}}_{\text{Pick WT}} \underbrace{\mu}_{\text{WT mutates}}$$

$$+ \underbrace{(N-k)}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_1} \underbrace{\frac{kr^*(x_1)}{(N-k)r_0(x_1) + kr^*(x_1)}}_{\text{Pick mutant}} \underbrace{(1-2\beta)}_{\text{No migration}}$$

$$+ \underbrace{(N-k)}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_1} \underbrace{\frac{jr^*(x_0)}{(N-j)r_0(x_0) + jr^*(x_0)}}_{\text{Pick left mutant}} \underbrace{\beta}_{\text{Migration}}$$

$$+ \underbrace{(N-k}_{\text{WT dies}} \underbrace{w(x_0)}_{\text{Pick } x_1} \underbrace{\frac{\ell r^*(x_2)}{(N-\ell)r_0(x_2) + \ell r^*(x_2)}}_{\text{Pick right mutant}} \underbrace{\beta}_{\text{Migration}} \quad \text{(A.2)}$$

and

$$T^+[x_2](j,k,\ell) = \underbrace{(N-\ell)}_{\text{WT dies}} \underbrace{w(x_2)}_{\text{Pick } x_2} \underbrace{\frac{(N-\ell)r_0(x_2)}{(N-\ell)r_0(x_2) + \ell}}_{\text{Pick WT}} \underbrace{\mu}_{\text{WT mutates}}$$

$$+ \underbrace{(N-\ell)}_{\text{WT dies}} \underbrace{w(x_2)}_{\text{Pick } x_2} \underbrace{\frac{\ell}{(N-\ell)r_0(x_2) + \ell}}_{\text{Pick mutant}} \underbrace{(1-\beta)}_{\text{No migration}}$$

$$+ \underbrace{(N-\ell)}_{\text{WT dies}} \underbrace{w(x_1)}_{\text{Pick } x_1} \underbrace{\frac{k}{(N-k)r_0(x_1) + k}}_{\text{Pick left mutant}} \underbrace{\beta}_{\text{Migration}} \quad . \quad \text{(A.3)}$$

Within the Moran process, we also have transitions that decrease the number of mutants in a given microhabitat. We denote $T^-[x_i]$ as the transition decreasing the number of mutants in microhabitat x_i. The explicit transition rates are given by

$$T^-[x_0](j,k,\ell) = \underbrace{j}_{\text{Kill mutant}} \underbrace{w(x_0)}_{\text{Pick } x_0} \underbrace{\frac{(N-j)r_0(x_0)}{(N-j)r_0(x_0) + j}}_{\text{Choose WT}} \underbrace{(1-\beta-\mu)}_{\text{WT grows without mutation or migration}}$$

$$+ \underbrace{j}_{\text{Kill mutant}} \underbrace{w(x_1)}_{\text{Pick } x_1} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1) + k}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Right migration}} \quad , \quad \text{(A.4)}$$

$$T^-[x_1](j,k,\ell) = \underbrace{k}_{\text{Kill mutant}} \underbrace{w(x_1)}_{\text{Pick } x_1} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1)+k}}_{\text{Choose WT}} \underbrace{(1-2\beta-\mu)}_{\text{WT grows without mutation or migration}}$$

$$+ \underbrace{k}_{\text{Kill mutant}} \underbrace{w(x_0)}_{\text{Pick } x_0} \underbrace{\frac{(N-j)r_0(x_0)}{(N-j)r_0(x_0)+j}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Left migration}}$$

$$+ \underbrace{k}_{\text{Kill mutant}} \underbrace{w(x_2)}_{\text{Pick } x_2} \underbrace{\frac{(N-\ell)r_0(x_2)}{(N-\ell)r_0(x_2)+\ell}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Right migration}} \;, \quad \text{(A.5)}$$

and

$$T^-[x_2](j,k,\ell) = \underbrace{\ell}_{\text{Kill mutant}} \underbrace{w(x_2)}_{\text{Pick } x_2} \underbrace{\frac{(N-\ell)r_0(x_2)}{(N-\ell)r_0(x_2)+\ell}}_{\text{Choose WT}} \underbrace{(1-\beta-\mu)}_{\text{WT grows without mutation or migration}}$$

$$+ \underbrace{\ell}_{\text{Kill mutant}} \underbrace{w(x_1)}_{\text{Pick } x_1} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1)+k}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Left migration}} \;. \quad \text{(A.6)}$$

With these transition rates, Ω is populated according to

$$\Omega_{m,n} = \begin{cases} W^{n \to m} & n \neq m \\ -\sum_{m' \neq m} W^{n \to m'} & n = m \end{cases} \quad \text{(A.7)}$$

where $W^{n \to m}$ is the transition rate from initial state n to final state m. Note that the number of mutants can change by one at most, so the backwards master equation operator is sparse.

We can then solve the mean first passage time equation

$$-1 = \sum_{m' \neq m_f} T(m_f|m')\Omega_{m',m_i} \quad \text{(A.8)}$$

where $T(m_f|m_i)$ is the mean time required for a system initially in state m_i to first reach state m_f. All of the natural boundary conditions are already imposed by Ω, but we do need to specify that

$$T((n^*(x_0), n^*(x_1), n^*(x_2))|(n^*(x_0), n^*(x_1), n^*(x_2))) = 0 \quad \text{(A.9)}$$

for any initial state $(n^*(x_0), n^*(x_1), n^*(x_2))$, which is intuitively obvious.

A.1.1 Choice of Model Weights

The microhabitat-specific weights used in the above transition rates allow us to explore different evolution models within this general framework. We investigate two different formulations for the model weights. The two formulations we consider are:

1. Weigh each microhabitat by its mean fitness normalized by the mean fitness of the entire system. This has the nice corollary that microhabitats in which cells replicate more frequently have more replication than other microhabitats.

2. Weigh all microhabitats equally. This leads to more simple model of evolution.

At this point, we would be remiss to not admit that neither choice of weights leads to a particularly realistic model of microbial evolution. Our goal is not to develop a high-fidelity model of evolution requiring knowledge of cell mechanisms, drug mechanisms, and chemical kinetic coefficients. Rather, we intend to develop a toy model of evolution that can exhibit interesting behavior despite its simplicity. Without relying on specific knowledge of mechanisms, we can hope that these results apply generally to microbial populations evolving in systems with spatial structure. With this motivation in mind, one can make a convincing argument for either choice of microhabitat weights—each leads to a relatively simple model of evolution motivated by observations of evolving microbial systems.

Let's explicitly calculate the microhabitat weights under the first formulation, in which mean fitness affects the relative replication rate. In this case, we write

$$w(x_i) = \frac{(N - n^*(x_i))r_0(x_i) + n^*(x_i)(1 - \epsilon)}{\sum_m \left[(N - n^*(x_m))r_0(x_m) + n^*(x_m) \right]}, \tag{A.10}$$

where $n^*(x_i)$ denotes the number of mutants in microhabitat x_i. If we define the mean fitness within a microhabitat as

$$\bar{r}(x_i) = \frac{(N - n^*(x_i))r_0(x_i) + n^*(x_i)}{N}, \tag{A.11}$$

we can equivalently write

$$w(x_i) = \frac{\bar{r}(x_i)}{\bar{r}(x_0) + \bar{r}(x_1) + \bar{r}(x_2)}. \tag{A.12}$$

Note that if all microhabitats have the same mean fitness at a particular state of the system, all microhabitats are equally likely to be chosen for the next replication event, and the transition rates match those we would obtain under the equally-weighted assumption, up to a scaling factor.
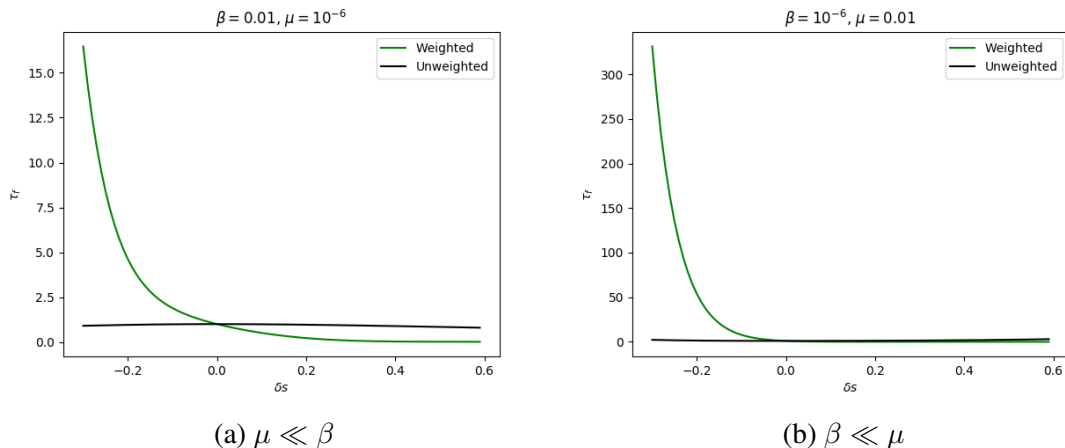
Figure A.1: The fixation times are calculated across different landscapes conserving $\langle s \rangle = 0.2$ using both the fitness-weighted and unity-weighted ("unweighted") transition rates. Although both models exhibit spatial heterogeneity in their respective fixation times, the fitness-weighted model shows an amplified effect of heterogeneity on the resulting fixation times.

With our weights defined, we can look at the results obtained with either model of evolution. The body of this dissertation extensively looks at the results obtained from assuming that all weights are unity. Recall that the principal observation of interest under this assumption is that the fixation time under different selection pressure landscapes depends critically on the distribution of selection pressure (δs), despite conserving the average selection pressure ($\langle s \rangle$). We observe that the distribution of selection pressure also affects the fixation times in the alternative model, in which each microhabitat is weighted according to its fitness. In Figure A.1, we plot the fixation times as a function of δs, assuming $\langle s \rangle = 1/6$ and $N = 25$, for both $\beta \ll \mu$ and $\mu \ll \beta$. The fixation curves for each model are normalized by their respective homogeneous landscape fixation time for ease of comparison. We again observe that fixation times can vary considerably with δs, significantly more so than the fixation times calculated using the unweighted model.

To explain the effect of spatial heterogeneity in this weighted model, we turn to more simple class of selection pressure landscapes that do not conserve $\langle s \rangle$. To begin we look at the fixation times calculated assuming a homogeneous fitness landscape, shown in Figure A.2. We see that in the limit $\mu \ll \beta$, the fixation times increase monotonically with the fitness value in the microhabitats. This is due to the genetic drift favoring the mutants becoming less strong as the wild-type fitness is increased. Because *de novo* mutations are rare in this limit, the fixation times grow rapidly as the mutants become more likely to go extinct due to genetic drift. The results with a homogeneous fitness landscape in the limit $\beta \ll \mu$, in which the microhabitats are treated as essentially independent with respect to migration, tell a very different story. Looking at the
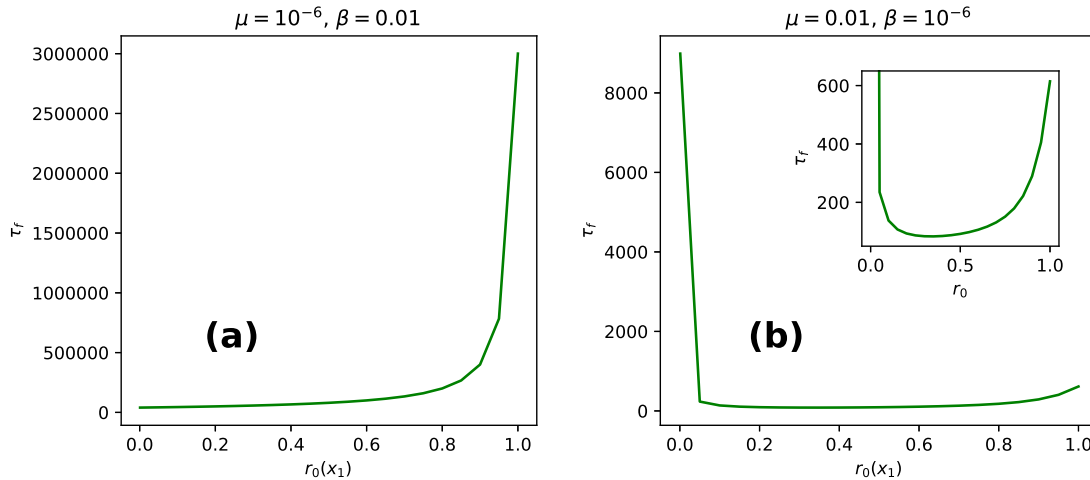
Figure A.2: (a) The fixation times for a homogeneous landscape of fitness r_0 using the weighted transition rates using $\mu = 10^{-6}$ and $\beta = 0.01$. (b) Fixation times using $\mu = 0.01$ and $\beta = 10^{-6}$, with inset showing zooming in the fixation times excluding $r_0 \ll 1$. All plots assume $N = 25$.

fixation times in Figure A.2(b), the most obvious feature is an artifact introduced by the weighted transition rate formulation when the wild-type fitness is near zero. When the wild-type fitness is very small, the mean fitness of the microhabitat grows considerably upon the addition of a single mutant. This causes the microhabitat with the mutant to be much more likely to be chosen for the next replication event, introducing a positive feedback into the system that can slow fixation across all of the microhabitats. For fitness values large enough to not be affected by this artifact, we observe a similar trend in fixation times to what we observed in the homogeneous landscape in the limit $\mu \ll \beta$.

We can use these results to understand landscapes with spatial heterogeneity. For a simple choice of landscapes, we look at those where the extrema fitness $r_0(x_0)$ is fixed to some specified value and the center fitness $r_0(x_1)$ is varied. In general, the observed behavior is a combination of the two counteracting forces that we observed in the homogeneous landscapes—as the wild-type fitness increases, the microhabitat is less likely to get "ignored" due to its small mean fitness without a mutant but also more likely to have a small mutant subpopulation go extinct due to genetic drift becoming less favorable to the mutants. This is illustrated in Figure A.3, which plots several representative fixation curves for limiting cases of the parameter values.

With results from both evolution models, we can revisit the relative merit of each model. The fitness-weighted transition rates appear to add an additional level of realism to the system. However, even this is not necessarily true in all cases. A potential issue is that both mutation and migration in this framework can only occur following replication. By making some microhabitats have more replication events than others, the effective mutation rates and migration rates also gain
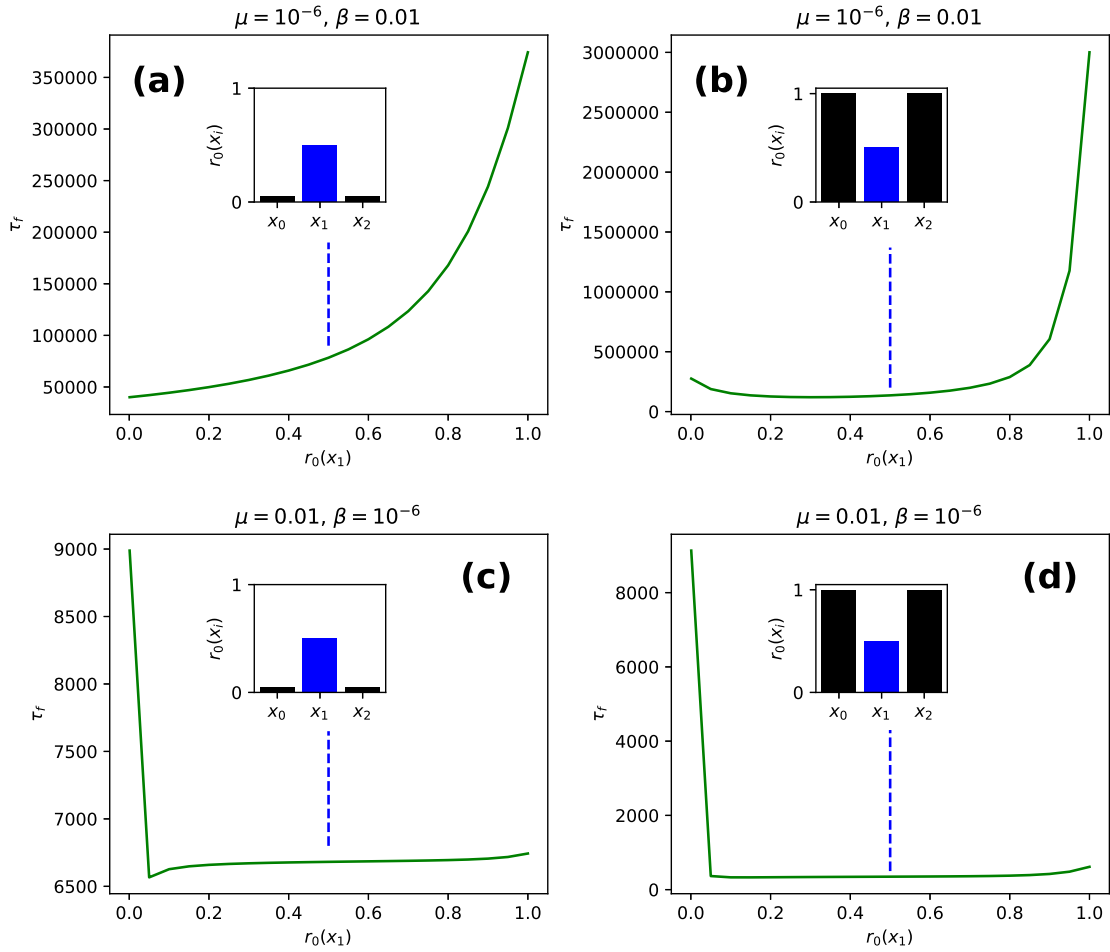
Figure A.3: (a) The fixation times for $\mu = 10^{-6}$, $\beta = 0.01$ fixing $r_0(x_0) = 10^{-3}$ using the weighted transition rates. (b) Fixation times using $\mu = 10^{-6}$, $\beta = 0.01$ fixing $r_0(x_0) = 1$. (c) Fixation times using $\mu = 0.01$, $\beta = 10^{-6}$ fixing $r_0(x_0) = 10^{-3}$. (d) Fixation times using $\mu = 0.01$, $\beta = 10^{-6}$ fixing $r_0(x_0) = 1$. All plots use $N = 25$.
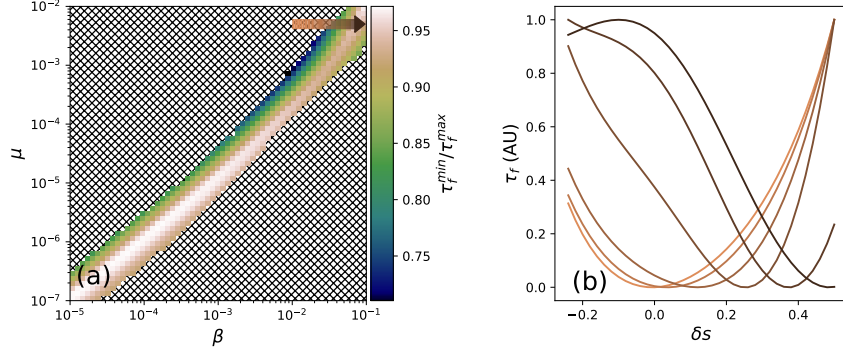
Figure A.4: (a) Minimum fixation times (τ_f^{min}) over different selection pressure distributions (relative to the maximum fixation times, τ_f^{max}) in the intermediate parameter region where fixation can be both accelerated and decelerated. $N = 25$ and $\langle s \rangle = 0.167$. (b) Across a specific trajectory in the intermediate region (arrow in panel (a)), the dependence of τ_f on heterogeneity (δs) transitions smoothly from a state with a minimum at $\delta s = 0$ (light test curve) to one with a maximum at $\delta s = 0$ (darkest curve). For ease of comparison, fixation times are scaled to arbitrary units between 0 and 1.

a spatial dependence. The observation that spatial heterogeneity can impact fixation times is also less surprising with this model. When the spatial distribution of fitness also controls the replication timescale within each microhabitat, the system lends itself to a trivial (although incomplete) explanation of the effect of spatial heterogeneity. With the unweighted transition rates, any deviation in fixation times caused by fitness distribution requires a more thoughtful explanation. For these reasons, we focus on the more simple model where all microhabitats are weighted equally. However, the fact that spatial heterogeneity impacts the fixation times calculated with both models only strengthens the applicability of the results.

A.2 Characterization of Intermediate Regime

To further investigate evolution in the intermediate regime, where heterogeneity can either speed or slow fixation, depending on the specific profile, we calculated the minimum and maximum fixation times ($\tau_f^{\min}$ and $\tau_f^{\max}$, respectively) as δs is modulated to create different selection pressure distributions (Fig. A.4a). While in many cases the fixation time is decreased by only a few percent, we do find larger effects in the high and low migration limits (i.e., on the edges) of the intermediate regime. To understand how the spatial τ_f profiles change over this intermediate region, we fix μ and traverse across a trajectory in β as shown by the arrow in Fig. A.4a. As we increase β, τ_f smoothly transitions from being minimized at $\delta s = 0$ to being maximized near $\delta s = 0$ (Fig. A.4b).

A.3 Alternative Landscape Parameters

The results in the main text were all generated with $\langle s \rangle = 0.167$. However, the existence of three regions of parameter space is not specific to this particular value of $\langle s \rangle$. In Fig. A.5, we see that we observe three regions of parameter space with a variety of values for $\langle s \rangle$. It is worth noting that the relative size and location of the intermediate region change with $\langle s \rangle$, and the magnitude of effects also changes with $\langle s \rangle$. However, the existence of these regions is constant throughout these different classes of landscapes.

Note that it a second intermediate region appears with larger values of $\langle s \rangle$. The existence of this additional intermediate region does not follow the same intuitive understanding and is more subtle. This second intermediate region arises due to numerical reasons. Upon discretizing our selection landscape in steps of $\Delta \delta s = 0.1$, it is possible for the "true" minimum to occur near $\delta s = 0$ but not exactly at this value. With these relatively large step sizes, it allows the "true" minimum to be mapped to $\delta s = 0$ even though the true minimum occurs at some small non-zero value. So the apparent area between the two intermediate regions is actually an artifact of the numerical discretization of our landscape. Further evidence comes from looking at the fixation times in this second intermediate region, which differ by less 1% from the spatially homogeneous landscape, so it is likely of little significance.

We can also vary N to confirm that the existence of the three regions of parameter space is resilient to changes in N. We investigated a few different values of N in Fig. A.6, and we again see the same three regions of parameter space.

A.4 Alternative Topologies

We can also confirm that the existence of three regions of parameter space is not predicated upon the specific topology chosen. The results of the main text assumed a nearest-neighbor connection without migration between the boundary microhabitats. We can easily imagine additional topologies for a specified selection landscape. We can remove one of the existing connections and allow migration between the boundary microhabitats (which is equivalent to cyclically permuting the selection landscape with the same connections). The resulting phase plot is shown in Fig. A.7a. We can also imagine a global topology in which all microhabitats are connected to all others. The phase plot from this global topology is shown in Fig. A.7b. Note that the phase plots differ quantitatively when the topology is changed, but all three topologies support three different regions of parameter space.

(a) $\langle s \rangle = 1/3$

(b) $\langle s \rangle = 1/2$

(c) $\langle s \rangle = 2/3$

(d) $\langle s \rangle = 5/6$
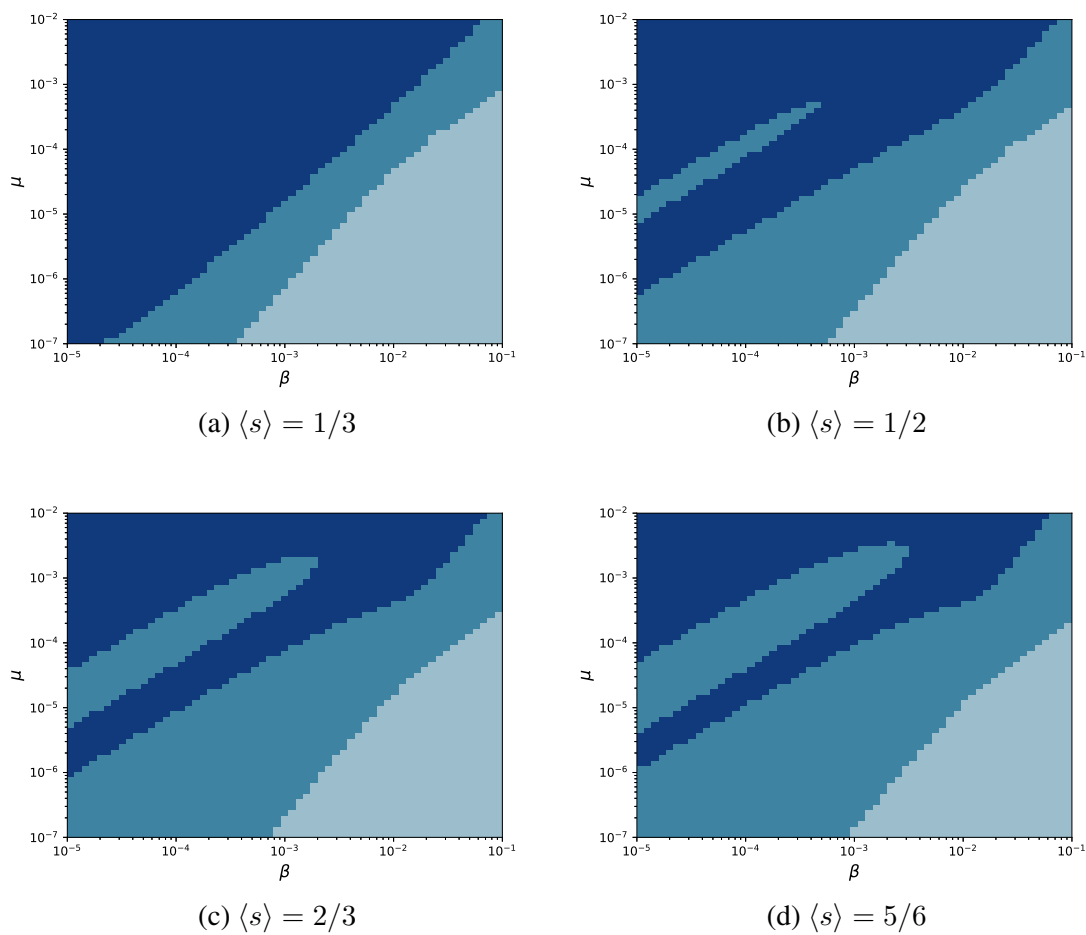
Figure A.5: Changing $\langle s \rangle$ does not change the existence of three regions of our parameter space. The appearance of a second intermediate region is artificial and due to numerical discretization.

(a) $N = 15$

(b) $N = 35$
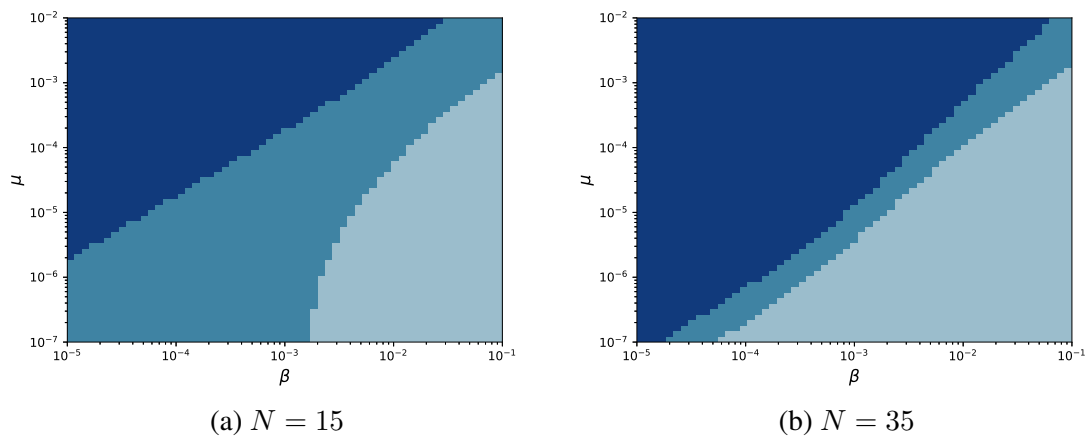
Figure A.6: Different population sizes N result in the same qualitative behavior.

(a) Permuted Topology   (b) Global Topology
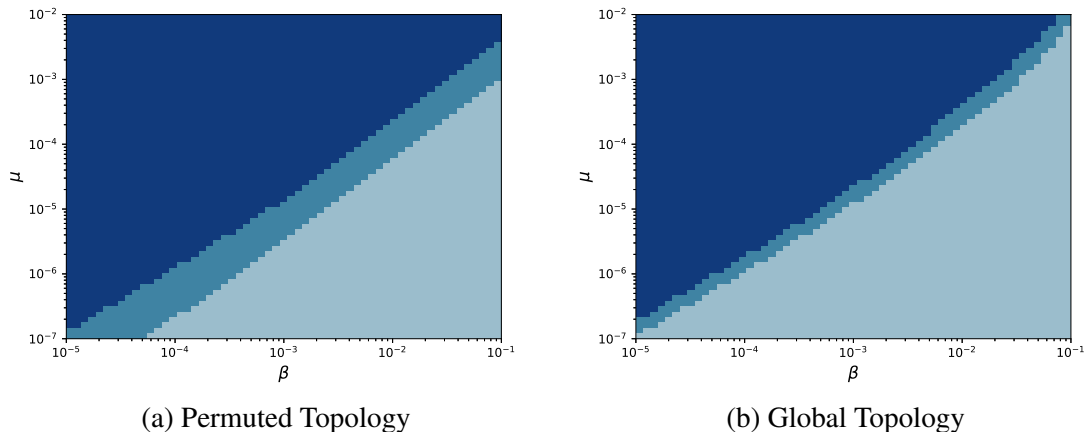
Figure A.7: Different topologies also produce three unique regions of parameter space.

A.5 Monotonic Landscape

We can also imagine using a monotonic selection landscape. A non-monotonic landscape potentially allows for more complicated dynamics due to the lack of a single direction of selection pressure gradient. In Fig. A.8, we see that a monotonic landscape also results in three regions of parameter space.

A.6 Analytical Approximation for Single Habitat

In a single microhabitat with selection pressure s, we can approximate the fixation time $\tau_f(s)$ in the limit $\mu \ll 1$ as

$$\tau_f(s) \approx \frac{1}{\lambda_{eff}} = \frac{1}{\mu N P_{fix}(s)}. \tag{A.13}$$

Equation A.13 provides a good approximation to $\tau_f(s)$ for a range of $\mu \ll 1$ (Figure A.9). This approximation assumes that mutations are sufficiently rare that fixation times are dominated by arrival and fixation of a single mutant (i.e., the timescale of fixation of a single mutant is fast compared to the expected arrival time of the next mutant via mutation). Therefore, the arrival time distribution is exponential with a rate parameter λ_{eff}. The factor μN is the transition rate $T^+(0)$ for increasing the number of mutants by one starting from zero mutants (see, for example, Equation A.1 with $j = 0$ and $\beta = 0$). P_{fix} is the probability of fixation of a single mutant in a habitat with selection pressure s and no mutation. The expression for $P_{fix}(s)$ is well-known, but for completeness, we briefly repeat the derivation here. To do so, we first write an iterative
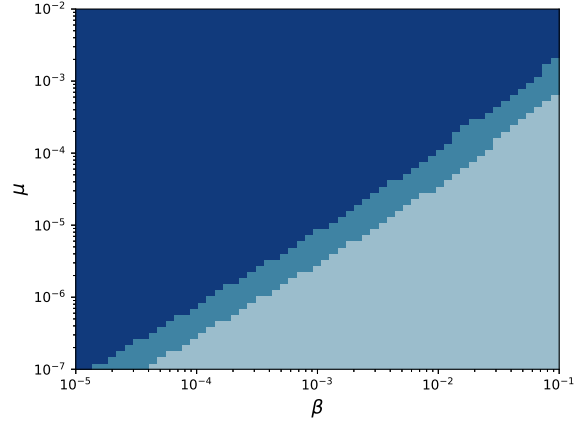
Figure A.8: A monotonic selection landscape produces a qualitatively similar phase plot to that generated with a non-monotonic landscape.

equation for p_i, the probability of fixation from a starting condition of i mutants. We have

$$
\begin{aligned}
p_i &= T^-(i)p_{i-1} + T^+(i)p_{i+1} + (1 - T^+(i) - T^-(i))p_i \\
&= \frac{\rho}{1+\rho}p_{i-1} + \frac{1}{1+\rho}p_{i+1}
\end{aligned}
\tag{A.14}
$$

where $\rho \equiv T^-(i)/T^+(i) = 1 - s$, and we have used the fact that

$$
\begin{aligned}
T^+(i) &= \frac{i(N-i)}{N - s(N-i)} \\
T^-(i) &= \frac{i(N-i)(1-s)}{N - s(N-i)}.
\end{aligned}
\tag{A.15}
$$

Equation A.14 is a second order linear difference equation with constant coefficients. It has general solution $p_i = c_1 + c_2\rho^i$, and given the boundary conditions $p_0 = 0$ and $p_N = 1$, we can solve for the constants c_1 and c_2 to arrive at

$$
p_i = \frac{1 - \rho^i}{1 - \rho^N}.
\tag{A.16}
$$

For $i = 1$ and $\rho = 1 - s$, we therefore have

$$
P_{fix}(s) \equiv p_1 = \frac{s}{1 - (1-s)^N}.
\tag{A.17}
$$

93
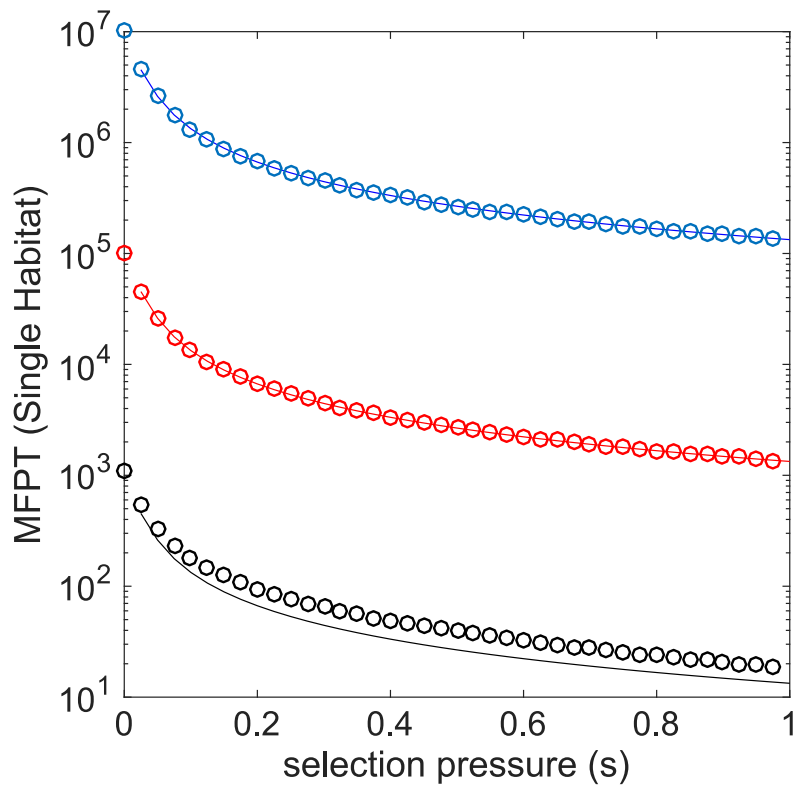
Figure A.9: Within a single vial system, the fixation time is well-approximated for sufficiently small μ by Equation A.13. Circles: exact calculation; solid line: approximation. Mutation rates are $\mu = 10^{-7}$ (blue), $\mu = 10^{-5}$ (red), and $\mu = 10^{-3}$ (black). $N = 75$ for all curves. Fixation times are measured in units of N^{-1}.

A.7 Analytical Approximation for Multiple Habitats

To derive an approximate expression for τ_f in the $M = 3$ vial array, we again consider the limit where the arrival of the first mutant in each vial dominates the fixation time, which now corresponds to $\mu, \beta \ll 1$. In this limit, the time to fixation within a single vial—given the presence of a single initial mutant—is again fast compared to the arrival time of that first mutant, which now can be due to either a mutation event (at rate μN) or a migration event (at rate $N\beta n_{fix}$, where n_{fix} is the integer number of (connected) neighboring vials that have already achieved fixation). Therefore, the different habitats effectively achieve fixation one at a time, and the fixation time within each vial is an exponential process governed by a rate

$$\lambda(s, n_{fix}) = N(\mu + \beta n_{fix})P_{fix}(s) \tag{A.18}$$

with $P_{fix}(s)$ again given by Equation A.17. For economy of notation, we introduce the shorthand $\lambda_i(n_{fix}) \equiv \lambda(s(x_i), n_{fix})$. The global fixation time, which depends on the distribution of selection pressures in the three habitats, can be approximated by

$$\tau_f \approx \tau_{min} + Q(x_0)\tau_{max}^\beta(x_1, x_2) + Q(x_1)\tau_{max}^\beta(x_0, x_2) + Q(x_2)\tau_{max}^\beta(x_0, x_1). \tag{A.19}$$

τ_{min} is the fixation time of the fastest vial in the absence of migration and is given by the expected minimum of three independent, exponentially distributed variables with rates $\lambda_i(0)$,

$$\tau_{min} = \frac{1}{\lambda_0(0) + \lambda_1(0) + \lambda_2(0)}. \tag{A.20}$$

$Q(x_i)$ is the probability that the vial at location x_i was the first to reach fixation in the absence of migration, and it is given by

$$Q(x_i) = \frac{\lambda_i(0)}{\sum_{j=0}^2 \lambda_j(0)} \tag{A.21}$$

The β dependence of Equation A.19 is contained in the terms $\tau_{max}^\beta(x_i, x_j)$, which also implicitly contain information about the connection topology. $\tau_{max}^\beta(x_i, x_j)$ is the expected maximum time of two independent fixation events occurring in habitats x_i and x_j, assuming that the third habitat $(x_k \neq x_i, x_j)$ has already achieved fixation and can therefore supply seed mutants at a rate βN. For example, we have

$$\tau_{max}^\beta(x_0, x_1) = \frac{1}{\lambda_0(0) + \lambda_1(1)} + \frac{\lambda_0(0)}{\lambda_0(0) + \lambda_1(1)}(\lambda_1(2))^{-1} + \frac{\lambda_1(1)}{\lambda_0(0) + \lambda_1(1)}(\lambda_0(1))^{-1} \tag{A.22}$$

where the first term is the expected minimum time of two independent processes occurring at rates $\lambda_0(0)$ and $\lambda_1(1)$; note that since the vial at x_2 has already achieved fixation, the rate for the vial at x_1 is taken at $n_{fix} = 1$, which accounts for migration from vial x_2 to x_1. The second term in Equation A.22 is a product of two terms: the probability that the vial at x_0 achieves fixation before the vial at x_1, and the rate at which the vial at x_1 would achieve fixation $(1/\lambda_1(2))$. The latter rate corresponds to $n_{fix} = 2$, as vials at x_0 and x_2 have already achieved fixation and can both contribute βN to the arrival rate of the first mutant. Finally, the last term is a product of two terms: the probability that the vial at x_1 achieves fixation before the vial at x_0, and the rate at which the vial at x_0 would achieve fixation $(1/\lambda_0(1))$.

Because of the symmetry of the selection pressure profiles considered here ($s(x_0) = s(x_2)$), we have $\tau_{max}^{\beta}(x_0, x_1) = \tau_{max}^{\beta}(x_1, x_2)$, while the remaining term in Equation A.19 is given by

$$
\begin{aligned}
\tau_{max}^{\beta}(x_0, x_2) &= \frac{1}{\lambda_0(1) + \lambda_2(1)} + \frac{\lambda_0(1)}{\lambda_0(1) + \lambda_2(1)}(\lambda_2(1))^{-1} + \frac{\lambda_2(1)}{\lambda_0(1) + \lambda_2(1)}(\lambda_0(1))^{-1} \\
&= \frac{3}{2\lambda_0(1)}
\end{aligned}
\tag{A.23}
$$

where the last line accounts for the symmetry of the selection profile.

By combining Equation A.19 with Equations A.17, A.18, and A.20-A.23, we arrive at a general equation for τ_f that depends on the selection pressure profile, μ, β, and N. The general expression is long and cumbersome, but it is straightforward to numerically evaluate τ_f for any value of the parameters. We find that the approximation performs surprisingly well over a large range of parameter values, qualitatively reproducing all features of the full MFPT calculation (Figure A.10, main panel; Compare to Figure 2, main text) and even providing excellent quantitative agreement in many cases (Figure A.10, bottom panels). To gain additional analytical insight into the model, we consider in what follows several limiting cases where additional analytical progress is tractable.

A.7.1 $\beta \ll \mu$ Limit

To investigate the limit $\beta \ll \mu$, we expand Equation A.19 in the small parameter $\epsilon_1 \equiv \beta/\mu$ and neglect terms of order ϵ_1 and higher. In this limit, τ_f in Equation A.19 reduces to τ_{max}, given by

$$
\tau_{max} = \frac{2}{\lambda_0(0)} + \frac{1}{\lambda_1(0)} - \frac{2}{\lambda_0(0) + \lambda_1(0)} - \frac{1}{\lambda_0(0) + \lambda_2(0)} + \frac{1}{2\lambda_0(0) + \lambda_1(0)},
\tag{A.24}
$$

which is the expected maximum of three independent exponentially distributed random variables; note that we have again taken into account the symmetry in the selection pressure profile ($\lambda_0 = \lambda_2$). In this limit, the three vial system acts effectively as three independent systems, with the overall fixation time corresponding to the slowest fixation. After rewriting τ_{max} in terms of $\langle s \rangle$
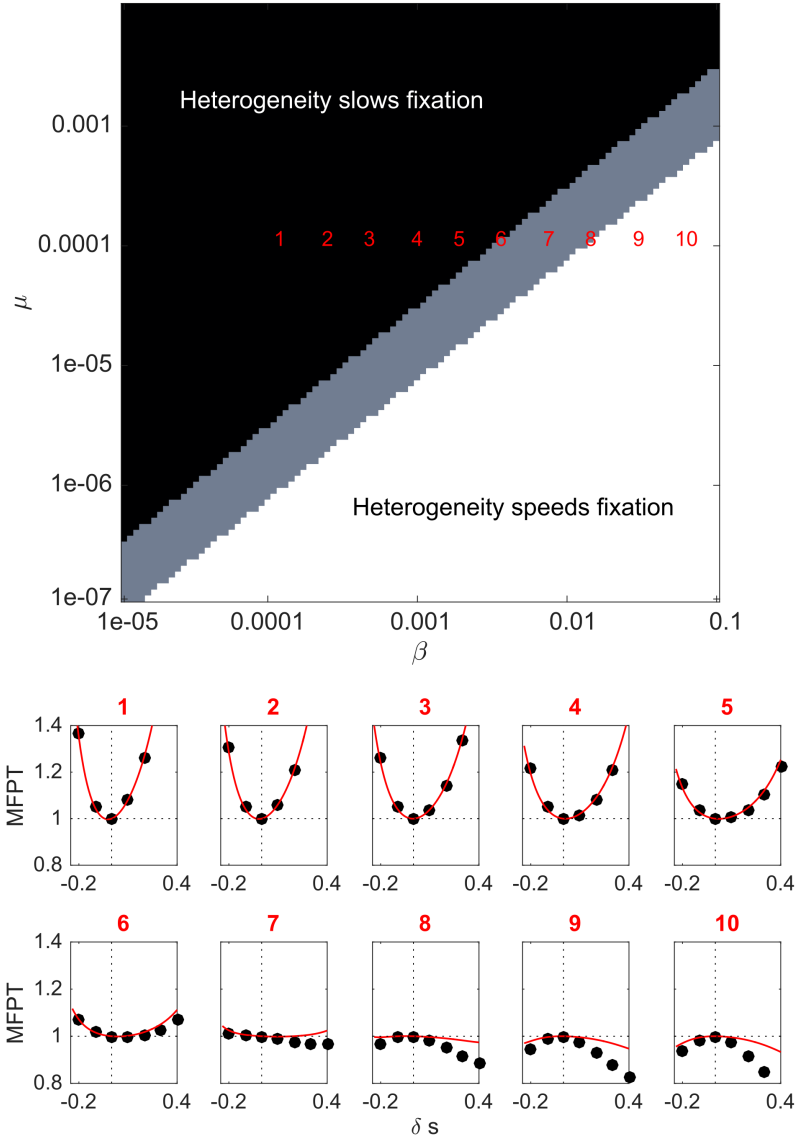
Figure A.10: Parameter regimes and fixation time curves calculated using approximate MFPT (compare to Figure 2, main text). MFPT approximations were performed for the indicated values of β and μ and for $-0.2 \leq \delta s \leq 0.5$ in steps of 0.025. Red numbers on the main plot correspond to curves shown in the subplots below. Black circles: exact calculation. Red curves: approximation.
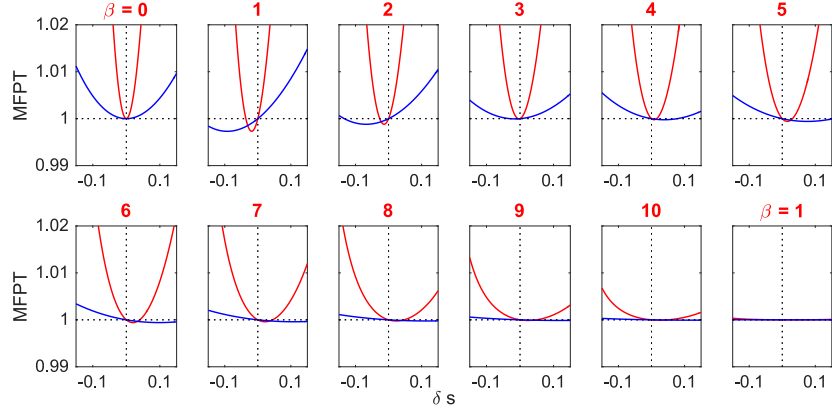
Figure A.11: Fixation times near the origin as a function of δs in the $N \to \infty$ limit. β and μ values for panels labeled 1-10 are the same as in Figure A.10; we've also included $\beta = 0$ and $\beta = 1$. $\langle s \rangle = 1/6$ (red) and $\langle s \rangle = 5/6$ (blue).

and δs, it is straightforward (though algebraically tedious) to show that $(\partial \tau_{max}/\partial \delta s)|_{\delta s=0} = 0$ and $(\partial^2 \tau_{max}/\partial \delta s^2)|_{\delta s=0} > 0$, indicating that the homogeneous landscape ($\delta s = 0$) minimizes the fixation time, consistent with results of the exact calculation. Intuitively, increasing heterogeneity reduces the minimum selection pressure in the spatial array, which in turn slows the expected maximum fixation time among the three habitats.

A.7.2 $\quad \beta \gg \mu$ Limit

To investigate the limit $\beta \gg \mu$, we expand Equation A.19 in the small parameter $\epsilon_2 \equiv \mu/\beta$. The dominant term is of order $1/\epsilon_2$, and we ignore terms of order unity and higher. In the limit $\beta \gg \mu$, τ_f in Equation A.19 reduces to the expected minimum of three independent exponential processes, leading to $\tau_f \approx \tau_{min}$ (Equation A.20). In this limit, the fixation time is dominated by dynamics in the vial that first achieves fixation; the remaining vials then rapidly achieve fixation due to fast migration. For large but finite N, the fixation time τ_{min} is maximized at $\delta s = 0$, indicating that heterogeneity always accelerates fixation, again consistent with the exact calculation (Fig. 2b, right panel). In this limit, the effective rate of fixation λ_{eff} is increased for all $\delta \neq 0$, as heterogeneity decreases fixation time in the vial with the fastest average fixation.

A.7.3 Large N Limit

While our primary focus is on finite populations, we briefly consider here the limit of $N \to \infty$, where Equation A.19 reduces (after simplification) to

$$\tau_f^\infty \approx \frac{1}{\mu(2s_0 + s_1)}\left[1 + \frac{3\mu s_1}{2(\beta + \mu)s_0} + \frac{2\mu\left(2\beta s_1(s_0 + s_1) + \mu(s_0^2 + s_0 s_1 + s_1^2)\right)}{(2\beta + \mu)s_1(\beta s_1 + \mu(s_0 + s_1))}\right] \quad \text{(A.25)}$$

Note that prior to taking the $N \to \infty$ limit, we first rescaled our time units by a factor of N. We've also used the shorthand $s_i \equiv s(x_i)$ and consider only symmetric landscapes ($s_0 = s_2$).

In the limit $s_1/s_0 \to 0$ (large selection valley), Equation A.25 reduces to $\tau_f^\infty = 1/(s_1(2\beta + \mu))$, which corresponds to selection in a single vial of selection pressure s_1 where seed mutants can arise due to mutation (μ) or due to migration from one of two vials that have already achieved fixation (2β). Fixation in the center vial is the rate limiting step for global fixation. Similarly, when $s_0/s_1 \to 0$ (large selection peak), we have $\tau_f^\infty = 3/(2s_0(\beta + \mu))$, which is the maximum of two independent exponential processes that each occur at rate $s_0(\beta + \mu)$. Fixation in the peripheral vials is the rate-limiting process.

To evaluate the magnitude of the change in fixation time as the selection profile is modulated, we fix the average selection pressure $\langle s \rangle$ and calculate the difference in fixation times $\Delta\tau$ between a selection profile with a large peak ($s_0 = s_2 \sim \epsilon_0$) and one with a large valley ($s_1 \sim 2\epsilon_0$), where $\epsilon_0 \ll 1$ is assumed to be small so that the vast majority of the selection pressure is concentrated either in the center vial or the edge vials. By expanding Equation A.25 in ϵ_0 and ignoring terms of order ϵ_0 or higher, we have

$$\frac{\Delta\tau}{\tau_v} = \frac{\tau_p - \tau_v}{\tau_v} = \frac{5 + 2\gamma}{1 + \gamma} > 0, \quad \text{(A.26)}$$

where τ_v (τ_p) corresponds to the selection profile with a large valley (peak) and $\gamma \equiv \mu/\beta$. The fixation time is always larger (slower) in a sharply peaked landscape than one with a sharp valley. Choosing the peaked landscape over the valley leads to a relative increase of 2-5 fold depending on the ratio of mutation to migration.

Following a change of variables in Equation A.25 from $\{s_i\} \to \langle s \rangle$ and δs, we find that the homogeneous landscape is an optimum ($(\partial\tau_f^\infty/\partial\delta s) = 0$) only for $\beta = 0$ and $\beta = \mu\left(7 + \sqrt{145}\right)/4$, and each corresponds to a minimum of the fixation time. Surprisingly, then, all other values of β correspond to the intermediate regime, where heterogeneity can either slow or speed fixation depending on the specific landscape. In practice, the value of δs for which a minimum occurs is often very close to zero, though it does depend on $\langle s \rangle$ (Figure A.11). Large heterogeneity–either a peak ($\delta s > 0$) or a valley ($\delta s < 0$)–tends to slow fixation. For $\beta \gg \mu$, the fixation time reduces to $\tau_f^\infty \approx (3\mu\langle s \rangle)^{-1}$ and no longer depends on δs; this corresponds to the fixation time of a single habitat with effective selection pressure $s_e = 3\langle s \rangle$, where the factor of 3 results from a tripling of

the population size (relative to a single vial).

APPENDIX B

Generalizing Toy Model for Imperfect Mutants

B.1 $M = 3$ MFPT Transition Rates

Here we enumerate the explicit transition rates for our system with $M = 3$ microhabitats. The transition rates depend on the system parameters: the mutation rate μ, the migration rate β, and the total population N, in addition to the wild-type fitness $r_0(x_i)$ and fitness cost ϵ such that $r^*(x_i) = 1 - \epsilon$. Note that these transition rates are the same as those in Appendix A.1 upon reducing $r^*(x_i)$ by ϵ and assuming weights $w(x_i) = 1$. The transitions increasing the number of mutants in microhabitat x_i starting from an initial state (j, k, ℓ) are then given by

$$
T^+[x_0](j, k, \ell) = \underbrace{(N-j)}_{\text{WT dies}} \underbrace{\frac{(N-j)r_0(x_0)}{(N-j)r_0(x_0) + j(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\mu}_{\text{WT mutates}}
$$
$$
+ \underbrace{(N-j)}_{\text{WT dies}} \underbrace{\frac{j(1-\epsilon)}{(N-j)r_0(x_0) + j(1-\epsilon)}}_{\text{Pick mutant}} \underbrace{(1-\beta)}_{\text{No migration}}
$$
$$
+ \underbrace{(N-j)}_{\text{WT dies}} \underbrace{\frac{k(1-\epsilon)}{(N-k)r_0(x_1) + k(1-\epsilon)}}_{\text{Pick right mutant}} \underbrace{\beta}_{\text{Migration}} , \quad \text{(B.1)}
$$

$$T^+[x_1](j,k,\ell) = \underbrace{(N-k)}_{\text{WT dies}} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1)+k(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\mu}_{\text{WT mutates}}$$

$$+ \underbrace{(N-k)}_{\text{WT dies}} \underbrace{\frac{k(1-\epsilon)}{(N-k)r_0(x_1)+k(1-\epsilon)}}_{\text{Pick mutant}} \underbrace{(1-2\beta)}_{\text{No migration}}$$

$$+ \underbrace{(N-k)}_{\text{WT dies}} \underbrace{\frac{j(1-\epsilon)}{(N-j)r_0(x_0)+j(1-\epsilon)}}_{\text{Pick left mutant}} \underbrace{\beta}_{\text{Migration}}$$

$$+ \underbrace{(N-k)}_{\text{WT dies}} \underbrace{\frac{\ell(1-\epsilon)}{(N-\ell)r_0(x_2)+\ell(1-\epsilon)}}_{\text{Pick right mutant}} \underbrace{\beta}_{\text{Migration}} \, , \quad \text{(B.2)}$$

and

$$T^+[x_2](j,k,\ell) = \underbrace{(N-\ell)}_{\text{WT dies}} \underbrace{\frac{(N-\ell)r_0(x_2)}{(N-\ell)r_0(x_2)+\ell(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\mu}_{\text{WT mutates}}$$

$$+ \underbrace{(N-\ell)}_{\text{WT dies}} \underbrace{\frac{\ell(1-\epsilon)}{(N-\ell)r_0(x_2)+\ell(1-\epsilon)}}_{\text{Pick mutant}} \underbrace{(1-\beta)}_{\text{No migration}}$$

$$+ \underbrace{(N-\ell)}_{\text{WT dies}} \underbrace{\frac{k(1-\epsilon)}{(N-k)r_0(x_1)+k(1-\epsilon)}}_{\text{Pick left mutant}} \underbrace{\beta}_{\text{Migration}} \, . \quad \text{(B.3)}$$

The corresponding transition rates decreasing the number of mutants in the microhabitat x_i are given by

$$T^-[x_0](j,k,\ell) = \underbrace{j}_{\text{Kill mutant}} \underbrace{\frac{(N-j)r_0(x_0)}{(N-j)r_0(x_0)+j(1-\epsilon)}}_{\text{Choose WT}} \underbrace{(1-\beta-\mu)}_{\text{WT grows without mutation or migration}}$$

$$+ \underbrace{j}_{\text{Kill mutant}} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1)+k(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Right migration}} \, , \quad \text{(B.4)}$$

$$T^-[x_1](j,k,\ell) = \underbrace{k}_{\text{Kill mutant}} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1)+k(1-\epsilon)}}_{\text{Choose WT}} \underbrace{(1-2\beta-\mu)}_{\text{WT grows without mutation or migration}}$$

$$+ \underbrace{k}_{\text{Kill mutant}} \underbrace{\frac{(N-j)r_0(x_0)}{(N-j)r_0(x_0)+j(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Left migration}}$$

$$+ \underbrace{k}_{\text{Kill mutant}} \underbrace{\frac{(N-\ell)r_0(x_2)}{(N-\ell)r_0(x_2)+\ell(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Right migration}} \,, \quad \text{(B.5)}$$

and

$$T^-[x_2](j,k,\ell) = \underbrace{\ell}_{\text{Kill mutant}} \underbrace{\frac{(N-\ell)r_0(x_2)}{(N-\ell)r_0(x_2)+\ell(1-\epsilon)}}_{\text{Choose WT}} \underbrace{(1-\beta-\mu)}_{\text{WT grows without mutation or migration}}$$

$$+ \underbrace{\ell}_{\text{Kill mutant}} \underbrace{\frac{(N-k)r_0(x_1)}{(N-k)r_0(x_1)+k(1-\epsilon)}}_{\text{Pick WT}} \underbrace{\beta}_{\text{Left migration}} \,. \quad \text{(B.6)}$$

B.2 Effect of Fitness Cost on Homogeneous Landscapes

To characterize the difference in the transition rates upon the addition of a fitness cost, we define

$$\Delta_\epsilon^+[x_i](j,k,\ell) \equiv T_\epsilon^+[x_i](j,k,\ell) - T_{\epsilon=0}^+[x_i](j,k,\ell), \quad \text{(B.7)}$$

$$\Delta_\epsilon^-[x_i](j,k,\ell) \equiv T_\epsilon^-[x_i](j,k,\ell) - T_{\epsilon=0}^-[x_i](j,k,\ell). \quad \text{(B.8)}$$

For notational convenience, we use the shorthand $\Delta^+[x_i] = \Delta_\epsilon^+[x_i](j,k,\ell)$ and $\Delta^-[x_i] = \Delta_\epsilon^-[x_i](j,k,\ell)$, where we have omitted the implied dependence on fitness cost ϵ and state (j,k,ℓ).

Looking first at the transition rates increasing the number of mutants, we will explicitly evaluate $\Delta^+[x_1]$. The difference in the transition rates for the extrema microhabitats will be similar but less complicated, since they are only connected to a single neighboring microhabitat. We first write the transition rates in terms of the fitness ratio

$$f \equiv \frac{r_0}{1-\epsilon}. \quad \text{(B.9)}$$

When both $\langle s \rangle$ and ϵ are specified, r_0 is completely determined for the homogeneous landscape. We must have

$$\langle s \rangle = (1-\epsilon) - r_0, \quad \text{(B.10)}$$

103

so that

$$r_0 = 1 - (\epsilon + \langle s \rangle), \tag{B.11}$$

for the homogeneous landscape. Then our fitness ratio becomes

$$f = \frac{1 - (\epsilon + \langle s \rangle)}{1 - \epsilon}. \tag{B.12}$$

We now write the positive transition rate as a function of this fitness ratio,

$$T_f^+[x_1](j,k,\ell) = (N-k)\left[\frac{N-k}{(N-k)+kf^{-1}}\mu + \frac{k}{(N-k)f+k}(1-2\beta) \right.$$
$$\left. + \frac{j}{(N-j)f+j}\beta + \frac{\ell}{(N-\ell)f+\ell} \right]. \tag{B.13}$$

Then we have

$$\Delta^+[x_1] = T_{f(\epsilon)}^+[x_1](j,k,\ell) - T_{f(\epsilon=0)}^+[x_1](j,k,\ell), \tag{B.14}$$

which is complicated when written out due to the non-linearity of the transition rates on the fitness ratio. Instead, we approximate the above expression using a Taylor expansion. Because enforcing conservation of $\langle s \rangle$ requires that r_0 also changes with ϵ, $f(\epsilon)$ changes relatively little across the range of fitness costs of interest. For this reason, we expand $f(\epsilon)$ about $\epsilon = 0$ to write

$$f(\epsilon) \simeq (1 - \langle s \rangle) - \langle s \rangle \epsilon - \langle s \rangle \epsilon^2 + \mathcal{O}(\epsilon^3). \tag{B.15}$$

Using this approximation, we can write

$$\Delta^+[x_1](j,k,\ell) = \langle s \rangle (N-k)\left[\frac{(N-j)j}{(\langle s \rangle(j-N)+N)^2}\beta + \frac{(N-k)k}{(\langle s \rangle(k-N)+N)^2}(1-2\beta-\mu) \right.$$
$$\left. + \frac{\ell(N-\ell)}{(\langle s \rangle(\ell-N)+N)^2}\beta \right]\epsilon + \mathcal{O}(\epsilon^2). \tag{B.16}$$

Note that in the above expression, even term is greater than or equal to zero. So this suggests that the transition rate increasing the mutants can only be larger with a non-zero fitness cost.

We now look at the transition rates decreasing the number of mutants. In terms of the fitness

ratio, this is given by

$$T_f^-[x_1](j, k, \ell) = k\left[\frac{N - k}{(N - k) + kf^{-1}}(1 - 2\beta - \mu) + \frac{N - j}{(N - j) + jf^{-1}}\beta\right.$$

$$\left. + \frac{N - \ell}{(N - \ell) + \ell f^{-1}}\beta\right]. \quad \text{(B.17)}$$

Upon expanding $f(\epsilon)$ about $\epsilon = 0$, we find

$$\Delta^-[x_1](j, k, \ell) = k\left[-\frac{k(N - k)\langle s\rangle}{(\langle s\rangle(k - N) + N)^2}(1 - \mu - 2\beta)\right.$$

$$\left. - \langle s\rangle\beta\left(\frac{(N - j)j}{(\langle s\rangle(j - N) + N)^2} + \frac{(N - \ell)\ell}{(\langle s\rangle(\ell - N) + N)^2}\right)\right]\epsilon + \mathcal{O}(\epsilon^2) \quad \text{(B.18)}$$

Now each of these terms is less than or equal to zero, so the transition rate for decreasing the number of mutants can only be smaller with the addition of a fitness cost.

Together these results explain the trend that we are seeing—with a non-zero fitness cost, the system is more likely to increase the number of mutants and also less likely to decrease the number of mutants. These two observations combine to predict that the offset term (the fixation time in a homogeneous landscape) should decrease as the fitness cost increases. Indeed, this is exactly what we observe in our data.

As an aside, we can further simplify each of the differences by taking both $\mu \to 0$ and $\beta \to 0$ to arrive at

$$\Delta^+[x_1](j, k, \ell) = \langle s\rangle\frac{(N - k)^2 k}{(\langle s\rangle(k - N) + N)^2}\epsilon, \quad \text{(B.19)}$$

$$\Delta^-[x_1](j, k, \ell) = -\langle s\rangle\frac{k^2(N - k)}{(\langle s\rangle(k - N) + N)^2}\epsilon, \quad \text{(B.20)}$$

which is a valid approximation across almost all parameters of interest. Upon taking this limit, the dependence on the microhabitat x_i disappears, so that the same expressions quantify the effect of adding a fitness cost to the transition rates in all of the microhabitats of the system.

B.3 Physical Landscape Geometry

Recall that selection pressure is defined as the difference between the mutant and wild-type fitness values:

$$s(x_i) = r^*(x_i) - r_0(x_i). \tag{B.21}$$

With a fitness cost ϵ, we have $r^* = 1 - \epsilon$. If we have an offset of δs in the selection pressure for the center microhabitat, we can write

$$r_0(x_i) = r_0 - \delta s \delta_{x_i, x_1}, \tag{B.22}$$

where δ_{x_i, x_1} is the Kronecker delta between microhabitat x_i and x_1. With this definition, we can write

$$S_{\text{total}} = \sum_i s(x_i), \tag{B.23}$$

$$= 3(1 - \epsilon) - 3r_0 + \delta s, \tag{B.24}$$

so that

$$3\langle s \rangle = 3(1 - \epsilon) - 3r_0 + \delta s. \tag{B.25}$$

This equation can be solved to get

$$r_0 = \frac{3(1 - \epsilon - \langle s \rangle) + \delta s}{3}. \tag{B.26}$$

B.3.1 Valid Range of δs

We first derive the range of valid δs values for a given choice of $\langle s \rangle$ and ϵ. Here we derive the bounds on the landscape of physical selection pressure landscapes. Starting with the center microhabitat, the fitness value is physical when it satisfies

$$0 \le r_0(x_1) \le 1. \tag{B.27}$$

Using our expression for r_0, this becomes

$$0 \le \frac{3(1 - \epsilon - \langle s \rangle) + \delta s}{3} - \delta s \le 1. \tag{B.28}$$

After some algebra, one finds

$$-\frac{3}{2}(\epsilon + \langle s \rangle) \leq \delta s \leq \frac{3}{2}(1 - \epsilon - \langle s \rangle). \tag{B.29}$$

Enforcing that the center microhabitat have physical fitness values produces this constraint on δs. We also require physical fitness values in our extrema microhabitats. That is, we require

$$0 \leq r_0 \leq 1. \tag{B.30}$$

Using our expression for r_0, one can derive the condition

$$-3(1 - \epsilon - \langle s \rangle) \leq \delta s \leq 3(\epsilon + \langle s \rangle). \tag{B.31}$$

Since a fitness landscape is only valid if all three of the microhabitats have physical fitness values, one can identify

$$\delta s_{min} = \max\left\{ -\frac{3}{2}(\langle s \rangle + \epsilon), -3(1 - \langle s \rangle - \epsilon) \right\}, \tag{B.32}$$

$$\delta s_{max} = \min\left\{ \frac{3}{2}(1 - \langle s \rangle - \epsilon), 3(\langle s \rangle + \epsilon) \right\}, \tag{B.33}$$

such that $\delta s_{min} \leq \delta s \leq \delta s_{max}$.

B.3.2 Neutral Genetic Drift

We next identify the values of δs that lead to neutral genetic drift for a given choice of $\langle s \rangle$ and ϵ. For any microhabitat, we achieve neutral genetic drift when

$$r_0(x_i) = 1 - \epsilon, \tag{B.34}$$

is satisfied. For the center microhabitat, this expression becomes

$$r_0 - \delta s = 1 - \epsilon, \tag{B.35}$$

$$\frac{3(1 - \epsilon - \langle s \rangle) + \delta s}{3} - \delta s = 1 - \epsilon, \tag{B.36}$$

$$\frac{-3\langle s \rangle + \delta s}{3} - \delta s = 0, \tag{B.37}$$

so that the condition for neutral genetic drift is independent of the fitness cost ϵ. Algebraic manipulations of this expression lead to the following condition for neutral genetic drift in the center

microhabitat:

$$-\frac{3}{2}\langle s \rangle = \delta s. \tag{B.38}$$

We can now determine when this condition results in a physical selection landscape. Physicality requires

$$0 \leq r_0(x_i) \leq 1. \tag{B.39}$$

For the center microhabitat, with neutral genetic drift we have $r_0(x_1) = 1 - \epsilon$ by construction, so this fitness value is physical for $0 \leq \epsilon < 1$. Because the fitness values of the extrema microhabitats are connected to that of the center microhabitat through imposing conservation of $\langle s \rangle$, we also need to check when these microhabitats have physical fitness values. These microhabitats have $r_0(x_0) = r_0(x_2) = r_0$, which must satisfy

$$0 \leq \frac{3(1 - \epsilon - \langle s \rangle) + \delta s}{3} \leq 1. \tag{B.40}$$

If we now impose $\delta s = -3\langle s \rangle / 2$ to enforce neutral genetic drift, this becomes

$$0 \leq \frac{3(1 - \epsilon - \langle s \rangle) - 3\langle s \rangle/2}{3} s \rangle \leq 1. \tag{B.41}$$

After some rearranging, one finds

$$-\frac{2}{3}\epsilon \leq \langle s \rangle \leq \frac{2}{3}(1 - \epsilon). \tag{B.42}$$

This constrains the values of $\langle s \rangle$ satisfying the neutrality condition that result in a physical fitness value for all of the microhabitats in our system.

We can perform the same analysis for the extrema microhabitats x_0 and x_2. For these microhabitats, we have $r_0(x_0) = r_0(x_2) = r_0$. Using our previous expression for r_0, neutral genetic drift is achieved in the extrema microhabitats when the system satisfies

$$\frac{3(1 - \epsilon - \langle s \rangle) + \delta s}{3} = 1 - \epsilon. \tag{B.43}$$

After some algebra, this condition can be rewritten as

$$\langle s \rangle = \frac{\delta s}{3}. \tag{B.44}$$

We can now determine when this leads to physical fitness values. The center microhabitat must

satisfy

$$0 \leq r_0 - \delta s \leq 1, \tag{B.45}$$

$$0 \leq (1 - \epsilon) - 3\langle s \rangle \leq 1, \tag{B.46}$$

upon enforcing neutrality. This can be simplified to constrain $\langle s \rangle$:

$$-\frac{1}{3}\epsilon \leq \langle s \rangle \leq \frac{1}{3}(1 - \epsilon). \tag{B.47}$$

B.3.3   Wild-Type Fitness Deviation

Using the same type of algebraic manipulations, we can write the wild-type fitness as a deviation from the mutant fitness to help us understand the different qualitative behaviors. In the center microhabitat, we have

$$r_0(x_1) = \frac{3(1 - \epsilon - \langle s \rangle) + \delta s}{3} - \delta s, \tag{B.48}$$

$$= (1 - \epsilon) + \left(\frac{\delta s}{3} - \langle s \rangle\right) - \delta s, \tag{B.49}$$

$$= (1 - \epsilon) + \left(-\frac{2}{3}\delta s - \langle s \rangle\right). \tag{B.50}$$

This allows us to identify

$$d[x_1](\delta s, \langle s \rangle) = -\frac{2}{3}\delta s - \langle s \rangle. \tag{B.51}$$

The expression for the wild-type fitness can also be written as a deviation from the mutant fitness. For the extrema microhabitats, we can write

$$r_0(x_0) = (1 - \epsilon) + \frac{\delta s}{3} - \langle s \rangle, \tag{B.52}$$

from which we identify the deviation

$$d[x_0](\delta s, \langle s \rangle) = \frac{\delta s}{3} - \langle s \rangle. \tag{B.53}$$

Due to the underlying symmetry of the selection pressure landscape, we have $d[x_2] = d[x_0]$.

APPENDIX C

Generalizing Toy Model to Larger Systems

C.1 Generalized Recursion Relation for Fixation Time Approximation

The fixation times are approximated by coarse-graining over all of intermediate states for each microhabitat, so that it is either fixated or unfixated. We neglect the time required for genetic drift to lead to fixation, so that the evolutionary time within a single microhabitat is insignificant after the first mutant arrives, provided that genetic drift does lead to fixation. This assumption only holds quantitatively when we are in the limit $\mu \ll 1$ and $\beta \ll 1$, so that the time to fixation is dominated by the arrival of the first mutant into the system (with $\mu \ll \beta$) or the time for the final microhabitat to reach fixation (with $\beta \ll \mu$), although we observed that this approximation worked well qualitatively outside of this approximation for $M = 3$. We treat fixation within each microhabitat as as exponential process with rate

$$\lambda_i(n) = N(\mu + \beta n)\mathcal{P}(s(x_i)), \tag{C.1}$$

where n is the number of fixated neighboring microhabitats and

$$\mathcal{P}(s(x_i)) = \frac{s(x_i)}{1 - (1 - s(x_i))^N}, \tag{C.2}$$

is the probability of the microhabitat x_i reaching fixation upon the introduction of a single mutant. The expected time for the first microhabitat to reach fixation through a *de novo* mutation is given by

$$\tau_{\min} = \frac{1}{\lambda_0(0) + \lambda_1(0) + \lambda_2(0)}, \tag{C.3}$$

which is the expected minimum of three exponential processes. Using this result, the fixation times can be approximated as

$$\hat{\tau}_f = \tau_{min} + Q(x_0)\tau^\beta_{\max}(x_0) + Q(x_1)\tau^\beta_{\max}(x_1) + Q(x_2)\tau^\beta_{\max}(x_2), \tag{C.4}$$

if $Q(x_i)$ represents the probability that microhabitat x_i is the first to reach fixation, given by

$$Q(x_i) = \frac{\lambda_i(0)}{\lambda_0(0) + \lambda_1(0) + \lambda_2(0)}, \tag{C.5}$$

and $\tau^\beta_{\max}$ denotes the expected maximum time for the remaining microhabitats to reach fixation. With $M = 3$, this term is simple enough that it can be explicitly written out:

$$\tau^\beta_{\max}(x_0) = \frac{1}{\lambda_1(1) + \lambda_2(0)} + \frac{\lambda_2(0)}{\lambda_1(1) + \lambda_2(0)}(\lambda_1(2))^{-1} + \frac{\lambda_1(1)}{\lambda_1(1) + \lambda_2(0)}(\lambda_2(1))^{-1}, \tag{C.6}$$

$$\tau^\beta_{\max}(x_1) = \frac{1}{\lambda_0(1) + \lambda_2(1)} + \frac{\lambda_0(1)}{\lambda_0(1) + \lambda_2(1)}(\lambda_2(1))^{-1} + \frac{\lambda_2(1)}{\lambda_0(1) + \lambda_2(1)}(\lambda_0(1))^{-1}, \tag{C.7}$$

and $\tau^\beta_{\max}(x_0) = \tau^\beta_{\max}(x_2)$ by symmetry.

We now wish to expand this approximation to hold for arbitrary values of M. To do this, we need to first generalize the notation of the important quantities to allow for multiple fixated microhabitats, $\{x_i\}$. We first express the expected waiting time for the first of the remaining microhabitats to fixate:

$$\tau_{\min}(\{x_i\}) = \frac{1}{\sum_{j \notin \{i\}} \lambda_j(n_j(\{x_i\}))}, \tag{C.8}$$

where n_j is the number of fixated neighbors for microhabitat x_j, given by

$$n_j(\{x_i\}) = \sum_{k \in \{i\}} (\delta_{j,k+1} + \delta_{j,k-1}). \tag{C.9}$$

Similarly, we can generalize the probability of microhabitat x_m being the next to reach fixation, assuming that microhabitats $\{x_i\}$ are already fixed:

$$Q(x_m | \{x_i\}) = \frac{\lambda_m(n_m)}{\sum_{j \neq \{i\}} \lambda_j(n_j)}. \tag{C.10}$$

These these expressions, we can now write our fixation time recursively according to

$$\hat{\tau}_f = \tau_{\max}^\beta(\{\}), \tag{C.11}$$

$$= \tau_{min}(\{\}) + \sum_{i=0}^{M-1} Q(x_i|\{\})\tau_{\max}^\beta(x_i). \tag{C.12}$$

We will then have M terms to evaluate, of the form

$$\tau_{\max}^\beta(x_j) = \tau_{min}(x_j) + \sum_{i \neq j} Q(x_i|x_j)\tau_{\max}^\beta(x_i, x_j), \tag{C.13}$$

$\forall j \in [0, M-1]$. The recursion relation continues until there is a single microhabitat that has not yet reached fixation. If this microhabitat is given by x_ℓ, we have

$$\tau_{\max}^\beta(\mathbb{X}\setminus x_\ell) = \frac{1}{\lambda_\ell(n_\ell(\mathbb{X}\setminus x_\ell))} + \tau_{\max}^\beta(\mathbb{X}), \tag{C.14}$$

since $Q(x_k|\mathbb{X}\setminus x_\ell) = 1$. So as long as we have $\tau_{\max}^\beta(\mathbb{X}) = 0$, our recursion successfully terminates. Note that intuitively, $\tau_{\max}^\beta(\mathbb{X}) = 0$ is obvious since the maximum waiting time once all of the microhabitats are fixated is zero.

C.1.1 Monte Carlo Estimate

Although our recursion relation allows us to write down an equation for the approximated fixation times for arbitrary M, it is not easy to evaluate this approximation. The reason is that the number of terms grows combinatorically with M, since it is effectively calculating the fixation time by assigning a weight to every possible trajectory reaching fixation.

Although we cannot calculate the weight of every trajectory, we can Monte Carlo sample different trajectories to form an estimate of the fixation times within the framework of the approximation. We can do this by drawing random samples from the exponential distributions. For microhabitat x_i, we draw a random sample $\xi_i(n_i) \sim \mathrm{Exp}[\lambda_i(n_i)]$. We can organize these results in an array

$$\Xi(\{\}) = \begin{pmatrix} \xi_0(0) \\ \xi_1(0) \\ \vdots \\ \xi_{M-1}(0) \end{pmatrix} \tag{C.15}$$

The next microhabitat to fixate in our system is the microhabitat with the smallest waiting time.

So the first microhabitat to reach fixation is given by

$$x_{i_1} = \operatorname{argmin}\{\Xi(\{\})\}. \tag{C.16}$$

After this microhabitat has reached fixation, all other microhabitats not neighboring the fixated microhabitat are adjusted to account for the passage of time $\tau_1 = \min\{\Xi(\{\})\}$. The microhabitats neighboring x_{i_1} require new random samples, since their distribution has changed. Following the fixation of the first microhabitat, our array is given by

$$\Xi(\{x_{i_1}\}) = \begin{pmatrix} \xi_0(0) - \tau_1 \\ \vdots \\ \xi_{i_1-2}(0) - \tau_1 \\ \xi_{i_1-1}(1) \sim \operatorname{Exp}[\lambda_{i-1}(1)] \\ \infty \\ \xi_{i_1+1}(1) \sim \operatorname{Exp}[\lambda_{i_1+1}(1)] \\ \xi_{i_2+2}(0) - \tau_1 \\ \vdots \\ \xi_{M-1}(0) - \tau_1 \end{pmatrix}. \tag{C.17}$$

This process then continues until all microhabitats have reached fixation. The fixation time for a given trajectory is simply given by summing over the individual random samples: $\tau_f = \sum_{i=1}^{M} \tau_i$.

Each fixation event typically requires us to draw two new random numbers, and we need to modify the remaining entries to subtract off the passage of time. Correcting the passage of time requires $\mathcal{O}(M)$ operations for each new fixation event. The more costly operation is the formation of the new exponential rates and then sampling from the distributions. With efficient random number generation from exponential distributions, each trajectory can be simulated quite quickly. We use GSL to perform efficient random number generation. We require a large number of trajectories to accurately Monte Carlo estimate the fixation times, so we take advantage of the embarrassingly parallel nature of simulating a large number of trajectories by using OpenMP. We find that we can simulate 10^8 trajectories with $M = 9$ in roughly 10 seconds and that the computational time required scales as a power law with M. The speed of this estimation provides us with a general-purpose algorithm to accurately estimate fixation times within the framework of the approximation developed, even if we cannot evaluate the recursion relation itself.

C.2 Bounding Approximating Fixation Times

We can bound the recursion relation below if we assume that the microhabitat with the smallest associated exponential rate is the next to reach fixation deterministically. In the framework of the generalized fixation time approximation, this would be equivalent to assigning probability 1 to the microhabitat with the largest value for Q. If we have fixated microhabitats $\{x_{i_1}, x_{i_2}, \ldots, x_{i_n}\}$, the next microhabitat to reach fixation is given by

$$x_{i_{n+1}} = \text{argmax}_{j \neq \{i_1, \ldots, i_n\}} Q(x_j | \{x_{i_1}, \ldots, x_{i_n}\}), \tag{C.18}$$

$$= \text{argmax}_{j \neq \{i_1, \ldots, i_n\}} \lambda_j(n_j(\{x_{i_1}, \ldots, x_{i_n}\})). \tag{C.19}$$

With this large simplification, the recursion relationship would be given by

$$\tau_{\text{max}}^\beta(\{x_{i_1}, \ldots, x_{i_n}\}) = \tau_{\text{min}}(\{x_{i_1}, \ldots, x_{i_n}\}) + \tau_{\text{max}}^\beta(\{x_{i_1}, \ldots, x_{i_n}\} \cup x_{i_{n+1}}). \tag{C.20}$$

There are no conditions on when this truncation of the generalized recursion relation can be applied. We chose to weight the contribution from all of the microhabitats being the first to reach fixation but used the truncation after a single microhabitat reached fixation. One can achieve a tighter lower bound by delaying the truncation until a later depth is reached but at the cost of more computation. We were satisfied with our results evaluating M truncated recursion relations.

In the same manner, one can form the upper bound by instead choosing the microhabitat with the largest exponential rate as the next to reach fixation. The math is identical to the above derivation, but with

$$x_{i_{n+1}} = \text{argmin}_{j \neq \{i_1, \ldots, i_n\}} \lambda_j(n_j(\{x_{i_1}, \ldots, x_{i_n}\})). \tag{C.21}$$

Because so much information was discarded in the formation of these bounds, they may appear to be only a mathematical novelty. However, there are specific systems in which we expect these bounds to accurately approximate the fixation times. In the limit $\mu \ll \beta$, when the fixation times are dominated by the arrival of the first mutant, both bounds serve as excellent approximations to the true fixation times. The bounds are also relevant near $\delta s = 0$, since all of the microhabitats have nearly identical exponential rates (for most choices of parameters). In this case, the exact order of events is not important for accurately estimating the time to fixation.

C.3 $M = 5$ MFPT Transition Rates

Here we enumerate over the $M = 5$ transition rates assuming a mutation rate μ, migration rate β, and population size N. If there exist $n^*(x_i)$ mutant cells in microhabitat x_i, each with fitness $r^*(x_i) = 1$, and $N - n^*(x_i)$ wild-type cells with fitness $r_0(x_i)$, the transition rates are given by

$$
\begin{aligned}
T^+[x_0] = (N - n^*(x_0)) &\frac{(N - n^*(x_0))r_0(x_0)}{(N - n^*(x_0))r_0(x_0) + n^*(x_0)}\mu \\
+ (N - n^*(x_0)) &\frac{n^*(x_0)}{(N - n^*(x_0))r_0(x_0) + n^*(x_0)}(1 - \beta) \\
+ (N - n^*(x_0)) &\frac{n^*(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)}\beta \quad \text{(C.22)}
\end{aligned}
$$

$$
\begin{aligned}
T^+[x_1] = (N - n^*(x_1)) &\frac{(N - n^*(x_1))r_0(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)}\mu \\
+ (N - n^*(x_1)) &\frac{n^*(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)}(1 - 2\beta) \\
+ (N - n^*(x_1)) &\frac{n^*(x_0)}{(N - n^*(x_0))r_0(x_0) + n^*(x_0)}\beta \\
+ (N - n^*(x_1)) &\frac{n^*(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)}\beta \quad \text{(C.23)}
\end{aligned}
$$

$$
\begin{aligned}
T^+[x_2] = (N - n^*(x_2)) &\frac{(N - n^*(x_2))r_0(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)}\mu \\
+ (N - n^*(x_2)) &\frac{n^*(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)}(1 - 2\beta) \\
+ (N - n^*(x_2)) &\frac{n^*(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)}\beta \\
+ (N - n^*(x_2)) &\frac{n^*(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)}\beta \quad \text{(C.24)}
\end{aligned}
$$

$$T^+[x_3] = (N - n^*(x_3)) \frac{(N - n^*(x_3))r_0(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)} \mu$$
$$+ (N - n^*(x_3)) \frac{n^*(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)} (1 - 2\beta)$$
$$+ (N - n^*(x_3)) \frac{n^*(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)} \beta$$
$$+ (N - n^*(x_3)) \frac{n^*(x_4)}{(N - n^*(x_4))r_0(x_4) + n^*(x_4)} \beta \quad \text{(C.25)}$$

$$T^+[x_4] = (N - n^*(x_4)) \frac{(N - n^*(x_4))r_0(x_4)}{(N - n^*(x_4))r_0(x_4) + n^*(x_4)} \mu$$
$$+ (N - n^*(x_4)) \frac{n^*(x_4)}{(N - n^*(x_4))r_0(x_4) + n^*(x_4)} (1 - \beta)$$
$$+ (N - n^*(x_4)) \frac{n^*(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)} \beta \quad \text{(C.26)}$$

and

$$T^-[x_0] = n^*(x_0) \frac{(N - n^*(x_0))r_0(x_0)}{(N - n^*(x_0))r_0(x_0) + n^*(x_0)} (1 - \beta - \mu)$$
$$+ n^*(x_0) \frac{(N - n^*(x_1))r_0(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)} \beta \quad \text{(C.27)}$$

$$T^-[x_1] = n^*(x_1) \frac{(N - n^*(x_1))r_0(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)} (1 - 2\beta - \mu)$$
$$+ n^*(x_1) \frac{(N - n^*(x_0))r_0(x_0)}{(N - n^*(x_0))r_0(x_0) + n^*(x_0)} \beta$$
$$+ n^*(x_1) \frac{(N - n^*(x_2))r_0(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)} \beta \quad \text{(C.28)}$$

$$T^-[x_2] = n^*(x_2) \frac{(N - n^*(x_2))r_0(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)} (1 - 2\beta - \mu)$$
$$+ n^*(x_2) \frac{(N - n^*(x_1))r_0(x_1)}{(N - n^*(x_1))r_0(x_1) + n^*(x_1)} \beta$$
$$+ n^*(x_2) \frac{(N - n^*(x_3))r_0(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)} \beta \quad \text{(C.29)}$$
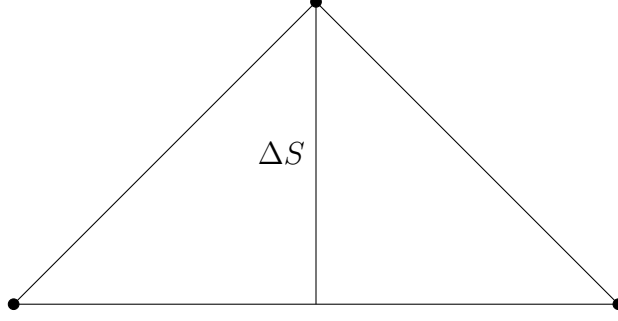
Figure C.1: The $M = 3$ selection landscape to be interpolated.

$$T^-[x_3] = n^*(x_3)\frac{(N - n^*(x_3))r_0(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)}(1 - 2\beta - \mu)$$
$$+ n^*(x_3)\frac{(N - n^*(x_2))r_0(x_2)}{(N - n^*(x_2))r_0(x_2) + n^*(x_2)}\beta$$
$$+ n^*(x_3)\frac{(N - n^*(x_4))r_0(x_4)}{(N - n^*(x_4))r_0(x_4) + n^*(x_4)}\beta \quad \text{(C.30)}$$

$$T^-[x_4] = n^*(x_4)\frac{(N - n^*(x_4))r_0(x_4)}{(N - n^*(x_4))r_0(x_4) + n^*(x_4)}(1 - \beta - \mu)$$
$$+ n^*(x_4)\frac{(N - n^*(x_3))r_0(x_3)}{(N - n^*(x_3))r_0(x_3) + n^*(x_3)}\beta \quad \text{(C.31)}$$

C.4 Interpolating $M = 3$ Landscape for Larger Systems

Upon considering systems with more microhabitats, a natural question arises regarding how to construct a landscape for arbitrary M that most closely resembles that in the $M = 3$ system. If our $M = 3$ selection landscape is given by $(s_0, s_0 + \Delta S, s_0)$, as illustrated in Figure C.1, which features do we want preserved across systems of all sizes? We want to look at only symmetric landscapes with odd M that have a single peak or valley in the selection pressure landscape at the center microhabitat. To further constrain the landscapes, we want all systems to have a selection pressure of s_0 for the extrema microhabitats and $s_0 + \Delta S$ for the center microhabitat. We also want to conserve $\langle s \rangle$ across all sizes.

For convenience, we will define

$$\delta s_j \equiv s(x_j) - s(x_{j-1}), \quad \text{(C.32)}$$

for $j \in [1, \lfloor M/2 \rfloor]$. This measures the difference in selection pressure between adjacent micro-habitats. Specifying the set $\{\delta s_j\}$ for $j \in [1, \lfloor M/2 \rfloor]$ is equivalent to fully specifying the selection

landscape, given the assumption of symmetry about $x_{\lfloor M/2 \rfloor}$ as well as $s(x_0) = S_0$. For a specific example, we can solve for the values for $M = 5$. Without loss of generality, we take $s_0 = 0$. Then our two constraints are given by

$$\delta s_1 + \delta s_2 = \Delta S, \tag{C.33}$$

$$\frac{1}{5}(3\delta s_1 + \delta s_2) = \frac{\Delta S}{3}. \tag{C.34}$$

These two equations can be solved to find

$$\delta s_1 = \frac{\Delta S}{3}, \tag{C.35}$$

$$\delta s_2 = \frac{2\Delta S}{3}. \tag{C.36}$$

For $M > 5$, however, we have $\lfloor M/2 \rfloor$ unknown quantities but only two equations for the constraints that we have defined up to this point. Among the infinite landscapes that can satisfy these constraints, we choose to minimize the variance of the variance of the differences $\{\delta_j\}$. That is, we minimize

$$\mathrm{Var}(\{\delta s_j\}) = \sum_{j=1}^{\lfloor M/2 \rfloor} (\delta s_j - \langle \delta s \rangle)^2, \tag{C.37}$$

to ensure that the selection pressure landscape is smooth for arbitrary M. We also want to write our constraints for general M. Our two constraints are

$$\sum_{j=1}^{\lfloor M/2 \rfloor} \delta s_j = \Delta S, \tag{C.38}$$

$$\frac{1}{M} \sum_{j=1}^{\lfloor M/2 \rfloor} (M - 2)\delta s_j = \Delta S/3, \tag{C.39}$$

where the mean selection pressure is written as

$$\langle s \rangle(\{\delta s_j\}) = \frac{1}{M} \left[(M - 2)\delta s_1 + (M - 4)\delta s_2 + \ldots + \delta s_{\lfloor M/2 \rfloor} \right] \tag{C.40}$$

$$= \frac{1}{M} \sum_{j=1}^{\lfloor M/2 \rfloor} (M - 2)\delta s_j. \tag{C.41}$$

We can now minimize the variance and use Lagrange multipliers to enforce these two general

constraints. We form the Lagrangian

$$\mathcal{L}(\{\delta s_j\}; \lambda_1, \lambda_2) = \sum_{j=1}^{\lfloor M/2 \rfloor} (\delta s_j - \langle \delta s \rangle)^2 + \lambda_1 \left(\sum_{k=1}^{\lfloor M/2 \rfloor} \delta s_k - \Delta S \right)$$

$$+ \lambda_2 \left(\frac{1}{M} \sum_{j=1}^{\lfloor M/2 \rfloor} (M-2)\delta s_j - \frac{\Delta S}{3} \right). \quad \text{(C.42)}$$

We then solve $\nabla \mathcal{L} = 0$ to calculate the set $\{\delta s_j\}$ to specify a selection landscape for arbitrary M. The relevant partial derivatives are given by

$$\partial_{\delta s_j} \mathcal{L} = 2\delta s_j - \frac{\Delta S}{M^2}(3M+1) + \lambda_1 + \lambda_2 \left(1 - \frac{2j}{M} \right), \quad \text{(C.43)}$$

$$\partial_{\lambda_1} \mathcal{L} = \sum_{j=1}^{\lfloor M/2 \rfloor} \delta s_j - \Delta S, \quad \text{(C.44)}$$

$$\partial_{\lambda_2} \mathcal{L} = \frac{1}{M} \sum_{j=1}^{\lfloor M/2 \rfloor} (M-2j)\delta s_j - \frac{\Delta S}{3}. \quad \text{(C.45)}$$

Setting the partial derivative with respect to δs_j equal to zero, we can write

$$\delta s_j = \frac{3M+1}{2M^2} \Delta S - \frac{\lambda_1}{2} - \frac{\lambda_2}{2} \left(1 - \frac{2j}{M} \right). \quad \text{(C.46)}$$

We then use this expression with our two constraints to form equations involving only λ_1 and λ_2. Using this expression for δs_j, our first constraint can be written as

$$\sum_{j=1}^{\lfloor M/2 \rfloor} \left[\frac{3M+1}{2M^2} \Delta S - \frac{\lambda_1}{2} - \frac{\lambda_2}{2} \left(1 - \frac{2j}{M} \right) \right] = \Delta S \quad \text{(C.47)}$$

Upon making use of the result

$$\sum_{j=1}^{\lfloor M/2 \rfloor} j = \frac{M^2 - 1}{8}, \quad \text{(C.48)}$$

one can write

$$\lambda_1 = -\frac{M(M-1)^2 \lambda_2 + 2\Delta S(M+1)^2}{2M^2(M-1)}. \quad \text{(C.49)}$$

119

Now using the second constraint, we have

$$\frac{1}{M}\sum_{j=1}^{\lfloor M/2\rfloor}(M-2j)\left[\frac{3M+1}{2M^2}\Delta S-\frac{\lambda_1}{2}-\frac{\lambda_2}{2}\left(1-\frac{2j}{M}\right)\right]=\frac{\Delta S}{3}. \tag{C.50}$$

Using the result

$$\sum_{j=1}^{\lfloor M/2\rfloor}j^2=\frac{1}{24}(M-1)M(M+1), \tag{C.51}$$

we can write

$$\lambda_1=\frac{-2M^2(M-2)(M-1)\lambda_2+(3+M(3+M(M-15)))\Delta S}{3M^2(M-1)^2}. \tag{C.52}$$

These two equations for λ_1 can be solved to find

$$\lambda_1=\frac{M(-5M^2+M-3)-1}{M^2(M^2-1)}\Delta S, \tag{C.53}$$

$$\lambda_2=\frac{8M}{M^2-1}\Delta S. \tag{C.54}$$

These two equations can now be substituted into our expression for δs_j. After simplification, we find

$$\delta s_j=\frac{8j}{M^2-1}\Delta S. \tag{C.55}$$

This assignment allows us to specify a selection pressure landscape for arbitrary M that is the analogue of the $M=3$ landscape. The individual differences $\{\delta s_j\}$ were useful for this calculation but we want to work with the selection pressure value for a given microhabitat. Using our derived expression for δs_j, one can find

$$s(x_j)=s_0+\frac{4j(j+1)}{M^2-1}\Delta S. \tag{C.56}$$

We illustrate the resulting selection pressure landscapes with $M=7$ and $M=9$ in Figure C.2 and Figure C.3, respectively.
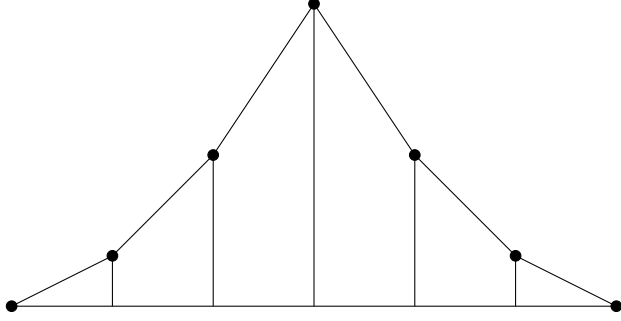
Figure C.2: The minimal variance $M = 7$ interpolated selection pressure landscape.
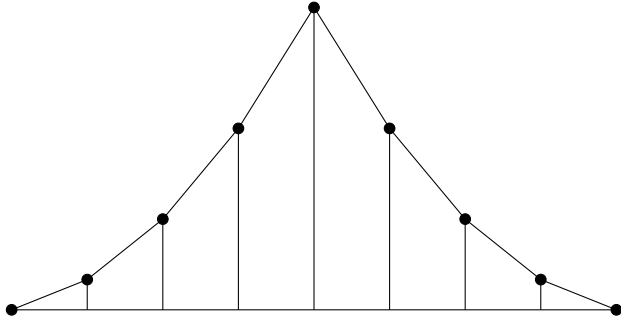
Figure C.3: The minimal variance interpolated selection landscape with $M = 9$.

C.5 Fixation Time Scaling with M

C.5.1 Minimum of M Exponential Processes

When we are in the regime $\mu \ll \beta$, the time to fixation is dominated by the time required for a *de novo* mutation to arise and fixate in the first microhabitat. In this regime, we can approximate

$$\tau_f \simeq \tau_{\min}(\{\}), \tag{C.57}$$

$$= \frac{1}{\sum_{i=0}^{M-1} \lambda_i(0)}, \tag{C.58}$$

which is just the expected minimum of M independent exponential rates with rates λ_i. Because the required for the first microhabitat is so much larger than the time required for the other microhabitats to reach fixation given a single fixated microhabitat, the migration occurs on a fast enough timescale that our expression is independent of β.

Using this result, we can understand how the fixation times for a system with arbitrary M

compare to those obtained with the traditional $M = 3$ system. We look at the ratio

$$\mathcal{T}_{\min}^M \equiv \frac{\tau_{\min}^M}{\tau_{\min}^{M=3}}, \tag{C.59}$$

which quantifies the effect of M on the fixation times in the limit when fixation can be considered the minimum of the M independent exponential processes. Using our exponential rates

$$\lambda_i(0) = N\mu \frac{s(x_i)}{1 - (1 - s(x_i))^N}, \tag{C.60}$$

we can write this ratio as

$$\mathcal{T}_{\min}^M = \frac{\sum_{i=0}^2 \frac{s_0(x_i)}{1-(1-s_0(x_i))^N}}{\sum_{i=0}^{M-1} \frac{s(x_i)}{1-(1-s(x_i))^N}}, \tag{C.61}$$

if we denote the $M = 3$ selection landscape as $(s_0(x_0), s_0(x_1), s_0(x_2))$ and the selection landscape in the arbitrary M system as $s(x_i)$ for $i \in [0, M - 1]$. This expression can be simplified if we use the knowledge of the selection pressure landscape for arbitrary M given that of the $M = 3$ system. We enforce

$$s(x_0) = s_0(x_0) \qquad s(x_{\lfloor M/2 \rfloor}) = s_0(x_1), \tag{C.62}$$

allowing us to write

$$\mathcal{T}_{\min}^M = \frac{\sum_{i=0}^2 \frac{s_0(x_i)}{1-(1-s_0(x_i))^N}}{\sum_{i=0}^2 \frac{s_0(x_i)}{1-(1-s_0(x_i))^N} + \sum_{i=1, \neq \lfloor M/2 \rfloor}^{M-2} \frac{s(x_i)}{1-(1-s(x_i))^N}}, \tag{C.63}$$

Using our expression for $\tau_{\min}^{M=3}$, we identify

$$\sum_{i=0}^2 \frac{s_0(x_i)}{1 - (1 - s_0(x_i))^N} = \left(N\mu\tau_{\min}^{M=3}\right)^{-1}. \tag{C.64}$$

We also simplify the remaining sum using the symmetry of the landscape so that our simplified ratio is given by

$$\mathcal{T}_{\min}^M = \frac{\left(N\mu\tau_{\min}^{M=3}\right)^{-1}}{\left(N\mu\tau_{\min}^{M=3}\right)^{-1} + 2\sum_{i=1}^{\lfloor M/2 \rfloor-1} \frac{s(x_i)}{1-(1-s(x_i))^N}}, \tag{C.65}$$

$$= \left(1 + 2\xi_M N\mu\tau_{\min}^{M=3}\right)^{-1}, \tag{C.66}$$

122

if we define

$$\xi_M \equiv \sum_{i=1}^{\lfloor M/2 \rfloor - 1} \frac{s(x_i)}{1 - (1 - s(x_i))^N}. \tag{C.67}$$

We can further simplify this term using our derivation of the selection pressure landscape for a system with arbitrary M. Using this knowledge, our expression becomes

$$\xi_M = \sum_{j=1}^{(M-3)/2} \frac{s_0(x_0) + \frac{4j(j+1)}{M^2-1}(s_0(x_1) - s_0(x_0))}{1 - \left(1 - s_0(x_0) - \frac{4j(j+1)}{M^2-1}(s_0(x_1) - s_0(x_0))\right)^N}, \tag{C.68}$$

which cannot be further simplified in general, although the sum is easy to evaluate numerically.

The validity of this scaling factor can be tested using our Monte Carlo estimation. The fixation times for arbitrary landscapes can be estimated using the Monte Carlo algorithm, and these values can be compared to what we obtain using the scaling factor, since we can write

$$\tau_{\min}^M = \mathcal{T}_{\min}^M \tau_{\min}^{M=3}. \tag{C.69}$$

That is, we calculate the fixation times in the $M = 3$ landscape and numerically evaluate the scaling factor $\mathcal{T}_{\min}^M$ using the calculated $\tau_{\min}^{M=3}$ value. The fixation times obtained in this manner can then be compared to the values obtained by directly Monte Carlo estimating fixation times with the arbitrary M selection pressure landscape. These two estimates are calculated for different selection pressure landscapes with $M = 11$ assuming $\mu = 10^{-8}$, $\beta = 0.1$, $\langle s \rangle = 0.2$, and $N = 25$ in Figure C.4.

Although the expression for $\mathcal{T}_{\min}^M$ cannot be further simplified in general, we can simplify it considerably if we look at a very simple landscape. If we look at the homogeneous landscape, for instance, we have $s(x_i) = \langle s \rangle \ \forall \ i$, so that

$$\xi_M = \sum_{j=1}^{(M-3)/2} \frac{\langle s \rangle}{1 - (1 - \langle s \rangle)^N}, \tag{C.70}$$

$$= \frac{M - 3}{2} \frac{\langle s \rangle}{1 - (1 - \langle s \rangle)^N}. \tag{C.71}$$

We can also simplify our expression for the fixation time with the $M = 3$ system:

$$\tau_{\min}^{M=3} = \left(3N\mu \frac{\langle s \rangle}{1 - (1 - \langle s \rangle)^N}\right)^{-1}. \tag{C.72}$$

Figure C.4: The estimated fixation time using the analytical scaling factor $\mathcal{T}_{\min}^M$ match what is obtained using Monte Carlo estimation with $\mu = 10^{-8}$, $\beta = 0.1$, $\langle s \rangle = 0.2$, and $N = 25$ with $M = 11$.

Using these two expressions, our scaling factor becomes

$$\mathcal{T}_{\min}^M = 3M^{-1},\tag{C.73}$$

after some algebra. This result is independent of not only β but also μ, $\langle s \rangle$, and N.

C.5.2 Maximum of M Exponential Processes

We can also look at our system when the fixation time can be approximated as the maximum of M independent exponential processes. This approximation is valid in the regime $\beta \ll \mu$. In this regime, migration contributes an insignificant flux of mutants, allowing us to again approximate the exponential rates as

$$\lambda_i = N\mu \frac{s(x_i)}{1 - (1 - s(x_i))^N},\tag{C.74}$$

even though we are in the opposite parameter regime as the earlier case. To form the scaling factor in this parameter regime, we need to calculate the expected maximum of M independent exponential processes. If we have M exponential random variables X_i for $i \in [1, M]$, we can calculate the expectation value of $X = \max\{X_1, \ldots, X_M\}$. We first form the cumulative distribution function

of X:

$$CDF_X(x) = P(\max\{X_1, \ldots, X_M\} < x), \tag{C.75}$$

$$= P(X_1 < x)P(X_2 < x) \cdots P(X_M < x), \tag{C.76}$$

using independence. Now this is expressed as the product of the cumulative distribution functions for single exponential processes, so we can write

$$CDF_X(x) = \left(1 - e^{\lambda_1 x}\right)\left(1 - e^{-\lambda_2 x}\right) \cdots \left(1 - e^{-\lambda_M x}\right), \tag{C.77}$$

allowing the expectation value to be calculated according to

$$E_M(x) = \int_0^\infty \left[1 - \prod_{k=1}^{M} \left(1 - e^{-\lambda_k x}\right)\right] dx. \tag{C.78}$$

This integral can be expressed as

$$E_M(x) = \sum_{k=1}^{M} (-1)^{k+1} \Sigma_M^{(k)}, \tag{C.79}$$

where

$$\Sigma_M^{(k)} \equiv \frac{1}{k!} \sum_{i_1=1}^{M} \sum_{i_2 \neq i_1} \sum_{i_3 \neq i_1, i_2} \cdots \sum_{i_M \notin \{i_1, \ldots, i_{M-1}\}} \frac{1}{\sum_{j=1}^{M} \lambda_j}. \tag{C.80}$$

This is clunky to write but expresses a simple idea. If we were to write out the expected maximum of three independent exponential processes, for instance, we have

$$E_3(\tau_f) = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} - \frac{1}{\lambda_1 + \lambda_2} - \frac{1}{\lambda_1 + \lambda_3} - \frac{1}{\lambda_2 + \lambda_3} + \frac{1}{\lambda_1 + \lambda_2 + \lambda_3}, \tag{C.81}$$

which is not dissimilar to the expression for the generalized recursion relation that we derived above. Due to how unwieldy this expectation value is, we cannot further the scaling factor for the system in the limit beyond

$$\mathcal{T}_{\max}^M \equiv \frac{\tau_{\max}^M}{\tau_{\max}^{M=3}}, \tag{C.82}$$

$$= \frac{\sum_{k=1}^{M} (-1)^{k+1} \Sigma_M^{(k)}}{\sum_{k=1}^{3} (-1)^{k+1} \Sigma_3^{(k)}}. \tag{C.83}$$
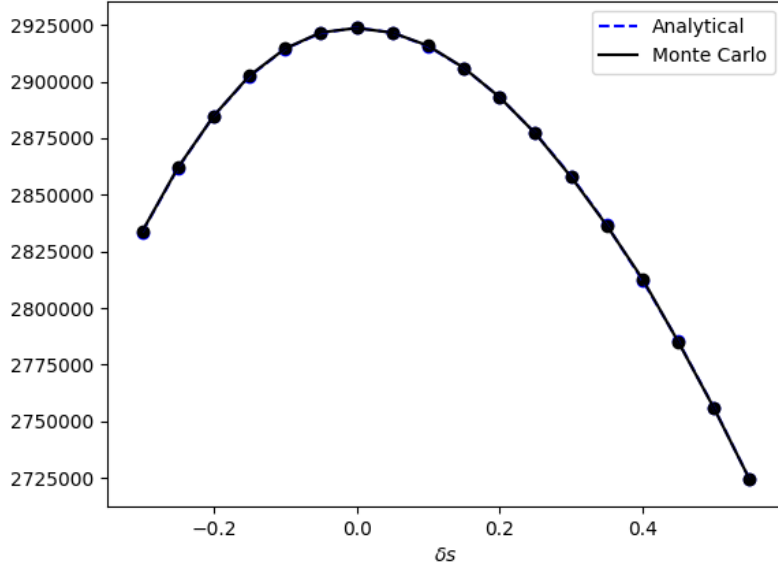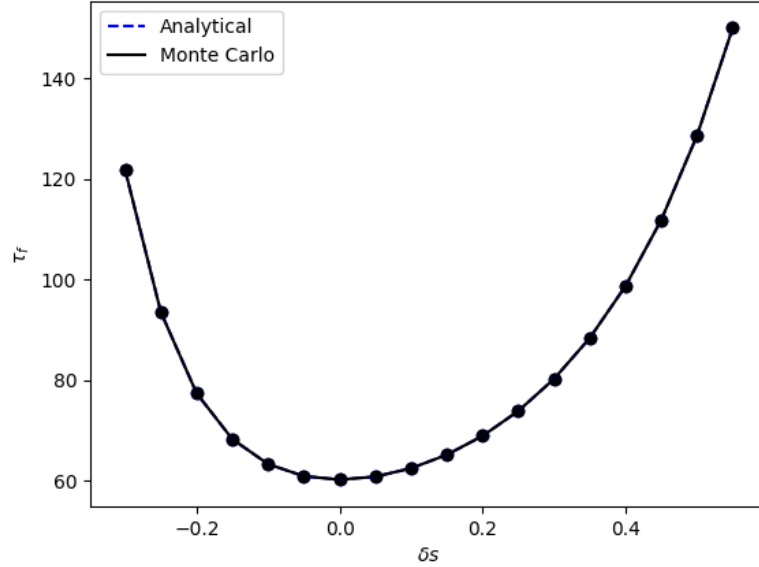
Figure C.5: The estimated fixation time using the analytical scaling factor $\mathcal{T}^{M}_{\max}$ match what is obtained using Monte Carlo estimation for a system with $M = 11$ and $\mu = 0.01$, $\beta = 10^{-8}$, $\langle s \rangle = 0.2$, and $N = 25$.

This scaling factor can again be calculated numerically. However, due to the combinatoric number of terms, it is difficult to evaluate for very large M. We evaluate the validity of the bound by again computing $\tau^{M}_{\max} = \mathcal{T}^{M}_{\max} \tau^{M=3}_{\max}$ and comparing the result to what we obtain by Monte Carlo estimating the fixation times. This comparison performed in Figure C.5 for $\mu = 0.01$, $\beta = 10^{-8}$, $\langle s \rangle = 0.2$, and $N = 25$ with $M = 11$ shows good agreement between the two calculations.

Although the scaling factor $\mathcal{T}^{M}_{\max}$ is very complicated in general, it becomes rather simple if we look at the homogeneous landscape. With insignificant migration and uniform selection values, all of the microhabitats have the same exponential rate given by

$$\lambda = N\mu \frac{\langle s \rangle}{1 - (1 - \langle s \rangle)^{N}}.$$
(C.84)

In our expression for $\Sigma^{(k)}_{M}$, all of the terms in the sum are then identical. By counting the number of terms that appear in the sum, we see

$$(\Sigma)^{(k)} = \frac{M!}{k!(M - k)!} \frac{1}{k\lambda},$$
(C.85)

126

so that our scaling factor is given by

$$\mathcal{T}_{\text{max}}^M = \frac{\sum_{k=1}^M (-1)^{k+1} \frac{M!}{k!(M-k)!} \frac{1}{k\lambda}}{\sum_{k=1}^3 (-1)^{k+1} \frac{3!}{k!(3-k)!} \frac{1}{k\lambda}}, \tag{C.86}$$

$$= \frac{M!}{3!} \frac{\sum_{k=1}^M \frac{(-1)^{k+1}}{k!(M-k)!k\lambda}}{\sum_{k=1}^3 \frac{(-1)^{k+1}}{k!(3-k)!k\lambda}}. \tag{C.87}$$

This equation can be simplified by identifying

$$\sum_{k=1}^M \frac{(-1)^{k+1}}{k!(M-k)!k} = \frac{H_M}{M!}, \tag{C.88}$$

where H_M is the Mth harmonic number. Thus, our scaling factor is given by

$$\mathcal{T}_{\text{max}}^M = \frac{H_M}{H_3}, \tag{C.89}$$

$$= \frac{6}{11} H_M, \tag{C.90}$$

which is independent of all system parameters besides M.

APPENDIX D

Kinetic Monte Carlo Simulation

D.1 Kinetic Monte Carlo Details

The kinetic Monte Carlo algorithm is quite straightforward for our system since the underlying atomic transitions are very simple. At each timestep

1. One of the N_{total} cells is randomly chosen for the next event. The microhabitat and genotype of the chosen cell are denoted by i and g, respectively.

2. The atomic rates are calculated according to $R_{\text{migration}} = \beta$, $R_{\text{growth}} = \phi_g(c, T_i)(1 - N_i/K_i)$, and $R_{\text{death}} = d$, which dictate the rates at which migration, growth, and death occur, respectively.

3. A uniform random number $u \in [0, R_{\text{max}}]$ is drawn, where R_{max} is chosen to be larger than $R_{\text{migration}} + R_{\text{growth}} + R_{\text{death}}$ to allow for the possibility of the state remaining unchanged.

4. The state of the system is updated according to the value of u. Upon replicating, a second random number r is drawn, such that a mutation occurs if $r < \mu$.

5. The simulation time is increased by an amount $\Delta t = 1/(N_{\text{total}} R_{\text{max}})$.

for migration rate β, mutation rate μ, carrying capacity K_i, growth rate $\phi_g(c, T_i)$, and death rate d. Note that the simulation requires a large amount of random number generation, which we perform using `GSL` for computational efficiency. In addition, we require a large number of independent simulations to arrive at meaningful estimates of the dynamics occurring during evolution. We use `OpenMP` to parallelize our implementation to distribute independent simulations across multiple processors.

Although the simulation loop is very straightforward, there are a number of considerations important for selecting the parameter values of interest. Here we briefly consider the impact of several of the parameters. First, we investigate μ, since this is a parameter common to the toy models of

evolution. Our results have only assumed asymmetric mutations, so that genotype g can mutate exclusively to genotype $g + 1$. We performed some simulations with symmetric mutations, so that $g \rightarrow g \pm 1$. The main difference we observed is that symmetric mutations can significantly slow the emergence of resistance. With asymmetric mutations, fixation is an absorbing state that will be reached in finite time. It is possible to construct landscapes in which there is a stationary state other than fixation with symmetric mutations for specific choices of MIC. This is similar to what was observed in the toy model of evolution upon performing a system size expansion of the master equation (Section 3.2.4. To reduce the computational burden, we have focused on the case with asymmetric mutations. Additionally, it is worthwhile to think about the number of mutants that arise per cell generation. If we make the simplifying assumption that each microhabitat has a large enough ϕ to support the carrying capacity K, we can expect approximately $\mu K L$ mutants per generation. This number, rather than simply μ, best characterizes the rarity of *de novo* mutations in our system. We can perform a similar estimate for the number of migration events per generation. However, there is a fundamental difference between mutation and migration in the atomic ruleset of the kinetic Monte Carlo algorithm, in that mutation can only occur during replication. This distinction makes approximating the number of migration events for a general landscape challenging, as the landscape of each particular genotype influences the number of timesteps required to replace a full generation of cells.

Finally, A related parameter is the death rate d, which determines the degree of population turnover. If d is very small, each microhabitat would very rapidly reach its carrying capacity K, at which point very little reproduction, and thus evolution, could occur. Alternatively, if d were very large, the number of cells would be far from the carrying capacity K. We tried a few different values, and $d = 0.1$ seemed to avoid either of these two pitfalls.

Note that we chose the same standard values for R_{max} and d as the authors of [32]. This was by design, as we wanted to build upon this work by studying the system with the added constraint of conserving $\langle \phi \rangle$. This added constraint ensures that any effect of heterogeneity across different landscapes is not the result of variable ϕ confounding the fixation times.

D.2   General Form of ϕ

If we assume that we have a monotonic temperature, chosen to be increasing without loss of generality, we can replace the temperature T with a spatial variable x that is monotonically increasing. With this substitution, our ϕ_m is given by

$$\phi_m(x) = \xi(x)\left[1 - \frac{c^2}{\varphi(x)^2 4^{2(m-1)}}\right] \tag{D.1}$$

The maxima occur at

$$\frac{d}{dx}\phi_m(x) = 0 \tag{D.2}$$

$$0 = \xi'(x)\left[1 - \frac{c^2}{\varphi^2(x)4^{2(m-1)}}\right] + \xi(x)\left(\frac{-c^2}{4^{2(m-1)}}\right)\frac{-2\varphi'(x)}{\varphi^3(x)} \tag{D.3}$$

$$\frac{d\xi}{\xi} = -2\left[\frac{\varphi'(x)}{\frac{4^{2(m-1)}}{c^2}\varphi^3(x) - \varphi(x)}\right]dx \tag{D.4}$$

$$\log\xi = -2\int_{x_i}^{x_f}\left[\frac{\varphi'(x)}{\frac{4^{2(m-1)}}{c^2}\varphi^3(x) - \varphi(x)}\right]dx \tag{D.5}$$

Note that we have implicitly assumed that at the extrema $\varphi(x^*) \neq 0$. Letting $\alpha \equiv 4^{2(m-1)}/c^2$ to simply notation, we can analytically take this integral to get

$$\log\xi = -2\left[-\frac{1}{2}\log(1 - \alpha\varphi^2(x)) - \log\varphi(x)\right]\Bigg|_{x_i}^{x_f} \tag{D.6}$$

$$= -2\left(\left[-\frac{1}{2}\log(1 - \alpha\varphi^2(x)) - \log\varphi(x_f)\right]\right) -$$
$$\qquad\qquad - 2\left(\left[-\frac{1}{2}\log(1 - \alpha\varphi^2(x_i)) - \log\varphi(x_i)\right]\right) \tag{D.7}$$

$$= \log\left[\frac{1 - \alpha\varphi^2(x_f)}{1 - \alpha\varphi^2(x_i)}\left(\frac{\varphi(x_f)}{\varphi(x_i)}\right)^2\right] \tag{D.8}$$

so that we easily identify the local maxima at

$$\xi(x^*) = \left(\frac{\varphi(x_f)}{\varphi(x_i)}\right)^2\frac{1 - \alpha\varphi^2(x_f)}{1 - \alpha\varphi^2(x_i)} \tag{D.9}$$

This result is remarkable in that the critical points x^* are fully determined by the endpoints evaluated with $\varphi(x)$ and not on the form of the coefficient $\varphi(x)$.

Let's now turn to the biological constraints on the system. We are assuming that

1. $\varphi(x)$ has one local maximum.

2. $\xi(x)$ has one local maximum.

We can use these constraints to further limit the form that can be assumed by $\xi(x^*)$.

The first condition prevents $\varphi(x)$ from having a bowl-shape, with both of the local maximum at the edge of the domain. Interestingly, if we relax our assumption and allow $\varphi(x)$ to have any number of local maxima that aren't at the endpoints, we don't actually relax the form that can be taken by $\xi(x^*)$.

The second constraint is more useful for restricting the class of functions we are interested in. For any set of endpoints, ξ^* evaluates to a single number. For any form of $\varphi(x)$, then, we can only have local extrema at

$$x^* = \xi^{-1}\left[\left(\frac{\varphi(x_f)}{\varphi(x_i)}\right)^2 \frac{1 - \alpha\varphi^2(x_f)}{1 - \alpha\varphi^2(x_i)}\right] \tag{D.10}$$

How many such extrema, particularly maxima, can exist given our constraints? Clearly, we could have a symmetric form of $\varphi(x)$, which would evaluate to the same argument of ξ^{-1}. However, since $\varphi(x) \in (0, 1]$, this does not lead to anything physical. If we imagine that $\xi(x)$ were symmetric, though, ξ would map two different values of x^* to the same value $\xi(x^*)$, leading to two extrema for a single set of endpoints.

If we take $\xi(x)$ to be flat for some $x \in [x_i, x_f]$, we would have an infinite degeneracy and therefore an infinite number of local extrema. This is intuitive in the trivial case where both $\varphi(x)$ and $\xi(x)$ have no spatial variation, since ϕ_m would be flat and every point would be a local extrema.

So if we assume that $\xi(x)$ is never flat in our domain, we know we can have two maxima if

$$\xi : x_1^*, x_2^* \to \xi(x^*) \tag{D.11}$$

Is it possible to have more? It is not. Since ξ can only have a single local maxima, it is not possible for it to map more than two values of x to the same range.

Therefore, the limited constraints on the system are sufficient to guarantee that the generalized form of ϕ_m has at most two peaks. It is obviously possible to have a single peak if $\varphi(x_i) = \varphi(x_f)$ as well. Since we have ruled out the case of the trivial forms of $\varphi(x)$ and $\xi(x)$, we must have at least one peak in our ϕ_m for any form of the coefficients that are never flat in the domain.

BIBLIOGRAPHY

[1] L. J. Piddock, "The crisis of no new antibioticswhat is the way forward?," *The Lancet infectious diseases*, vol. 12, no. 3, pp. 249–253, 2012.

[2] S. Sengupta, M. K. Chattopadhyay, and H.-P. Grossart, "The multifaceted roles of antibiotics and antibiotic resistance in nature," *Frontiers in microbiology*, vol. 4, p. 47, 2013.

[3] S. B. Levy and B. Marshall, "Antibacterial resistance worldwide: causes, challenges and responses," *Nature medicine*, vol. 10, no. 12s, p. S122, 2004.

[4] C. Holohan, S. Van Schaeybroeck, D. B. Longley, and P. G. Johnston, "Cancer drug resistance: an evolving paradigm," *Nature reviews. Cancer*, vol. 13, no. 10, p. 714, 2013.

[5] M. M. Gottesman, "Mechanisms of cancer drug resistance," *Annual review of medicine*, vol. 53, no. 1, pp. 615–627, 2002.

[6] J. Davies and D. Davies, "Origins and evolution of antibiotic resistance," *Microbiology and Molecular Biology Reviews*, vol. 74, no. 3, pp. 417–433, 2010.

[7] H. C. Neu, "The crisis in antibiotic resistance," *Science*, vol. 257, no. 5073, pp. 1064–1073, 1992.

[8] C. Årdal, M. Balasegaram, R. Laxminarayan, D. McAdams, K. Outterson, J. H. Rex, and N. Sumpradit, "Antibiotic development–economic, regulatory and societal challenges," *Nature Reviews Microbiology*, vol. 18, no. 5, pp. 267–274, 2020.

[9] R. Nelson, "Antibiotic development pipeline runs dry," *The Lancet*, vol. 362, no. 9397, pp. 1726–1727, 2003.

[10] K. Bush, P. Courvalin, G. Dantas, J. Davies, B. Eisenstein, P. Huovinen, G. A. Jacoby, R. Kishony, B. N. Kreiswirth, E. Kutter, *et al.*, "Tackling antibiotic resistance," *Nature Reviews Microbiology*, vol. 9, no. 12, pp. 894–896, 2011.

[11] R. Laxminarayan, A. Duse, C. Wattal, A. Zaidi, H. Wertheim, N. Sumpradit, E. Vlieghe, G. Hara, I. Gould, H. Goossens, C. Greko, A. So, M. Bigdeli, G. Tomson, W. Woodhouse, E. Ombaka, A. Peralta, F. Qamar, F. Mir, S. Kariuki, Z. Bhutta, A. Coates, R. Bergstrom, G. Wright, E. Brown, and O. Cars, "Antibiotic resistance-the need for global solutions," *The Lancet Infectious Diseases*, vol. 13, pp. 1057–1098, Dec. 2013.

[12] J. M. Blair, M. A. Webber, A. J. Baylay, D. O. Ogbolu, and L. J. Piddock, "Molecular mechanisms of antibiotic resistance," *Nature reviews microbiology*, vol. 13, no. 1, pp. 42–51, 2015.

[13] L. Fernández and R. E. Hancock, "Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance," *Clinical microbiology reviews*, vol. 25, no. 4, pp. 661–681, 2012.

[14] H. Nikaido, "Multidrug resistance in bacteria," *Annual review of biochemistry*, vol. 78, pp. 119–146, 2009.

[15] G. D. Wright, "Molecular mechanisms of antibiotic resistance," *Chemical communications*, vol. 47, no. 14, pp. 4055–4061, 2011.

[16] D. I. Andersson, N. Q. Balaban, F. Baquero, P. Courvalin, P. Glaser, U. Gophna, R. Kishony, S. Molin, and T. Tnjum, "Antibiotic resistance: turning evolutionary principles into clinical reality," *FEMS Microbiology Reviews*, vol. 44, pp. 171–188, 01 2020.

[17] D. Schillaci, V. Span, B. Parrino, A. Carbone, A. Montalbano, P. Barraja, P. Diana, G. Cirrincione, and S. Cascioferro, "Pharmaceutical approaches to target antibiotic resistance mechanisms," *Journal of Medicinal Chemistry*, vol. 60, no. 20, pp. 8268–8297, 2017. PMID: 28594170.

[18] M. Baym, L. K. Stone, and R. Kishony, "Multidrug evolutionary strategies to reverse antibiotic resistance," *Science*, vol. 351, no. 6268, p. aad3292, 2016.

[19] I. Bozic, J. G. Reiter, B. Allen, T. Antal, K. Chatterjee, P. Shah, Y. S. Moon, A. Yaqubie, N. Kelly, D. T. Le, *et al.*, "Evolutionary dynamics of cancer in response to targeted combination therapy," *Elife*, vol. 2, p. e00747, 2013.

[20] Y. Iwasa, M. A. Nowak, and F. Michor, "Evolution of resistance during clonal expansion," *Genetics*, vol. 172, no. 4, pp. 2557–2566, 2006.

[21] E. Toprak, A. Veres, J.-B. Michel, R. Chait, D. L. Hartl, and R. Kishony, "Evolutionary paths to antibiotic resistance under dynamically sustained drug selection," *Nature genetics*, vol. 44, no. 1, pp. 101–105, 2012.

[22] E. A. Yurtsev, H. X. Chao, M. S. Datta, T. Artemova, and J. Gore, "Bacterial cheating drives the population dynamics of cooperative antibiotic resistance plasmids," *Molecular systems biology*, vol. 9, no. 1, p. 683, 2013.

[23] R. Chait, A. Craney, and R. Kishony, "Antibiotic interactions that select against resistance," *Nature*, vol. 446, no. 7136, p. 668, 2007.

[24] J. P. Torella, R. Chait, and R. Kishony, "Optimal drug synergy in antimicrobial treatments," *PLoS computational biology*, vol. 6, no. 6, p. e1000796, 2010.

[25] E. Hansen, R. J. Woods, and A. F. Read, "How to use a chemotherapeutic agent when resistance to it threatens the patient," *PLoS biology*, vol. 15, no. 2, p. e2001110, 2017.

[26] A. Fischer, I. Vázquez-García, and V. Mustonen, "The value of monitoring to control evolving populations," *Proceedings of the National Academy of Sciences*, vol. 112, no. 4, pp. 1007–1012, 2015.

[27] H. Chung, T. D. Lieberman, S. O. Vargas, K. B. Flett, A. J. McAdam, G. P. Priebe, and R. Kishony, "Global and local selection acting on the pathogen stenotrophomonas maltophilia in the human lung," *Nature communications*, vol. 8, p. 14078, 2017.

[28] D. Nichol, P. Jeavons, A. G. Fletcher, R. A. Bonomo, P. K. Maini, J. L. Paul, R. A. Gatenby, A. R. Anderson, and J. G. Scott, "Steering evolution with sequential therapy to prevent the emergence of bacterial antibiotic resistance," *PLoS computational biology*, vol. 11, no. 9, p. e1004493, 2015.

[29] M. Gralka, D. Fusco, S. Martis, and O. Hallatschek, "Convection shapes the trade-off between antibiotic efficacy and the selection for resistance in spatial gradients," *Physical Biology*, vol. 14, no. 4, p. 045011, 2017.

[30] E. Pavlovskiĭ, *Natural Nidality of Transmissible Diseases; With Special Reference to the Landscape Epidemiology of Zooanthroponoses*. Urbana : University of Illinois Press, 1966.

[31] F. Fu, M. A. Nowak, and S. Bonhoeffer, "Spatial heterogeneity in drug concentrations can facilitate the emergence of resistance to cancer therapy," *PLoS Comput Biol*, vol. 11, no. 3, p. e1004142, 2015.

[32] P. Greulich, B. Waclaw, and R. J. Allen, "Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells," *Physical Review Letters*, vol. 109, no. 8, p. 088101, 2012.

[33] R. Hermsen, J. B. Deris, and T. Hwa, "On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient," *Proceedings of the National Academy of Sciences*, vol. 109, no. 27, pp. 10775–10780, 2012.

[34] T. B. Kepler and A. S. Perelson, "Drug concentration heterogeneity facilitates the evolution of drug resistance," *Proceedings of the National Academy of Sciences*, vol. 95, no. 20, pp. 11514–11519, 1998.

[35] S. Moreno-Gamez, A. L. Hill, D. I. Rosenbloom, D. A. Petrov, M. A. Nowak, and P. S. Pennings, "Imperfect drug penetration leads to spatial monotherapy and rapid evolution of multidrug resistance," *Proceedings of the National Academy of Sciences*, vol. 112, no. 22, pp. E2874–E2883, 2015.

[36] Q. Zhang, G. Lambert, D. Liao, H. Kim, K. Robin, C.-k. Tung, N. Pourmand, and R. H. Austin, "Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments," *Science*, vol. 333, no. 6050, pp. 1764–1767, 2011.

[37] R. Hermsen and T. Hwa, "Sources and sinks: a stochastic model of evolution in heterogeneous environments," *Physical review letters*, vol. 105, no. 24, p. 248104, 2010.

[38] M. Baym, T. D. Lieberman, E. D. Kelsic, R. Chait, R. Gross, I. Yelin, and R. Kishony, "Spatiotemporal microbial evolution on antibiotic landscapes," *Science*, vol. 353, no. 6304, pp. 1147–1151, 2016.

[39] A. I. Minchinton and I. F. Tannock, "Drug penetration in solid tumours," *Nature Reviews Cancer*, vol. 6, no. 8, pp. 583–592, 2006.

[40] M. Müller, A. dela Peña, and H. Derendorf, "Issues in pharmacokinetics and pharmacodynamics of anti-infective agents: Distribution in tissue," *Antimicrobial Agents and Chemotherapy*, vol. 48, no. 5, pp. 1441–1453, 2004.

[41] P. S. Stewart, "Diffusion in biofilms," *Journal of Bacteriology*, vol. 185, no. 5, pp. 1485–1491, 2003.

[42] O. Trdan, C. M. Galmarini, K. Patel, and I. F. Tannock, "Drug Resistance and the Solid Tumor Microenvironment," *JNCI: Journal of the National Cancer Institute*, vol. 99, pp. 1441–1454, 10 2007.

[43] S. T. Pickett and M. L. Cadenasso, "Landscape ecology: spatial heterogeneity in ecological systems," *Science*, vol. 269, no. 5222, p. 331, 1995.

[44] M. C. Whitlock and R. Gomulkiewicz, "Probability of fixation in a heterogeneous environment," *Genetics*, vol. 171, no. 3, pp. 1407–1417, 2005.

[45] L. A. Real and R. Biek, "Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes," *Journal of the Royal Society Interface*, vol. 4, no. 16, pp. 935–948, 2007.

[46] V. Manem, M. Kohandel, N. Komarova, and S. Sivaloganathan, "Spatial invasion dynamics on random and unstructured meshes: Implications for heterogeneous tumor populations," *Journal of Theoretical Biology*, vol. 349, pp. 66 – 73, 2014.

[47] V. S. K. Manem, K. Kaveh, M. Kohandel, and S. Sivaloganathan, "Modeling invasion dynamics with spatial random-fitness due to micro-environment," *PLOS ONE*, vol. 10, pp. 1–20, 10 2015.

[48] T. Maruyama, "On the fixation probability of mutant genes in a subdivided population," *Genetics Research*, vol. 15, no. 2, pp. 221–225, 1970.

[49] R. S. Singer, M. P. Ward, and G. Maldonado, "Can landscape ecology untangle the complexity of antibiotic resistance?," *Nature Reviews Microbiology*, vol. 4, no. 12, pp. 943–952, 2006.

[50] R. Allen and B. Waclaw, "Antibiotic resistance: a physicist's view," *Physical Biology*, vol. 13, p. 045001, aug 2016.

[51] D. Fisher, M. Lssig, and B. Shraiman, "Evolutionary dynamics and statistical physics," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2013, p. N01001, feb 2013.

[52] K. Wood, "Microbial ecology: Complex bacterial communities reduce selection for antibiotic resistance," *Current Biology*, vol. 29, no. 21, pp. R1143–R1145, 2019.

[53] M. Khasin, E. Khain, and L. M. Sander, "Fast migration and emergent population dynamics," *Physical review letters*, vol. 109, no. 24, p. 248102, 2012.

[54] M. Khasin, B. Meerson, E. Khain, and L. M. Sander, "Minimizing the population extinction risk by migration," *Physical review letters*, vol. 109, no. 13, p. 138104, 2012.

[55] J. N. Waddell, L. M. Sander, and C. R. Doering, "Demographic stochasticity versus spatial variation in the competition between fast and slow dispersers," *Theoretical population biology*, vol. 77, no. 4, pp. 279–286, 2010.

[56] Y. T. Lin, H. Kim, and C. R. Doering, "Demographic stochasticity and evolution of dispersion i. spatially homogeneous environments," *Journal of Mathematical Biology*, vol. 70, no. 3, pp. 647–678, 2015.

[57] Y. T. Lin, H. Kim, and C. R. Doering, "Demographic stochasticity and evolution of dispersion ii: Spatially inhomogeneous environments," *Journal of Mathematical Biology*, vol. 70, pp. 679–707, Feb 2015.

[58] G. W. Constable and A. J. McKane, "Fast-mode elimination in stochastic metapopulation models," *Physical Review E*, vol. 89, no. 3, p. 032141, 2014.

[59] G. W. Constable and A. J. McKane, "Population genetics on islands connected by an arbitrary network: An analytic approach," *Journal of theoretical biology*, vol. 358, pp. 149–165, 2014.

[60] G. W. Constable and A. J. McKane, "Stationary solutions for metapopulation moran models with mutation and selection," *Physical Review E*, vol. 91, no. 3, p. 032711, 2015.

[61] E. Peterson and P. Kaur, "Antibiotic resistance mechanisms in bacteria: Relationships between resistance determinants of antibiotic producers, environmental bacteria, and clinical pathogens," *Frontiers in Microbiology*, vol. 9, p. 2928, 2018.

[62] S. Okasha, "Emergence, hierarchy and top-down causation in evolutionary biology," *Proceedings of the National Academy of Sciences*, vol. 2, no. 1, pp. 49–54, 2011.

[63] J. Ross and A. P. Arkin, "Complex systems: From chemistry to systems biology," *Proceedings of the National Academy of Sciences*, vol. 106, no. 16, pp. 6433–6434, 2009.

[64] C. S. Gokhale, Y. Iwasa, M. A. Nowak, and A. Traulsen, "The pace of evolution across fitness valleys," *Journal of Theoretical Biology*, vol. 259, no. 3, pp. 613–620, 2009.

[65] Y. Iwasa, F. Michor, and M. A. Nowak, "Stochastic tunnels in evolutionary dynamics," *Genetics*, vol. 166, no. 3, pp. 1571–1579, 2004.

[66] D. M. Weinreich, L. Chao, and P. Phillips, "Rapid evolutionary escape by large populations from local fitness peaks is likely in nature," *Evolution*, vol. 59, no. 6, pp. 1175–1182, 2005.

[67] D. B. Weissman, M. M. Desai, D. S. Fisher, and M. W. Feldman, "The rate at which asexual populations cross fitness valleys," *Theoretical population biology*, vol. 75, no. 4, pp. 286–300, 2009.

[68] S. Farhang-Sardroodi, A. Darooneh, M. Nikbakht, N. Komarova, and M. Kohandel, "The effect of spatial randomness on the average fixation time of mutants.," *PLoS computational biology*, vol. 13, no. 11, p. e1005864, 2017.

[69] A. Kaznatcheev, J. G. Scott, and D. Basanta, "Edge effects in game-theoretic dynamics of spatially structured tumours," *Journal of The Royal Society Interface*, vol. 12, no. 108, p. 20150154, 2015.

[70] B. Allen, G. Lippner, Y.-T. Chen, B. Fotouhi, N. Momeni, S.-T. Yau, and M. A. Nowak, "Evolutionary dynamics on any population structure," *Nature*, vol. 544, no. 7649, pp. 227–230, 2017.

[71] K. S. Korolev, M. Avlund, O. Hallatschek, and D. R. Nelson, "Genetic demixing and evolution in linear stepping stone models," *Reviews of modern physics*, vol. 82, no. 2, p. 1691, 2010.

[72] C. Li, "The statistical processes of evolutionary theory," *American journal of human genetics*, vol. 14, no. 4, p. 438, 1962.

[73] C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry and the natural sciences*, vol. 13 of *Springer Series in Synergetics*. Berlin: Springer-Verlag, third ed., 2004.

[74] N. V. Kampen, *Stochastic processes in physics and chemistry*. North Holland, 2007.

[75] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *Journal of Mathematics and Physics*, vol. 6, no. 1–4, pp. 164–189, 1927.

[76] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of eckart-young decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[77] R. A. Harshman, "Foundations of the parafac procedure: Models and conditions for an" explanatory" multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, no. 84, pp. 1–84, 1970.

[78] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[79] G. Tomasi and R. Bro, "Parafac and missing values," *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 2, pp. 163–180, 2005.

[80] D. B. Clewell, M. S. Gilmore, Y. Ike, and N. Shankar, *Enterococci: from commensals to leading causes of drug resistant infection*. Massachusetts Eye and Ear Infirmary, 2014.

[81] M. G. De Jong and K. B. Wood, "Tuning spatial profiles of selection pressure to modulate the evolution of drug resistance," *Phys. Rev. Lett.*, vol. 120, p. 238102, Jun 2018.

[82] S. Gagneux, C. D. Long, P. M. Small, T. Van, G. K. Schoolnik, and B. J. M. Bohannan, "The competitive cost of antibiotic resistance in mycobacterium tuberculosis," *Science*, vol. 312, no. 5782, pp. 1944–1946, 2006.

[83] A. H. Melnyk, A. Wong, and R. Kassen, "The fitness costs of antibiotic resistance mutations," *Evolutionary Applications*, vol. 8, no. 3, pp. 273–283, 2015.

[84] L. Rosso, J. Lobry, and J. Flandrois, "An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model," *Journal of Theoretical Biology*, vol. 162, no. 4, pp. 447–463, 1993.

[85] P. B. Price and T. Sowers, "Temperature dependence of metabolic rates for microbial growth, maintenance, and survival," *Proceedings of the National Academy of Sciences*, vol. 101, no. 13, pp. 4631–4636, 2004.

[86] P. A. Mackowiak, M. Marling-Cason, and R. L. Cohen, "Effects of Temperature on Antimicrobial Susceptibility of Bacteria," *The Journal of Infectious Diseases*, vol. 145, pp. 550–553, 04 1982.

[87] R. R. Regoes, C. Wiuff, R. M. Zappala, K. N. Garner, F. Baquero, and B. R. Levin, "Pharmacodynamic functions: a multiparameter approach to the design of antibiotic treatment regimens," *Antimicrobial Agents and Chemotherapy*, vol. 48, no. 10, pp. 3670–3676, 2004.

[88] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The journal of physical chemistry*, vol. 81, no. 25, pp. 2340–2361, 1977.

[89] J. Monod, "The growth of bacterial cultures," *Annual Review of Microbiology*, vol. 3, no. 1, pp. 371–394, 1949.

[90] M. J. Brauer, C. Huttenhower, E. M. Airoldi, R. Rosenstein, J. C. Matese, D. Gresham, V. M. Boer, O. G. Troyanskaya, and D. Botstein, "Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast," *Molecular Biology of the Cell*, vol. 19, no. 1, pp. 352–367, 2008. PMID: 17959824.

[91] K. Valgepea, K. Adamberg, A. Seiman, and R. Vilu, "Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins," *Mol. BioSyst.*, vol. 9, pp. 2344–2358, 2013.

[92] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, "Interdependence of cell growth and gene expression: Origins and consequences," *Science*, vol. 330, no. 6007, pp. 1099–1102, 2010.

[93] J. Monod, "La technique de culture continue, theorie et applications," *Ann. Inst. Pasteur*, vol. 79, no. 4, pp. 390–410, 1950.

[94] A. J. Saldanha, M. J. Brauer, and D. Botstein, "Nutritional homeostasis in batch and steady-state culture of yeast," *Molecular Biology of the Cell*, vol. 15, no. 9, pp. 4089–4104, 2004. PMID: 15240820.

[95] C. N. Takahashi, A. W. Miller, F. Ekness, M. J. Dunham, and E. Klavins, "A low cost, customizable turbidostat for use in synthetic circuit characterization," *ACS Synthetic Biology*, vol. 4, no. 1, pp. 32–38, 2015. PMID: 25036317.

[96] Z. J. Heins, C. P. Mancuso, S. Kiriakov, B. G. Wong, C. J. Bashor, and A. S. Khalil, "Designing automated, high-throughput, continuous cell growth experiments using evolver," *Journal of Visualized Experiments*, vol. 147, 2019.

[97] J. A. Ryan *et al.*, *Understanding and managing cell culture contamination*. Corning Incorporated, 1994.

[98] W. Arber, "Genetic variation: molecular mechanisms and impact on microbial evolution," *FEMS Microbiology Reviews*, vol. 24, pp. 1–7, 01 2000.

[99] A. Hilfinger and J. Paulsson, "Separating intrinsic from extrinsic fluctuations in dynamic biological systems," *Proceedings of the National Academy of Sciences*, vol. 108, no. 29, pp. 12167–12172, 2011.

[100] L. S. Tsimring, "Noise in biology," *Reports on Progress in Physics*, vol. 77, p. 026601, jan 2014.

[101] N. Eling, M. D. Morgan, and J. C. Marioni, "Genetic variation: molecular mechanisms and impact on microbial evolution," *Nature Reviews Genetics*, vol. 20, pp. 536–548, 2019.

[102] R. Selander and B. Levin, "Genetic diversity and structure in escherichia coli populations," *Science*, vol. 210, no. 4469, pp. 545–547, 1980.

[103] S. M. McBride, V. A. Fischetti, D. J. LeBlanc, R. C. Moellering, Jr., and M. S. Gilmore, "Genetic diversity among enterococcus faecalis," *PLOS ONE*, vol. 2, pp. 1–22, 07 2007.

[104] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79 – 86, 1951.

[105] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[106] C. J. Geyer, "Practical markov chain monte carlo," *Statistical science*, pp. 473–483, 1992.

[107] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[108] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré, "Markov chain monte carlo without likelihoods," *Proceedings of the National Academy of Sciences*, vol. 100, no. 26, pp. 15324–15328, 2003.