Statistical Fluctuations in Evolutionary and Population Dynamics

Citation

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493257

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current/terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Statistical Fluctuations in Evolutionary
and Population Dynamics

A dissertation presented
by
Thiparat Chotibut
to
The Department of Physics
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Physics
Harvard University
Cambridge, Massachusetts

March 2016
Statistical Fluctuations in Evolutionary and Population Dynamics

Abstract

In this thesis, we study collective phenomena that arise from microscopic fluctuations at the individual level of two different living populations. First, we study evolutionary dynamics of two-species competitions in a well-mixed environment subject to population size fluctuations. We demonstrate a mechanism for neutral evolution such that population size fluctuations favor a fixation of one species over the other. An effective evolutionary dynamics for fluctuation-induced selection is derived. We then investigate strong mutualism, in a limit where a varying population size can strongly influence the evolutionary dynamics. We determine fixation probabilities as well as mean fixation times taking into account the population size degree of freedom. The results elucidate the interplay between population size fluctuations and evolutionary dynamics in well-mixed systems. Second, we investigate single species marine population subject to a constant flow field and quenched random spatially fluctuating growth rates. We show that the non-equilibrium steady-state population density of a generalized Fisher-Kolmogorov-Petrovsky-Piscounov (FKPP) equation develops a flow-driven striation pattern. The striations are highly asymmetric with a longitudinal correlation length that diverges linearly with the flow speed and a transverse correlation length that approaches a finite velocity-independent value. The findings suggest that, although the growth disorder can be spatially uncorrelated, correlated population structures with striations emerge naturally at sufficiently strong advection.
## Contents

Title Page i

Abstract iii

Table of Contents iv

Citations to Previously Published Work vi

Acknowledgements vii

1 Introduction 1

1.1 Evolutionary Dynamics in a Well-mixed Environment with Fluctuating Population Sizes 2

1.2 Spatial Population Dynamics in Disordered and Advective Environments 15

2 Population Genetics with Fluctuating Population Sizes 23

2.1 Competitive Lotka-Volterra Model 24

2.2 Mean Field Description 25

2.3 Growth of Population Size and Mapping to the Replicator Dynamics 27

2.4 Stochastic Description 31

2.5 Replicator Dynamics with Genetic Drift and Population Size Fluctuations 35

2.6 Fluctuation-induced Selection in Quasi-Neutral Evolution 39

2.7 A Naive Approximation 42

2.8 Effective Evolutionary Dynamics from Adiabatic Elimination of the Fast Population Size Variable 44

2.9 Dimensional Reduction: the Fixation Probability and the Mean Fixation Time 47
Citations to Previously Published Work

The manuscript based on Chapter 2 is in preparation for submission.

Sections 2.1-2.5 and Chapter 3 appear in their entirety in

**Evolutionary dynamics with fluctuating population sizes and strong mutualism**
T. Chotibut, and D. R. Nelson

Chapter 4 will appear in its entirety in

**Striated populations in disordered environments with advection**
T. Chotibut, D. R. Nelson, and S. Succi
Acknowledgements

This academic journey would never be completed without remarkable people whom I have met and learned from. First and foremost, I would like to express my deepest gratitude to my advisor, Prof. David R. Nelson, for his inspiring guidance, support, and encouragement throughout my graduate study. His vast and deep knowledge combined with his penetrating insights never cease to amaze me, and regularly point me to exciting avenues of research. I thank him for providing abundant resources and autonomy for me to explore new ideas, and for his meticulous comments on my work, from which I have learned to grow as a scientist. I truly treasure our discussions, which have been and will always be a source of my inspiration for doing science. His incisive thinking and his ability to spot subtle connections between disparate subjects continue to serve as a gold standard for me to look up to in my future endeavors. It is a great honor to be mentored by David.

I am also very grateful for the other members of my thesis committee, Prof. Subir Sachdev and Prof. Michael P. Brenner. The course on many-body quantum theory and the reading course on quantum phase transition with Prof. Subir Sachdev provide useful theoretical tools for studying collective phenomena, some of which appears in Chapter 4 of this thesis. Subir’s style of combining deep physical arguments with skillful execution of mathematics has convinced me that condensed matter physics is enjoyable and elegant. I also thank Subir for helping me smoothly transition into a different research direction as I pondered my future research career in Thailand during the second year of graduate school. I have learned the essence of applied mathematics and mathematical modeling from Prof. Michael P. Brenner’s APM201 and group meetings. Michael’s unlimited enthusiasm for diverse scientific problems broadens my perspectives on science, and his career advice during the last
year of my graduate study has been very helpful.

I would also like to thank teachers and colleagues at Harvard. Prof. Sauro Succi kindly and patiently taught me the Lattice Boltzmann method, and sparked my interests in numerical computation as well as in the history of computer programming in the punched card era. Many discussions with Sauro help refine Chapter 4 of this thesis. Matched asymptotic methods I learned from serving as a teaching fellow for Prof. Lakshminarayanan Mahadevan’s beautiful course on physical mathematics also help me develop analytical techniques for Chapter 3. The soft matter group provides a very inspiring and stimulating work environment; in particular, I thank Prof. Ariel Amir, Prof. Andrej Kosmrlj, Prof. Chris Rycroft, Severine Atis, Amir Azadi, Daniel Beller, Thomas Fai, Max Lavrentovich, Andrew Maranthan, Wolfram Mobius, Melanie Mueller, Jayson Paulose, Bryan Weinstein, and David Zwicker for many stimulating discussions. Salute to Max Lavrentovich for being a fantastic office-mate for many years. The FAS Division of Science Research Computing Group at Harvard University provides crucial computational resources for this thesis. Jennifer Bastin, Lisa Cacciabaudo, and Carol Davis, your assistance in the administrative work of the department has been a big support.

I am also thankful for teachers and friends I met elsewhere. My interests in theoretical physics are triggered by beautiful lectures on introductory quantum mechanics by Prof. Paul Fendley at the University of Virginia. I particularly thank Prof. Seunghun Lee and Prof. Blaine E. Norum at the University of Virginia who offered research opportunities during my undergraduate years and introduced me to research in condensed matter and high-energy physics, respectively. Seunghun successfully persuaded me to pursue a graduate study in condensed matter physics, and I am very fortunate that I took his advice. I also thank the hospitality of the Center for Models of Life at the Niels Bohr Institute, the University of Copenhagen,
where a portion of research in this thesis was conducted. I would like to acknowledge the Development and Promotion of Science and Technology Talents Project from the Royal Thai government for sponsoring the scholarship throughout my years in the United States. Also, a big cheer to the Thai communities in Charlottesville, Cambridge, and Boston for making the United States feel like home.

Finally, I am indebted to my parents, for giving me the freedom to explore the world, for their unlimited love and support throughout my education, and for instilling in me persistence and resilience. Although their only child has received aortic valve replacement surgery and needs special medical attention, they audaciously let him follow his dreams half a world away from home. Without all the medical miracles, and the trusts from my parents, this thesis would never exist. I wholeheartedly dedicate this thesis to my parents.
To my mom, my dad, and the Chotibut family
Chapter 1
Introduction

The study of noise and fluctuations is a central theme of statistical physics. When a large number of interacting constituents, individually subject to microscopic fluctuations, form a macroscopic system, the collective effect of fluctuations can be dramatic, and fascinating macroscopic collective phenomena can arise. Living populations, too, are macroscopic systems consisting of interacting self-replicating individuals, each can be subject to noise and fluctuations. Not surprisingly, fascinating macroscopic phenomena also arise in living populations. For instance, in microbial populations, fluctuations in individual growth rate can result in a dramatic effect on the genetic compositions of the populations [1–3]. Figure 1.1 demonstrates such an effect whereby stochastic growth rates (genetic drift) cause genetic segregation and eventually destroy the coexistence of the two neutral species during a spatial range expansion.

Although advances in experimental evolution and microbiology reveal the details and the complexity of living populations, standard theoretical models are often oversimplified and neglect many salient features of living populations. One of the most important yet often overlooked feature is temporal and spatial variations in the population sizes. This thesis investigates the effects of these fluctuations in two different living populations: well-mixed two-species populations with temporal fluctuations in the population size, and single-species marine microbial populations with spatial fluctuations in the population size. Chapter 2 and 3 investigate the former, and Chapter 4 studies the latter. We show that incorporating these salient
Figure 1.1: A range expansion experiment on an agar plate with a razor blade inoculation (adapted from Ref. [1].) Beginning with an equal fraction of well-mixed green and red fluorescently labeled neutral E.coli, fluctuations in the individual reproduction rates (genetic drift) lead to genetic segregation, and a single species can eventually take over the populations by chance. Chapter 2 and 3 of this thesis concerns the effect of genetic drift, selection, as well as temporal fluctuations in the population sizes on evolutionary dynamics in a well-mixed environment.

features result in interesting macroscopic phenomena in living populations. The rest of this introductory chapter provides motivations and necessary backgrounds for the main chapters.

1.1 Evolutionary Dynamics in a Well-mixed Environment with Fluctuating Population Sizes

Evolutionary processes are ubiquitous in living systems. Individuals reproduce and pass on its genes to their descendants. Mutation, which is the error in gene duplication, may arise and give rise to new mutants. Depending on environmental conditions and interactions among individuals in living populations, the fitter individuals tend to reproduce faster by natural selection. However, fitter individuals may also give birth to fewer descendants by random chances; these statistical fluctuations in the reproduction rates are termed genetic drift in population genetics. Selection, mutation, and genetic drift influence the genetic compositions of
the populations. Quantitative understanding of these ideas, which are the focus of evolutionary dynamics and population genetics, may lead to practical biomedical applications, for example, drug developments and treatments of cancer. The rest of this section introduces the standard concept of selection and genetic drift in well-mixed populations of fixed size, which are fundamental to the generalization to fluctuating population sizes in Chapter 2 and 3. The discussion of mutation can be found in textbooks of population genetics and evolutionary dynamics [4–6].

In a well-mixed environment such as in a chemostat or in a shaken test tube, relative genotypic or phenotypic abundances (frequency) of interacting microbes can be measured over time. Many mathematical descriptions are proposed to describe the dynamics in the frequency space due to natural selection. Among the widely studied frameworks for the well-mixed populations is evolutionary game theory [7, 8], which was originally proposed to study animal behaviors [9]. Using dynamical system approaches, evolutionary game theory extends the notion of classical game theory in economics [10, 11] to describe how the relative abundance of strategies in the populations changes, assuming a payoff from microscopic interaction. In the context of biological evolution, strategies are translated into species or phenotypes, and the payoffs are interpreted as reproductive fitness.

The central equation of evolutionary game theory that prescribes deterministic time evolution of relative frequency $f_i(t)$ of species $i$ in a well-mixed and infinitely large populations is the replicator dynamics [12, 13]:

$$\frac{df_i}{dt} = [w_i(f) - \bar{w}(f)] f_i. \quad (1.1)$$

In Eq.(1.1), $w_i(f)$ is the frequency-dependent fitness of species $i$, and $\bar{w}(f) =$
Chapter 1: Introduction

\[ \sum_j f_j w_j(f) \] is the mean fitness of all interacting species \([6, 7, 14]\). Replicator dynamics encapsulates frequency-dependent natural selection; a fitter species flourishes and a weaker species succumbs to evolutionary forces. The fitness of species \(i\) is often defined as a constant background plus the total payoff from interactions, assumed to be linear in \(\{f_i(t)\}\), \(w_i(f) = 1 + \sum_j a_{ij} f_j\), where \(a_{ij}\) is a phenomenological payoff matrix characterizing interactions with species \(j\).

For two interacting species, which is typical in competition experiments \([15, 16]\) and is the focus of this thesis, the frequency \(f(t)\) of species 1 fully specifies the state of the populations, since the frequency of species 2 is just \(1 - f(t)\). In this case, the replicator dynamics determine the time evolution of \(f(t)\) from Eq. (1.1):

\[
\frac{df}{dt} = \left[ \alpha_1 (1 - f) - \alpha_2 f \right] (1 - f) f \equiv v_E(f), \tag{1.2}
\]

where \(\alpha_1 = a_{12} - a_{22}\) and \(\alpha_2 = a_{21} - a_{11}\). A rich variety of competition scenarios emerge from this simple description of evolutionary games. Depending on the payoff differences \(\alpha_1\) and \(\alpha_2\), Eq. (1.2) exhibits 5 qualitatively different competition scenarios, schematically sketched in Fig. 1.2.

For positive \(\alpha\)’s (first quadrant of Fig. 1.2), a stable fixed point corresponding to a species coexistence appears at \(f^* = \alpha_1 / (\alpha_1 + \alpha_2)\), lying between the unstable fixed points \(f = 0\) and \(f = 1\). This scenario is commonly referred to as “snowdrift game” in game theory or mutualism in the context of evolution \([3, 14, 17]\). For negative \(\alpha\)’s (third quadrant of Fig. 1.2), the fixed point \(f^*\) becomes unstable while the fixed points 0 and 1 are stable. This bistability situation is known by “coordination game” in game theory or antagonism in our context. When \(\alpha\)’s have opposite signs (second quadrant and fourth quadrant of Fig. 1.2), scenarios in game
Figure 1.2: Four competition scenarios in the two-species replicator dynamics represented by the four quadrants. The fifth scenario, neutral evolution ($\alpha_1 = \alpha_2 = 0$), in which every point $f \in [0,1]$ is a fixed point, arises at the origin.

theory are either called “harmony” or “prisoner’s dilemma”. The fixed point $f^*$ becomes physically irrelevant. The only two physical fixed points are $f = 0$ and $f = 1$; only one of them is stable. For $\alpha_1 > 0$ and $\alpha_2 < 0$, the fixed point $f = 1$ is stable and species 1 dominates. For $\alpha_2 > 0$ and $\alpha_1 < 0$, the fixed point $f = 0$ is stable and species 2 dominates. Lastly, when $\alpha_1 = \alpha_2 = 0$, every point is a fixed point. We shall refer to this scenario as neutral evolution, representing situations when the two interacting species are neutral variants of each other.

In addition to the deterministic force of natural selection, genetic compositions can also be influenced by fluctuating evolutionary forces from random birth and death events in finite populations. Random fluctuations in the reproductive rate, or genetic drift, is one of the central concepts in population genetics embodied in

5
the work of Fisher [18] and Wright [19]. In the Wright-Fisher model consisting of
$N$ neutral haploid individuals, the parent generation are randomly chosen to give
birth to the daughter generation. The generations are assumed non-overlapping,
and the random sampling process repeats until the daughter generation contains
$N$ individuals. This random sampling with replacement imitates random birth and
death events of the parental generation. For two neutral species 1 and 2, this process
generates fluctuations in the species frequency as follows: Let $X_t$ denote the number
of species 1 in generation $t$; the conditional probability that $X_{t+1} = n$ given $X_t = m$
is the binomial distribution:

$$P(X_{t+1} = n|X_t = m) = \binom{N}{n} \left( \frac{m}{N} \right)^n \left( 1 - \frac{m}{N} \right)^{N-n}.$$ \hspace{1cm} (1.3)

By the property of the binomial distribution, the mean species fraction $f_t$ remains
unchanged. However, the species frequency $f_t \equiv X_t/N$ fluctuates with the variance
that depends on the population size $N$ and the species frequency of the parent
generation $f$ as

$$\text{Var}(f_{t+1}|f_t = f) = \frac{f(1-f)}{N}.$$ \hspace{1cm} (1.4)

This discrete-time unbiased random walk in the genetic composition of populations
exemplifies genetic drift, whose effect becomes more pronounced at smaller popu-
lation size. Despite being neutral (identical reproduction rate on average), one of
the species can take over the populations (fixation) by chance. Fig. 1.3 illustrates
genetic drift for the Wright-Fisher model.

A variant of the Wright-Fisher model for genetic drift is the Moran model, which
does not assume non-overlapping generations. Although the original model is for-
mulated in discrete time [21, 22], we introduce the continuous time version here as it
is more relevant to statistical physics in the context of the Master equation. Recall
that, in a continuous-time discrete-state Markov process, the time evolution of the probability distribution \( P(n, t) \) for finding the system in a discrete state \( n \) at time \( t \) evolves according to the Master equation [23, 24]:

\[
\partial_t P(n, t) = \sum_{n'} \left[ W(n|n')P(n', t) - W(n'|n)P(n, t) \right],
\]

where \( W(n|n') \) is the transition rate from the configuration \( n' \) to \( n \). The Moran model is a Markov process that specifies the transition rate by a continuous-time sampling with replacement. In a finite population of size \( N \) with \( n \) species 1 and \( N - n \) species 2, two individuals are sampled at a rate \( \mu \); one is chosen to reproduce and the other is chosen to die to ensure the population size is constant. The transition rates by reproduction of species 1 (death of species 2) and by death of species 1
(reproduction of species 2) are thus given by, respectively,

\[ W(n+1|n) = \mu \left( 1 - \frac{n}{N} \right) \left( \frac{n}{N} \right), \quad (1.6) \]

\[ W(n-1|n) = \mu \left( 1 - \frac{n}{N} \right) \left( \frac{n}{N} \right). \quad (1.7) \]

The Master equation describing the dynamics of species 1 in the Moran model reads

\[
\partial_t P(n, t) = [W(n|n + 1)P(n + 1, t) + W(n|n - 1)P(n - 1, t)]
- [W(n + 1|n)P(n, t) + W(n - 1|n)P(n, t)],
\]

where the transition rates are given by Eqs. (1.6) and (1.7).

In the large \( N \) limit, we may promote the species frequency \( f = n/N \) to a continuous variable and approximate the discrete Master equation (1.8) by the Fokker-Planck equation in \( f \). Systematic methods for the Fokker-Planck approximation of the discrete Master equation, which we shall employ in Chapter 2 and 3, are the Kramers-Moyal expansion and the Van-Kampen’s system size expansion \([23, 24]\). Here, it suffices to Taylor expand Eq. (1.8) to \( O(1/N^2) \):

\[
\partial_t P(f, t) = \mu \left( f - \frac{1}{N} \right) \left( 1 - f + \frac{1}{N} \right) \left( 1 - \frac{1}{N} \partial_f + \frac{1}{2N^2} \partial_f^2 \right) P(f, t)
+ \mu \left( f + \frac{1}{N} \right) \left( 1 - f - \frac{1}{N} \right) \left( 1 + \frac{1}{N} \partial_f + \frac{1}{2N^2} \partial_f^2 \right) P(f, t)
- 2\mu f (1 - f) P(f, t) + O(1/N^3).
\]

Upon defining one generation time as \( \tau_g = N\mu^{-1} \), which represents \( N \) sampling
events, the final result is a Fokker-Planck equation for genetic drift:

$$\frac{\partial_t P(f, t)}{1} = \frac{g}{\tau} \frac{\partial^2 f}{\partial f^2} \left[ \frac{D_g(f)}{N} P(f, t) \right], \quad (1.9)$$

where the frequency-dependent diffusion coefficient is

$$\frac{D_g(f)}{N} = \frac{f(1-f)}{N}, \quad (1.10)$$

similar to the variance per generation time of Eq. (1.4) in the Wright-Fisher model. The Fokker-Planck equation (1.9) describes genetic drift as an unbiased random walk in the frequency space, provided the two competing species are neutral.

In finite populations with natural selection, the Moran model must be modified to account for the selection bias. Deterministic selection from evolutionary game theory can be combined with the stochastic process of genetic drift by adjusting the transition rate appropriately. In a fixed population size $N$ with $n$ individuals of species 1 and $N-n$ of species 2, the transition rates by reproduction and death of species 1 are, respectively, [6]

$$W(n+1|n) = \mu \left[ \frac{nw_1}{nw_1 + (N-n)w_2} \right] \left( 1 - \frac{n}{N} \right), \quad (1.11)$$
$$W(n-1|n) = \mu \left[ \frac{(N-n)w_2}{nw_1 + (N-n)w_2} \right] \left( \frac{n}{N} \right), \quad (1.12)$$

where $w_1$ and $w_2$ are the fitness weights of species 1 and 2 determined from the payoff matrix by

$$w_1 = \frac{a_{11}(n) + a_{12}(N-n)}{N}, \quad (1.13)$$
$$w_2 = \frac{a_{21}(n) + a_{22}(N-n)}{N}. \quad (1.14)$$
Chapter 1: Introduction

Upon assuming large population size \( N \gg 1 \), small payoffs \( a_{ij} \ll 1 \), and Taylor expanding the discrete Master equation (1.8) in the species frequency \( f = n/N \) as in the case of the Moran model, the result is a Fokker-Planck equation with both biased motion and genetic drift:

\[
\partial_t P(f, t) = -\frac{1}{\tau_g} \partial_f [v_E(f) P(f, t)] + \frac{1}{\tau_g} \partial_f^2 \left[ \frac{D_g(f)}{N} P(f, t) \right],
\]

(1.15)

where the bias arises from the natural selection term from the replicator dynamics in Eq. (1.2)

\[
v_E(f) = [\alpha_1(1 - f) - \alpha_2(f)](1 - f)f,
\]

(1.16)

where \( \alpha_1 = a_{12} - a_{22} \) and \( \alpha_2 = a_{21} - a_{11} \) (similar to the definition of the replicator dynamics), and the diffusion coefficient is that of the Moran model in Eq. (1.10) [2, 17]. In the case of neutral evolution \( (\alpha_1 = \alpha_2 = 0) \), we recover the unbiased random walk result of Eq. (1.9).

A representation of Eq. (1.15) that reveals the underlying continuous time stochastic dynamics is the stochastic differential equation

\[
\frac{df}{dt} = \frac{1}{\tau_g} v_E(f) + \sqrt{\frac{2D_g(f)}{\tau_g N}} \Gamma(t),
\]

(1.17)

where \( \Gamma(t) \) is the Gaussian white-noise with zero mean \( \langle \Gamma(t) \rangle \) and unit variance \( \langle \Gamma(t) \Gamma(t') \rangle = \delta(t - t') \) [2, 25]. To recover the Fokker-Planck equation (1.15) from Eq. (1.17), the Ito’s interpretation must be employed. In particular, the choice of \( t^* \in [t, t + dt] \) in the standard calculus interpretation of Eq.(1.17) as

\[
f(t + dt) \approx f(t) + \left\{ \frac{1}{\tau_g} v_E[f(t^*)] + \sqrt{\frac{2D_g[f(t^*)]}{\tau_g N}} \Gamma(t^*) \right\} dt
\]

(1.18)
can not be arbitrary, as one takes the limit $dt \to 0$. A different choice of $t^*$ in an infinitesimal interval $[t, t+dt]$ will lead to a different Fokker-Planck equation [25, 26]. To recover population genetics, where the fluctuations in a given generation are entirely determined by the statistics of the preceding generation, we need the Ito’s interpretation where $t^*$ must be $t$ [26]. The Ito’s interpretation also leads to the Ito’s change of variable formula that is important for studying the fluctuations-induced drift phenomena of Chapter 2.

The stochastic differential equation (1.17) implies that once the system reaches either $f = 0$ or $f = 1$, the dynamics completely stop; both the replicator dynamics $v_E(f) = [\alpha_1(1 - f) - \alpha_2(f)](1 - f) f$ and the genetic drift, whose strength is proportional to the diffusion coefficient $D_g(f)/N = f(1 - f)/N$, vanish at these states. Since fluctuations can drive the system into but not away from $f = 0$ and $f = 1$, these are absorbing states. These absorbing states arise from the state-dependent diffusion coefficient (multiplicative noise) of Eq.(1.17) and are common in non-equilibrium statistical dynamics [27–29]. The importance of genetic drift is now clear; genetic drift eventually destroys coexistence even when the two species are neutral variants or interact mutualistically. In these two cases, a series of microscopic fluctuations lead to an irreversible macroscopic change in the population compositions that would otherwise never happen; those who survive are the fortunate ones!

Two important quantities quantify the fate of the surviving species: the fixation probability $u(f)$ and the mean fixation time $\tau(f)$. These are, respectively, the probability that a species of interest takes over the population and the average time required for this to happen, given an initial composition $f$. Throughout this thesis, we shall refer to fixation as the situation when species 1 takes over, which is equivalent to the situation that the system eventually reaches the absorbing state $f = 1$. 

11
This first passage problem is more conveniently studied from the backward time formulation (backward Kolmogorov equation) with a target state in mind, rather than the forward time formulation in the Fokker-Planck equation (forward Kolmogorov equation) [20, 25, 26]. The standard results of $u(f)$ and $\tau(f)$ for evolutionary dynamics in a fixed population size for different selection scenarios are briefly discussed in Chapter 2.

Although replicator dynamics with genetic drift embodied in Eq. (1.17) and its generalizations have received increasing attention for studying the interplay between selection and genetic drift, e.g. the dilemma of cooperation [30], rare fluctuation effects in mutualism [31, 32], the crossover from the mean-field behavior to fluctuations-dominated behavior in quantum game theory [33], as well as competition and cooperation in spatial range expansions [2, 17, 34, 35], this framework enforces a strictly fixed population size $N$, which has several drawbacks. First, it imposes an artificial growth constraint: the birth of one species necessitates the death of the other even when the two species are neutral variants. Furthermore, population size fluctuations away from a preferred carrying capacity often arise in laboratory experiments, such as in evolution experiments by serial transfers [36] depicted in Fig. 1.4, as well as in natural environments. In addition, understanding how effectively compressible oceanic flows affect population genetics of marine organisms such as phytoplankton and cyanobacteria [37–40] requires time-dependent description of local population size, determined by a fluid flow structure. Incorporating spatially dependent population sizes into the evolutionary dynamics of Eq. (1.18) raises important technical and conceptual challenges [41, 42].

In Chapter 2 and 3, we study a two-species competitive Lotka-Volterra model that accounts for natural population growth and encompasses the selection scenarios of the replicator dynamics. Instead of artificially enforcing a strictly fixed population
Chapter 1: Introduction

Figure 1.4: Schematics of evolution experiments by serial transfers (adapted from Ref. [36].) Competing populations grow in fresh media (brown flasks), until the nutrient is near depletion when a portion of nearly saturated populations sampled to regrow in another fresh media. The transfer process repeats, and the dynamics of the genetic frequency can be measured over time. The population size is not constant as often assumed in theoretical population genetics and evolutionary dynamics. In Chapter 2 and 3 of this thesis, we show how population size fluctuations can change the results of the replicator dynamics with genetic drift in a strictly fixed population size.

size $N$, the population size becomes a dynamical variable $N(t)$ and couples to the evolutionary dynamics $f(t)$. The number of individuals of each species stochastically changes by both logistic growth and interspecies competition. These growth and competition events determine the transition rates of the Master equation (1.5) without fixing the population size.

In Chapter 2, we first discuss the limit when long-time dynamics is governed by weak population size fluctuations around a fixed stable equilibrium population size $N_e$. If two competing species reproduce in the dilute limit at an equal rate, $f(t)$ and $N(t)$ approximately decouple near the equilibrium population size $N_e$. In this case, the effective evolutionary dynamics near the equilibrium population size is described by replicator dynamics with genetic drift of Eq. (1.17). Despite population size fluctuations, Moran model results with and without selection are recovered. Pigolotti et al. utilized this limit to extend Eq. (1.17) to study population genetics.
in aquatic environments, where population size also varies in both time and space [41].

We then study the interesting limit where \( f(t) \) and \( N(t) \) are coupled, although \( N(t) \) still fluctuates around a fixed stable equilibrium size \( N_e \). A novel mechanism emerges as a result of the coupling, such that neutral evolution acquires a fluctuation-induced selection term in Eq. (1.17). Species with a reproductive disadvantage in the dilute limit far from the equilibrium population size acquire a selective advantage for competitions at long times near \( N_e \). After adiabatic elimination of the fast population size variable and careful application of the Ito calculus, the effective evolutionary dynamics of quasi-neutral evolution near the equilibrium population size \( N_e \) is determined.

In chapter 3, we study the limit when \( f(t) \) and \( N(t) \) are coupled and competitions take place with a systematically varying population size as opposed to fluctuations around a fixed equilibrium population size. We focus on the strong mutualism scenario, the usual approach, embodied in Eq. (1.17), fails to predict the fixation probability due to a strong coupling between \( f(t) \) and \( N(t) \). The problem can be restated as a far from equilibrium escape problem to absorbing boundaries from an attractive mutualistic fixed point in a two dimensional phase space. The method of matched asymptotic expansions is employed to construct both the fixation probability and the mean fixation time, taking into account both evolutionary and population dynamics. The details of the asymptotic expansions are provided in Appendix B. These results elucidate an intriguing interplay between evolutionary dynamics and population size fluctuations.
1.2 Spatial Population Dynamics in Disordered and Advec-
tive Environments

How growth, competition, and dispersal affect the spatial structure of living pop-
ulations has been a keystone to understanding biodiversity and stability of ecosys-
tems [43–47]. Numerous works demonstrate that simple dispersal mechanisms, such
as diffusive spreading and chemotaxis, when coupled with spatial heterogeneity of
resources, may result in intricate far-from-equilibrium spatial patterning in diverse
populations, including, for example, bacterial patterns on a Petri dish [48–50], soft-
sediment mussel beds [51–53], and vegetation patterns in arid ecosystems [46, 54, 55].
More complicated dispersal mechanisms such as advective transport may also in-
fluence spatial distributions in natural populations. For instance, patchiness and
filamentation of planktonic communities on ocean surfaces can arise from the com-
plex interplay between growth, competition, diffusion, and advection by turbulence
and chaotic oceanic flows [52, 56–58]. Two-dimensional effectively compressible tur-
bulence restricts growth and competition to thin filaments, which can lead to a
dramatic decrease in the global carrying capacity in models of microorganism pop-
ulations [59, 60]. Even in a simple effectively unidirectional flow field, the combined
effect of spatial variations in resources and advection on the spatial structure of
populations is very rich [61–70].

A general framework for systematic exploration of the growth and diffusion of
single-species populations in an advective, spatially varying growth environments is
the reaction-diffusion-advection equation [45, 71]:

\[ \partial_c + \nabla \cdot [v(x, t)c] = \nabla \cdot [D(x, t)\nabla c] + f[c]. \]  

(1.19)
Chapter 1: Introduction

In an ecological context, Eq. (1.19) is a nonlinear partial differential equation for the time evolution of the coarse-grained density \( c \equiv c(x, t) \) of populations consisting of individuals that are transported by advection with the velocity field \( v(x, t) \), diffuse (for example by, say, a run and tumble mechanism in the case of bacteria [72]) with a space and time-dependent diffusivity \( D(x, t) \), and reproduce with a density-dependent nonlinear growth rate \( f[c] \) that depends on a local environment.

When advection is absent, the diffusivity is constant, \( D(x, t) = D \), and the resources necessary for growth are homogeneously distributed, the density of populations that reproduce with the logistic growth rate \( f[c] = ac - bc^2 \) (with \( a \) and \( b \) constant) obey the Fisher-Kolmogorov-Petrovsky-Piscounov (FKPP) equation [73, 74]

\[
\partial_t c = D \nabla^2 c + ac - bc^2. \tag{1.20}
\]

The linear reaction term \( ac \) and the non-linear reaction term \( -bc^2 \) describe exponential growth and death by competition for limited resources in a locally well-mixed environment. Although first studied as a model of advantageous gene spreading in spatial population genetics [74], the FKPP equation is a prototypical reaction-diffusion equation that admits a stable traveling non-linear wave solution describing a propagation of an autocatalytic-reaction front in homogeneous media [45, 73–75].

In population dynamics and ecology, this wave corresponds to a spatial range expansion of growing and competing populations that advance from a region saturated at the local carrying capacity \( K \equiv a/b \) (the stable fixed point of the locally well-mixed logistic growth process \( \dot{c}(t) = ac(t) - bc^2(t) \)), into unoccupied territory. For most boundary and initial conditions, this nonlinear population wave propagates at a typical Fisher velocity [75]

\[
v_F = 2\sqrt{Da}, \tag{1.21}
\]
Figure 1.5: Experiments of bacteriophage T7 populations diffusing and reproducing on disordered E.Coli lawn: the yellow fluorescent labeled region are susceptible E. Coli populations (favourable growth habitats) while the red fluorescent labeled rhombus are resistant E. coli populations (less favourable growth habitats.) Spatially disordered growth environments distort the spreading of viral population front (shaded region) that would otherwise be described by a flat Fisher wave front of the FKPP equation (1.20) (adapted from Ref. [76].)

which, in one dimension, is the minimal possible speed [73]. The traveling population wave eventually saturates the domain at long times; the population density becomes featureless and equal to the local carrying capacity $K$, the stable steady state of Eq. (1.20).

In most natural environments, however, resources necessary for growth are not uniformly distributed. Populations tend to cluster around favourable growth hot spots, and ecological niches emerge [77, 78]. This spatial heterogeneity also influences the dynamics of population front during spatial range expansions, for example, as shown in Fig. 1.5 for viral spreading in disordered growth environments [76]. In addition to heterogeneous resource distribution, natural populations such as marine microbes also experience fluid advection. Interesting phenomena and non-trivial spatial population structure at long times can arise when a constant flow field $\mathbf{v}(x, t) \equiv \mathbf{v}$ advects the population across a spatially quenched growth landscape of the form $f[c] = a(x)c - bc^2$ [61–63, 65], such that the simplest generalization of the
FKPP equation reads

\[ \partial_t c + \mathbf{v} \cdot \nabla c = D \nabla^2 c + a(\mathbf{x})c - bc^2. \]  

(1.22)

Since advection can carry the populations away from favourable growth hot spots and diffusion tends to spread populations out, the density \(c(\mathbf{x}, t)\) deviates from the profile of the local carrying capacity \(K(\mathbf{x}) \equiv a(\mathbf{x})/b\) in regions where \(a(\mathbf{x}) > 0\). For instance, in a landscape with a finite-size favourable growth hot spot surrounded by an unfavourable growth environment, the steady state population structure, instead of localizing near the hot spot, can be driven by advection to delocalize or even become extinct [63]. This localization-extinction transition has been observed in microbial experiments [66, 79]; see, for example, Fig. 1.6 for an experiment on growing bacteria subject to simulated advection and a spatially quenched localized favourable growth hot spot.

On landscapes where multiple growth habitats are present, but with minimal spatial correlations between them, the local growth rate can be regarded as a spatially quenched uncorrelated random variable

\[ a(\mathbf{x}) \equiv \langle a(\mathbf{x}) \rangle = a_0 + \delta a(\mathbf{x}), \]  

(1.23)

where \(a_0\) is the mean background growth rate, \(\delta a(\mathbf{x})\) represents a frozen-in spatial fluctuations with \(\langle \delta a(\mathbf{x}) \rangle = 0\) and \(\langle \delta a(\mathbf{x})\delta a(\mathbf{x}') \rangle \sim \Delta^2 \delta^d(\mathbf{x} - \mathbf{x}')\), \(\langle \cdot \rangle\) represents a spatial average, \(\Delta\) characterizes the strength of growth rate fluctuations, and \(l^d\) measures a typical habitat size. The early time dynamics of populations growing on such disordered and advective environments is well-studied [61, 65]. When the local population size \(c(\mathbf{x}, t)\) is still small at early times, the non-linear competition term
Figure 1.6: Growth of *Bacillus subtilis* colony exposed to ultraviolet radiation with a slowly rotating ultraviolet radiation shield (region confined by the white dashed lines), adapted from Ref. [79]: The region is a fertile growth environment, shielded from harmful UV radiation that penetrates elsewhere. Rotation of the UV-shielded region mimics the effect of advection on the populations growing on a spatially quenched disordered environment of Eq. (1.22) with an advection speed that increases linearly in the radial direction away from the center of rotation (red dot) as $v = \omega r$. At larger radius (strong advection), the populations can not keep up and detach from the localized favourable growth region. Because the average environment is hostile, populations become extinct away from the center of rotation (strong advection), but are localized under the moving shield below a critical radius (below a critical advection speed).

$-bc^2$ can be neglected and the time evolution of the generalized FKPP equation (1.22) is dominated by the linearized growth operator:

$$\mathcal{L}[c(x,t)] \equiv [D \nabla^2 - \mathbf{v} \cdot \nabla + a(x)] [c(x,t)]. \quad (1.24)$$

The right eigenfunctions $\{\phi_n^R(x)\}$ and the associated eigenvalues $\{\lambda_n\}$ of $\mathcal{L}$ determine, respectively, the early-time spatial structures and the associated time evolu-
tion by the usual time evolution operator in the eigenfunction basis expansions:

\[ c(x, t) \approx \exp(\mathcal{L}t)c(x, 0) \]

\[ = \sum_n c_n \exp(\lambda_n t)\phi_n(x), \]  

(1.25)

where \( c_n \) is the projection of \( c(x, 0) \) onto the eigenfunction \( \phi_n^R(x) \), and the approximation sign indicates that the non-linear term in equation (1.22) is neglected. Growth eigenfunctions and the associated random non-Hermitian spectra of the discrete approximation of \( \mathcal{L} \) exhibit remarkable localization-delocalization properties [65, 70, 80–83]. In strongly advective environments in which the spatially averaged growth rate is positive, all growth eigenfunctions are delocalized and the spreading dynamics of populations at the frontier in the direction transverse to the flow is expected to become super-diffusive [65].

While the early-time properties of growth eigenfunctions and growth dynamics in a spatially quenched random growth rate are relatively well understood, less is known about the structure of long-time steady-state populations. Only close to extinction (in an overall hostile growth environment in which the growth rates \( a(x) \) are mostly negative), when the early-time growth eigenfunctions do not significantly overlap, the population structure at long times can be predicted by the early-time growth eigenfunctions; see Fig. 1.7 and its caption for a discussion. However, more generally, significant distortions due to the mode-coupling triggered by the non-linear competition \(-bc^2\) are possible, especially when advection is strong [65].

In Chapter 4, we study steady-state populations described by Eq. (1.22) in a spatially fluctuating and overall fertile growth landscape. Specifically, we study the long-time steady-state population density \( c^*(x) \equiv c(x, t \to \infty) \) satisfying the
Figure 1.7: Comparison between the long-time population structure in one-dimensional disordered hostile environment without advection (average growth rate is negative) from a numerical solution of Eq. (1.22) (left) and from the eigenfunctions \( \{ \phi^R_n(x) \} \) whose eigenvalues have positive real parts (right). The average growth rate \( a_0 \) is negative (hostile) and only four growth hot spots that result from fluctuations \( \delta a(x) \) can support life. Here, only four eigenvalues have positive real parts, and the associated early-time growth eigenfunctions do not significantly overlap. These four eigenfunctions track the long-time steady state of Eq. (1.22) (adapted from [65]). In chapter 4, we study the steady-state populations in fertile environments (\( a_0 \) is positive) with advection, such that most eigenvalues of the operator \( \mathcal{L} \) in Eq.(1.24) have positive real parts and the relevant eigenfunctions have extended spatial components and unavoidably overlap.

Because the non-linear term \( -bc^2 \) introduces interactions between the spatially overlapping modes, the early-time eigenfunction expansion in Eq. (1.25) no longer correlate with the long-time population structure of the non-linear Eq. (1.22). We analyze spatial correlations of the resulting steady states in Chapter 4.

The nonlinear equation,

\[
0 = D \nabla^2 c^*(x) - \mathbf{v} \cdot \nabla c^*(x) + a(x)c^*(x) - bc^2(x),
\]

arising from a weak spatially-quenched random growth landscape described by Eq. (1.23). Here, we set \( a(x) \equiv a_0 + \delta a(x) \), where \( a_0 \) is a constant positive background growth rate, and \( \delta a(x) \) is a weak spatially uncorrelated fluctuations. We show that, although growth rates fluctuations can be spatially uncorrelated, correlated steady-
Figure 1.8: Satellite images of macroscopic filamental structure in phytoplankton blooms. Left: the Atlantic coast of Patagonia (Norman Kuring, Ocean Color Web, NASA). Right: East of Tasmania (Jeff Schmaltz/MODIS Rapid Response Team, NASA/GSFC). Growth, competition for resources, diffusion, advection, and spatially disordered environments are present in these natural phenomena. We show in chapter 4 that the long-time steady state of the generalized FKPP equation (1.22) with spatially quenched uncorrelated fluctuating growth rates $a(x)$ can lead to correlated population structure with striations induced by sufficiently strong recirculating flows.

State population structure with striations reminiscent of the filamental structure of the planktonic community shown in Fig. 1.8, emerge naturally, exemplifying the central theme of this thesis, that microscopic fluctuations can lead to interesting macroscopic phenomena in living populations.
Chapter 2
Population Genetics with Fluctuating Population Sizes

Game theory ideas provide a useful framework for studying evolutionary dynamics in a well-mixed environment. However, as discussed in the introductory chapter, this approach typically enforces a strictly fixed overall population size, deemphasizing natural growth processes. In this chapter, we study a competitive Lotka-Volterra model, with number fluctuations, that accounts for natural population growth and encompasses interaction scenarios typical of evolutionary games. Secs. 2.1-2.4 present the mean-field and stochastic description of the competitive Lotka-Volterra model. The phase portraits of the model and of the replicator dynamics are compared and contrasted. The emphasis is on parameter values such that an attractive line of approximately fixed population size dominates the long-time dynamics. This limit enables us to identify the mapping between the model and the two-species replicator dynamics (1.2).

In Sec. 2.5, we discuss the limit when the replicator dynamics with genetic drift allows independent population size fluctuations. Standard population genetics results such as the fixation probability and the mean fixation time in different selection scenarios are recovered. In Sec. 2.6, we study fluctuation-induced selection phenomena in quasi-neutral evolution. Adiabatic elimination of a fast variable is employed to derive an effective dynamics near an equilibrium population size. The fixation probability and the mean fixation time beginning with an arbitrary initial
population size and an initial frequency are inferred analytically and numerically.

## 2.1 Competitive Lotka-Volterra Model

The competitive Lotka-Volterra model accounts for natural population growth with limited resources; each individual of the same species $S_i$ undergoes a logistic growth process:

$$S_i \xrightarrow{\mu_i} S_i + S_i,$$

$$S_i + S_i \xrightarrow{\lambda_{ii}} S_i,$$

where $\mu_i$ is the reproduction rate of species $i$, and $\lambda_{ii}$ is the rate of intraspecies competition. The combination of (2.1), which describes an exponential growth of population in abundant resources, and (2.2), which dominates when the population size is large, leads to saturation of population size at the carrying capacity $N_i^* = \mu_i/\lambda_{ii}$. Experiments show a logistic growth accurately captures the growth dynamics of a single yeast strain in a well-mixed culture [84].

Interspecies interactions are modeled by additional competition

$$S_i + S_j \xrightarrow{\lambda_{ij}} S_j,$$

where $\lambda_{ij}$ is the rate at which species $j$ wins in the competition for limited resources with species $i$. In general, $\lambda_{ij} \neq \lambda_{ji}$ for $i \neq j$ although $\lambda_{ij}$ and $\lambda_{ji}$ must both be nonnegative in this model. The interaction (2.3) encapsulates situations when one species suffers from the presence of the others, for example, by secretions of toxins or competitions for the same resources. As we will now show, there are 5 generic competition scenarios analogous to replicator dynamics. The population
size, however, is not strictly fixed since the reactions (2.1-2.3) do not conserve the overall population size.

### 2.2 Mean Field Description

In a well-mixed environment with an infinitely large population size, Eqs. (2.1)-(2.3) can be regarded as chemical reactions and determine the mean field dynamics of the number of species $i$, $N_i$, as

\[
\frac{dN_1}{dt} = (\mu_1 - \lambda_{11}N_1 - \lambda_{12}N_2)N_1, \tag{2.4}
\]

\[
\frac{dN_2}{dt} = (\mu_2 - \lambda_{22}N_2 - \lambda_{21}N_1)N_2, \tag{2.5}
\]

where we set the reaction volume to 1. Without interspecies competition, each species $i$ independently grows up and saturates at the carrying capacity $N_i^* = \mu_i/\lambda_{ii}$. Although the carrying capacity of the two species can be different in general, we focus on the case when $N_1^* = N_2^* = N$ for simplicity. By introducing $c_i = N_i/N$, which represents the number of species $i$ relative to its carrying capacity, Eqs. (2.4) and (2.5) can be non-dimensionalized to read

\[
\frac{1}{(1 + s_o)} \frac{dc_1}{d\tilde{t}} = c_1 \left(1 - c_1 - c_2\right) + \beta_1 c_1 c_2, \tag{2.6}
\]

\[
\frac{dc_2}{d\tilde{t}} = c_2 \left(1 - c_1 - c_2\right) + \beta_2 c_1 c_2, \tag{2.7}
\]

where $\tilde{t}$ is the dimensionless time $\mu_2 t$, $s_o$ is the reproductive advantage of species 1 near the origin defined by $1 + s_o \equiv \mu_1/\mu_2$, and the interspecies competitions are absorbed into $\beta_1 \equiv 1 - \left(\frac{\lambda_{12}}{\lambda_{22}}\right)\left(\frac{\mu_2}{\mu_1}\right)$ and $\beta_2 \equiv 1 - \left(\frac{\lambda_{21}}{\lambda_{11}}\right)\left(\frac{\mu_1}{\mu_2}\right)$. Note that the $\{\beta_i\}$ can not exceed unity if $\{\mu_i\}$ and $\{\lambda_{ij}\}$ are positive. Three dimensionless parame-
ters $s_o, \beta_1, \beta_2$ control the phase portraits in the $c_1-c_2$ plane, which always contain at least 3 physically relevant fixed points at $(0,0), (1,0)$, and $(0,1)$, corresponding to the total extinction, the saturation of species 1, and the saturation of species 2, respectively. The fixed point $(0,0)$ is always unstable with the straight heteroclinic trajectories connecting $(0,0)$ to $(1,0)$ and $(0,0)$ to $(0,1)$ describing the logistic growth of a single species in the absence of the other species.

Two dimensionless parameters $\beta_1$ and $\beta_2$ dictate competition scenarios similar to those described by $\alpha_1$ and $\alpha_2$ in the replicator dynamics, provided an initial condition contains non-zero population of both species. However, the overall population size is now allowed to change. These mean field competition scenarios are illustrated in Figs. 2.1, 2.2, and 2.3. If $\beta_1 \beta_2 < 0$, the species $i$ with positive $\beta_i$ dominates. The fixed point corresponding to the saturation of the dominating species is a stable node and the fixed point corresponding to the saturation of the extinct species is a saddle point.

When $\beta_1 \beta_2 > 0$, another dynamically relevant fixed point appears at $c^* = \frac{1}{\beta_1 + \beta_2 - \beta_1 \beta_2}(\beta_1, \beta_2)$. If both $\beta_1$ and $\beta_2$ are negative, we have a bistable situation similar to antagonism. Initial conditions that lie on the basin of attraction of the fixed point $(1,0)$ and $(0,1)$ result in the total domination (i.e., fixation) of species 1 and species 2, respectively. The coexistence fixed point $c^*$ is a saddle point whose stable 1-d manifold consists of the separatrices. Here, coexistence is fragile and only possible for initial conditions lying exactly on the separatrices.

When both $\beta_1$ and $\beta_2$ are positive, stable coexistence emerges at the stable fixed point $c^*$ similar to mutualism. Although we shall refer to this scenario as mutualism to conform to Refs. [17] and [41], we emphasize that interspecies interactions actually arise from underlying competitive interactions. In our case, interspecies interactions reduce the growth rate per capita of both species and restrict $\lambda_{ij} > 0$.
or equivalently $\beta_i < 1$. Stable coexistence can persist despite the competition. The population size at $c^*$, however, reduces to $\frac{\beta_1 + \beta_2}{\beta_1 + \beta_2 - \beta_1 \beta_2} N$ with an upper bound $2N$ attained in the absence of interspecies competition ($\lambda_{12} = \lambda_{21} = 0$, or equivalently $\beta_1 = \beta_2 = 1$.)

Lastly, the exceptional case $\beta_1 = \beta_2 = 0$ resembles neutral evolution such that every point on a one dimensional line $c_1 + c_2 = 1$ is a fixed point. We shall call this fixed line scenario quasi-neutral evolution as the two species will not be neutral variants in the dilute limit if $s_o \neq 0$: A reproductive advantage near the origin does not destroy the coexistence line $c_1 + c_2 = 1$, but instead modifies the relative abundance of the two species as population size grows and saturates somewhere on the fixed line $c_1 + c_2 = 1$. The next section discusses the approach toward population size saturation in detail.

2.3 Growth of Population Size and Mapping to the Replicator Dynamics

Despite the rough similarity of the scenarios above to those of replicator dynamics, the competitive Lotka-Volterra model contains the overall population size as a dynamical variable. In general, growth and competition together do not conserve the population size, as illustrated in Fig. 2.3. In the limit $|\beta_1| \ll 1$ and $|\beta_2| \ll 1$, however, there is an attractive 1-d manifold of approximately fixed population size $c_1 + c_2 \approx 1$ on which conventional replicator dynamics determines the ultimate outcome. We shall refer to the competition near the line $c_1 + c_2 = 1$ in this limit as the competition under the replicator condition. Under the replicator condition, the strict balance between growth and competitive death results in an effective replicator dynamics with an approximately fixed population size, which we discuss below.
Figs. 2.1 and 2.2 illustrate competitions in this limit.

We now discuss the growth of population size toward the replicator condition and the eventual mapping onto the replicator dynamics. Upon using $c_T \equiv c_1 + c_2$ to measure the overall population size and defining $f \equiv c_1/c_T$ as the frequency of species 1, we obtain the following coupled dynamics of $c_T$ and $f$ from Eq. (2.4) and

\[ c_T \equiv c_1 + c_2 \]
\[ f \equiv c_1/c_T \]

\[ \beta_1 c_1 + \beta_2 c_2 = 0 \]

\[ \beta_1, \beta_2 > 0 \]

\[ \beta_1, \beta_2 < 0 \]

\[ s_o = 1, |\beta_2| = 0.028 \text{ and } |\beta_1| = 0.014 \]

The competition scenario depends on the sign of $\beta_1$ and $\beta_2$ similar to $\alpha_1$ and $\alpha_2$ in Fig. 1.2. (a),(b),(c), and (d) correspond to species 2-domination, mutualism, antagonism, and species 1-domination, respectively. Red circles represent fixed points and blue lines correspond to mean field trajectories of Eqs. (2.6) and (2.7) solved numerically. For $s_o = 1$, population size relaxes toward the replicator condition ($c_T \approx 1$) along a trajectory of constant $\rho$, which forms an upper branch of the parabola $c_1 = \rho^2 c_2^2$. Deviation from a trajectory of fixed $\rho$ only becomes apparent close to the line $c_T = 1$. Once the replicator condition is reached, the replicator dynamics at a fixed population size takes over.
Eq. (2.5),
\[
\frac{dc_T}{dt} = (1 + s_0 f)v_G(c_T) + (\alpha_1 + \alpha_2)f(1 - f)c_T^2, \tag{2.8}
\]
\[
\frac{df}{dt} = v_E(f) + s_0 f(1 - f)(1 - c_T), \tag{2.9}
\]
where the growth term \(v_G(c_T) \equiv c_T(1 - c_T)\) describes the logistic growth of population size, and the evolutionary dynamics term \(v_E(f) \equiv [\alpha_1 + (\alpha_1 + \alpha_2)f]f(1 - f)\) resembles the frequency-dependent selection in Eq. (1.2) with the identification
\[
\alpha_1 = (1 + s_0) \beta_1, \quad \text{and} \quad \alpha_2 = \beta_2. \tag{2.10}
\]
We first analyze the quasi-neutral evolution scenario when \(\beta_1 = \beta_2 = 0\), and then treat the case \(0 < |\beta_i| \ll 1\) as a weak perturbation. In the quasi-neutral scenario, \(\alpha_1 + \alpha_2 = 0\) and \(c_T\) obeys \(\frac{dc_T}{dt} = (1 + s_0 f)v_G(c_T)\). For any non-zero initial population size, \(c_T\) eventually saturates at \(c_T = 1\), which is an attractive 1-d manifold of fixed points in the original \(c_1-c_2\) phase space. As the population size grows from \(c_T(0) < 1\) or declines from \(c_T(0) > 1\) to saturate at \(c_T = 1\), \(c_1(t)\) and \(c_2(t)\) change to conserve the variable \(\rho\) defined by
\[
\rho \equiv c_2(t)/c_1(t)^{1/(1+s_0)},
\]
\[
= c_2(0)/c_1(0)^{1/(1+s_0)}, \tag{2.11}
\]
because Eq. (2.6) and Eq. (2.7) with \(\beta_1 = \beta_2 = 0\) implies \(d\rho/d\tilde{t} = 0\). To see how the frequency of each species changes as the population size approaches \(c_T = 1\), it’s helpful to rewrite \(\rho\) in terms of \(f\) and \(c_T\) as
\[
\rho = c_T(t)^{s_0/(1+s_0)}[1 - f(t)]/f(t)^{1/(1+s_0)}. \tag{2.12}
\]
Since \( \rho \) is a conserved variable, Eq. (2.12) implies that the frequency of a reproductively advantageous species increases as \( c_T(t) \) grows toward \( c_T = 1 \) when \( c_T(0) \ll 1 \). On the other hand, the frequency of a reproductively advantageous species decreases as \( c_T(t) \) declines toward \( c_T = 1 \) when \( c_T(0) \gg 1 \). If both species grow up at an equal rate \( (s_o = 0) \), the frequency of each is independently conserved, regardless of \( c_T(t) \).

For \( 0 < |\beta_1| \ll 1 \) and \( 0 < |\beta_2| \ll 1 \), the dynamics of population size away from \( c_T = 1 \) still obeys \( \frac{d\rho}{dt} \approx (1 + s_o f) v_G(c_T) \) since \( (1 + s_o f) v_G(c_T) \gg [\alpha_1 + \alpha_2] f (1 - f) c_T^2 \) in Eq. (2.8). Moreover, the approach toward \( c_T = 1 \) again follows a trajectory of approximately constant \( \rho \) since, away from \( c_T = 1 \),

\[
\left| \frac{d}{dt} \ln \rho \right| = \left| c_T [(\beta_2 - \beta_1) + (\beta_2 + \beta_1) f] \right| \\
\ll \left| (1 + s_o f) (1 - c_T) + [(1 + s_o) \beta_1 + \beta_2] f (1 - f) c_T \right| \\
= \left| \frac{d}{dt} \ln c_T \right| .
\]

Once \( c_T \) is in close proximity to 1, \( \rho \) is no longer conserved. The thin neighborhood of \( c_T = 1 \) in which conservation is violated, however, becomes vanishingly small in the limit \( |\beta_i| \ll 1 \). Accordingly, we can set \( c_T = 1 \) in Eq. (2.8) and Eq. (2.9) to find

\[
\frac{dc_T}{dt} \approx 0 \quad \text{and} \quad \frac{df}{dt} = v_E(f) ,
\]

which reproduces the replicator dynamics of a fixed population size \( N \). The mean field trajectories in Figs. 2.1 and 2.2 depict the approach toward the replicator condition in which the replicator dynamics at \( c_T = 1 \) determines how the frequency of each species changes. Fig. 2.1 illustrates growth along the bent trajectories \( c_1(t) = \rho^2 c_2(t)^2 \) that arises when species 1 has a reproductive advantage near the origin \( (s_o = 1) \), while Fig. 2.2 depicts growth along a set of straight lines of fixed
species’ frequency when $s_o = 0$.

### 2.4 Stochastic Description

In finite populations, the ultimate fate of the coupled system depends not only on the 3 dimensionless parameters $s_o, \beta_1, \beta_2$ and the initial condition, but also on fluctuations from the mean-field dynamics due to microscopic stochasticity. We can quantify the stochastic dynamics by regarding the microscopic rates in Eqs. (2.1)-(2.3) as Markov processes. The joint probability distribution of finding $N_i$ individuals of species $i$ at time $t$, $P(N_1, N_2, t)$, obeys the Master equation

$$
\partial_t P(N_1, N_2, t) = \mu_1(N_1 - 1)P(N_1 - 1, N_2, t) + \mu_2(N_2 - 1)P(N_1, N_2 - 1, t) 
+ \lambda_{11}N_1(N_1 + 1)P(N_1 + 1, N_2, t) + \lambda_{22}N_2(N_2 + 1)P(N_1, N_2 + 1, t) 
+ \lambda_{12}N_2(N_1 + 1)P(N_1 + 1, N_2, t) + \lambda_{21}N_1(N_2 + 1)P(N_1, N_2 + 1, t) 
- \left[ \mu_1N_1 + \mu_2N_2 + \lambda_{11}N_1(N_1 - 1) + \lambda_{22}N_2(N_2 - 1) 
+ \lambda_{12}N_2N_1 + \lambda_{21}N_1N_2 \right] P(N_1, N_2, t).
$$

In the limit $1/N \ll 1$ (recall that $N = \mu_1/\lambda_{11} = \mu_2/\lambda_{22}$), this discrete Master equation can be approximated by the Fokker-Planck equation for the continuous probability distribution $P(c_1, c_2, t)$ via the Kramers-Moyal expansions or the Van-Kampen expansions [24, 26]. The corresponding Fokker-Planck equation reads

$$
\partial_t P(e, t) = \sum_{i=1}^2 \left( -\partial_{c_i}[v_i(e)P(e, t)] + \frac{1}{2N} \partial^2_{c_i}D_i(e)P(e, t) \right),
$$

(2.14)
Chapter 2: Population Genetics with Fluctuating Population Sizes

where the deterministic drift and $N$-independent diffusion coefficients are

$$v_1(c) = \mu_1 c_1 (1 - c_1 - c_2) + \mu_1 \beta_1 c_1 c_2,$$  \hspace{1cm} (2.15)

$$v_2(c) = \mu_2 c_2 (1 - c_1 - c_2) + \mu_2 \beta_2 c_1 c_2,$$  \hspace{1cm} (2.16)

$$D_1(c) = \mu_1 c_1 (1 + c_1 + c_2) - \mu_1 \beta_1 c_1 c_2,$$  \hspace{1cm} (2.17)

$$D_2(c) = \mu_2 c_2 (1 + c_1 + c_2) - \mu_2 \beta_2 c_1 c_2.$$  \hspace{1cm} (2.18)

An equivalent representation in terms of the Ito calculus [26] prescribes stochastic dynamics of the $c_i(t)$ that resembles a set of coupled Langevin equations:

$$\frac{dc_i}{dt} = v_i(c) + \sqrt{\frac{D_i(c)}{N}} \Gamma_i(t),$$  \hspace{1cm} (2.19)

where $\Gamma_i(t)$ is a Gaussian white-noise with $\langle \Gamma_i(t) \Gamma_j(t') \rangle = \delta_{i,j} \delta(t-t')$ and $\langle \Gamma_i(t) \rangle = 0$. In the limit of infinitely large population size $N$, the noise term of order $\sqrt{1/N}$ in Eq. (2.19) vanishes and we recover the mean-field description of Eq. (2.6) and Eq. (2.7).

For finite $N$, number fluctuations alter the mean-field description and can lead to outcomes different from the deterministic predictions. For instance, fluctuations will eventually drive one of the two species to fixation and destroy stable coexistence for mutualism. Regardless of the deterministic phase portraits, the eventual fate of the system at long times is fixation of a single species. Once one species becomes fixed, the dynamics of the fixed species follow stochastic logistic growth while the other species remains forever extinct as is easily checked from Eqs. (2.15)-(2.19). The $c_1$ and $c_2$ axes are thus absorbing boundaries corresponding to the fixation of species 1 and of species 2, respectively. The only absorbing state in the phase space is total extinction $(0,0)$, which is inaccessible since we don’t allow for the death process.
$S_i \to \phi$ in our simulations. Had we included the death process, the total extinction would nevertheless be extremely unlikely because the mean time to extinction for logistic growth grows exponentially with $N$ [85].

If time is non-dimensionalized to $\tilde{t} = \mu_2 t$ in Eqs. (2.15)-(2.19), one can see that the stochastic dynamics depend on 4 dimensionless parameters: $s_o, \beta_1, \beta_2$, and $1/N$. The parameters $s_o, \beta_1, \beta_2$ control both the mean-field phase portrait and the diffusion.

Figure 2.2: Replicator dynamics with genetic drift and population size fluctuations when $c_T \approx 1$. (a),(b),(c), and (d) correspond to species 2-domination, mutualism, antagonism, and species 1-domination, respectively with $s_o = 0$, and $|\beta_1| = |\beta_2| = 0.03$. In this case, the mean field trajectories approach the replicator condition ($c_1 + c_2 = c_T \approx 1$) as straight lines that preserve the frequency $f$. Each orange stochastic trajectory, simulated from the Gillespie’s algorithm with the initial condition $(f, c_T) = (0.5, 0.5)$ (i.e., $c_1 = c_2 = 0.25$) and $N = 100$ individuals, demonstrates typical fixation events. Stochastic trajectories show rapid growth of population size toward the replicator condition where selection, genetic drift, and population size fluctuations ultimately determine the competition outcome.
coefficients, while the parameter $1/N$ sets the strength of fluctuations relative to deterministic drift.

In the following sections and in the next Chapter, we investigate the dynamics in different ranges of $s_o, \beta_1, \beta_2$ assuming small fluctuations $1/N \ll 1$. In Sec. 2.5, we study the limit when $|\beta_i| \ll 1$ and $s_o = 0$, with typical stochastic trajectories shown

![Phase portraits for $s_o = 0$, and $|\beta_1| = |\beta_2| = 0.5$, which includes the case of strong mutualism of Chapter 3. Competition at long times no longer takes place close to the line $c_1 + c_2 = c_T = 1$, depicted as the red dashed line in (b) and (c), but with a varying overall population size. (a),(b),(c), and (d) correspond to species 2-domination, mutualism, antagonism, and species 1-domination, respectively. The initial condition for the stochastic simulation is $(f, c_T) = (0.5, 0.5)$ (i.e., $c_1 = c_2 = 0.25$) with $N = 50$ individuals. In contrast to mutualism under the replicator condition (e.g., Fig. 2.2(b)), the coexistence fixed point in strong mutualism shown in (b) is highly stable and fixation becomes a rare event even when $N$ is as small as 50.](image-url)
in Fig. 2.2. In Sec. 2.6, we analyze the quasi-neutral evolution case \( \beta_1 = \beta_2 = 0 \), but allowing \( s_o \), the reproductive advantage near the origin, to be large. Both Secs. 2.5 and 2.6 describe dynamics under the replicator condition, with the only difference being the \( s_o \)-dependent and initial-condition-dependent trajectories describing the approach to \( c_T = 1 \). In Chapter 3, we study a strong mutualism scenario where \( \beta_1 \) and \( \beta_2 \) are \( O(1) \) and we set \( s_o = 0 \) for simplicity. Fig. 2.3 illustrate the stochastic dynamics in this case, where the time dependence of the overall population size plays a crucial role. In this situation, replicator dynamics is never an appropriate description.

### 2.5 Replicator Dynamics with Genetic Drift and Population Size Fluctuations

We now study competitive Lotka-Volterra dynamics under the replicator condition (\( c_T \approx 1 \) and \( |\beta_i| \ll 1 \)) with \( s_o = 0 \) and \( 1/N \ll 1 \), thus extending Eq. (1.17) to include fluctuations in the overall population size. We focus on the dynamics of the frequency of species 1 \( f(t) \) and the total population size \( c_T(t) \). When \( s_o = 0 \) and \( |\beta_i| \ll 1 \), Appendix A.1 shows that the coupled stochastic dynamics of \( f \) and \( c_T \) for \( c_T \approx 1 \) are similar to those studied by Pigolotti et al. [41], which read

\[
\frac{df}{dt} = \mu v_E(f) c_T + \sqrt{\frac{\mu D_g(f)}{N} \left( \frac{1 + c_T}{c_T} \right)} \Gamma_f(t), \tag{2.20}
\]

\[
\frac{dc_T}{dt} = \mu v_G(c_T) + \sqrt{\frac{\mu c_T(1 + c_T)}{N}} \Gamma_{c_T}(t), \tag{2.21}
\]

where \( \Gamma_i(t) \) is an uncorrelated Gaussian white noise with zero mean and unit variance, \( D_g(f) = f(1 - f) \) is the frequency-dependent genetic drift coefficient [4, 86],
\( v_E(f) \) and \( v_G(c_T) \) are the usual selection function and the logistic growth function mentioned earlier. These stochastic differential equations, which arise from a more general dynamics with \( s_o \neq 0 \) discussed in Appendix A.1, must be interpreted in terms of Ito calculus \([24, 26]\) in order to correctly reproduce the Fokker-Planck Equation (2.14). We have retained the original unit of time to make the reproduction timescale explicit and denoted \( \mu = \mu_2 \) for brevity.

In Eq. (2.21), the dynamics of population size is \( f \)-independent and exhibits the balance between fast approximately deterministic relaxation toward the equilibrium line \( c_T = 1 \) and slow fluctuations with variance \( 1/N \) around this equilibrium. On the other hand, the dynamics of \( f \) depends on \( c_T \). Nevertheless, it is mostly influenced by the mean population size \( \langle c_T \rangle = 1 \) since the variance of \( c_T \) about \( c_T = 1 \) is \( 1/N \ll 1 \). Thus, the effective dynamics of \( f \), accurate to first order in \( 1/N \), can be approximated by replacing \( c_T = 1 \), resulting in

\[
\frac{df}{dt} = \mu v_E(f) + \sqrt{2\mu N D_g(f) \Gamma_f(t)}.
\] (2.22)

Eq. (2.21) and Eq. (2.22) together describe the dynamics near the replicator condition when \( 1/N \ll 1 \), which is precisely Eq. (1.17) including an independently fluctuating population size around the fixed mean \( c_T = 1 \) (after identifying \( \mu^{-1} \) as the generation time \( \tau_g \) in Eq. (1.17).)

Note that the variance per generation time of Eq. (2.22) given by \( f(1 - f)/N \) is independent of both the population size fluctuations away from \( c_T = 1 \) and the selection mechanism in the vicinity of this line. In fact, the variance resembles that of a Wright-Fisher or Moran model [4]. Thus, the variance effective population size deduced from genetic drift is equivalent to the mean population size \( N \) despite fluctuations in the overall population size. However, as we show in the next section,
different results obtain when \( s_o \neq 0 \) and the variance differs from that of a Wright-Fisher or Moran model.

From the closed form Eq. (2.22) for \( f(t) \), we can recover results for the fixation probability \( u(f) \) and the mean fixation time \( \tau(f) \) which are, respectively, the probability that species 1 become fixed (instead of species 2) and the average time to lose heterozygosity provided species 1 initially has a frequency \( f \) at \( c_T = 1 \). These quantities obey ordinary differential equations,

\[
\begin{align*}
v_E(f) \frac{du}{df} + \frac{D_g(f)}{N} \frac{d^2 u}{df^2} &= 0, \\
v_E(f) \frac{d\tau}{df} + \frac{D_g(f)}{N} \frac{d^2 \tau}{df^2} &= -\frac{1}{\mu},
\end{align*}
\]

subject to the boundary conditions \( u(0) = 0, u(1) = 1 \) and \( \tau(0) = \tau(1) = 0 \) [24, 26]. The differential equations can be integrated directly leading to close form solutions which read

\[
\begin{align*}
u(f) &= \int_0^f e^{-N\Psi(x)} \frac{dx}{\int_0^1 e^{-N\Psi(x)} \frac{dx}{dx}}, \\
\tau(f) &= I(1)u(f) - I(f),
\end{align*}
\]

where \( \Psi(x) \equiv \int_0^x v_E(y)/D_g(y) dy \), and

\[
I(f) \equiv \left[ \int_0^f dx e^{-N\Psi(x)} \int_1^x dy e^{N\Psi(y)/D_g(y)} \right](N/\mu).
\]

We now review the implications of Eq. (2.22) for different selection scenarios. In neutral evolution \( (\beta_1 = \beta_2 = 0) \), Eq. (2.22) becomes (with \( D_g(f) = f(1-f) \))

\[
\frac{df}{dt} = \sqrt{\frac{2\mu}{N} D_g(f) \Gamma_f(t)},
\]

which is a continuous approximation of the Moran model or the Wright-Fisher sam-
Chapter 2: Population Genetics with Fluctuating Population Sizes

pling in population genetics [4, 5, 86–88]. Only genetic drift participates in the dynamics and fixation events are results of an unbiased random walk with \( c_T \approx 1 \) toward \( f = 0 \) or \( f = 1 \), absorbing boundaries, and independent fluctuations of population size about the mean \( N \). In this case, direct evaluation of Eqs. (2.25) and (2.26) gives [87]

\[
\begin{align*}
    u_{\text{neutral}}(f) &= f, \\
    \tau_{\text{neutral}}(f) &= -\left( \frac{N}{\mu} \right) \left[ f \ln f + (1 - f) \ln(1 - f) \right].
\end{align*}
\]  

(2.28) (2.29)

For selection that favors domination of one species, the special case \( \tilde{s} \equiv \beta_1 = -\beta_2 \) reproduces the Moran process with an effective selective advantage \( \tilde{s} \) (provided \( c_T \approx 1 \)), described by

\[
\frac{df}{dt} = \mu \tilde{s} f (1 - f) + \sqrt{\frac{2\mu}{N} D_{\theta}(f) \Gamma_f(t)}. 
\]  

(2.30)

We emphasize that the growth rates of the two species when \( c_T \ll 1 \) in this particular competitive Lotka-Volterra dynamics are strictly identical (\( s_o = 0 \)), but the species with positive \( \beta_i \) behaves near \( c_T = 1 \) as if it has a selective advantage \( \tilde{s} \). Upon evaluating Eq. (2.25), we arrive at the celebrated Kimura result for the fixation probability [89]

\[
u(f) = \frac{1 - e^{-\tilde{s}Nf}}{1 - e^{-\tilde{s}N}}. 
\]  

(2.31)

A lengthy closed-form formula for the mean fixation time can also be obtained, see for example Ref. [30].

For antagonistic or mutualistic interactions, the effective dynamics of \( f \) reads

\[
\frac{df}{dt} = \mu \beta f (1 - f)(f^* - f) + \sqrt{\frac{2\mu}{N} D_{\theta}(f) \Gamma_f(t)}, 
\]  

(2.32)
where \( f^* = \beta_1 / (\beta_1 + \beta_2) \) is the coexistence fixed point with \( c_T = 1 \) and \( \tilde{\beta} \equiv (\beta_1 + \beta_2) \) controls the stability of \( f^* \). The parameter \( \tilde{\beta} \) is positive and negative for mutualism and antagonism, respectively. In either case, the fixation probability directly follows from Eq. (2.25), and is given by

\[
u(f) = \frac{\int_0^f e^{\frac{N\tilde{\beta}}{2}(f^*-f)^2} df}{\int_0^1 e^{\frac{N\tilde{\beta}}{2}(f^*-f)^2} df},
\]

in agreement with Ref. [17]. The mean fixation time from Eq. (2.26) can not be simplified further, and must be evaluated numerically.

Pigolotti et al. simulated the fixation probability for different competition scenarios under the replicator condition with \( s_o = 0 \) and found good agreement with these predictions of the fixation probabilities even for fairly small population sizes of \( O(N) \sim 100 \) individuals [41]. Constable et al. also studied this limit using a different mathematical technique and found good agreement between theories and simulations of both the fixation probability and the mean fixation time [90]. These results confirm that the competitive Lotka-Volterra model reduces to replicator dynamics with genetic drift and an independently fluctuating overall population size, \( provided s_o = 0, |\beta_i| \ll 1, \text{ and } c_T \approx 1. \) The remainder of this chapter and in chapter 3 explore regimes where this is no longer the case.

### 2.6 Fluctuation-induced Selection in Quasi-Neutral Evolution

In this section, we discuss a fluctuation-induced selection mechanism triggered by the feedback between population size fluctuations and evolutionary game dynamics near the equilibrium line \( c_T = 1 \) when \( s_o \neq 0 \). We focus on quasi-neutral evolution
(i.e. $\beta_i = 0$, but $s_o \neq 0$) so that the effect of fluctuation-induced selection is most pronounced due to the absence of the usual selection mechanism of population genetics [4, 5, 87]. As shown in Appendix A.1, the coupled stochastic dynamics of $f$ and $c_T$ near $c_T = 1$ for quasi-neutral evolution read

$$\frac{df}{dt} = v_R(f, c_T) + \sqrt{\frac{D_g(f)}{N}} \left( 1 + \frac{c_T}{c_T} \right) \left( 1 + s_o(1 - f) \right) \Gamma_f(t), \quad (2.34)$$

$$\frac{dc_T}{dt} = (1 + s_o f) v_G(c_T) + \sqrt{\frac{c_T(1 + c_T)}{N}} (1 + s_o f) \Gamma_{c_T}(t), \quad (2.35)$$

Figure 2.4: A schematic illustration of fluctuation-induced selection at $c_T \approx 1$ for $s_o > 0$ in quasi-neutral evolution. A combination of slow population size fluctuations and fast relaxation toward $c_T = 1$ along a warped trajectory $c_2 = (\rho c_1)^{1+s_o}$ of conserved $\rho$, depicted by the blue curves, generates an effective selection term $\tilde{v}(f)$ at $c_T \approx 1$. Remarkably, fluctuation-induced selection disfavors fixation of species with a reproductive advantage near the origin (i.e., $\tilde{v}(f) < 0$ for $s_o > 0$). The effective stochastic dynamics of $f$ at $c_T \approx 1$ after adiabatic elimination of the fast variable $c_T$ is given by Eq. (2.46). The effective selection term $\tilde{v}(f)$ given by Eq. (2.44) differs from the naive approximation $v_R(f, c_T = 1)$ in Eq. (2.36) by a factor of $2(1 + s_o f)^2/(1 + s_o)$, leading to the improved agreement between simulation and theory when $\mathcal{O}(s_o) \sim 1$ in Fig. 2.5.
where \( v_R(f, c_T) = s_o f (1 - f) \left[ (1 - c_T) - \frac{1}{N} \left( \frac{1 + c_T}{c_T} \right) \right] \) is the deterministic drift due to the reproductive advantage near the origin \( s_o \) and \( v_G(c_T) = c_T (1 - c_T) \) is the usual logistic growth of population size. When \( s_o = 0 \), Eqs. (2.34) and (2.35) reduce to the Moran model for neutral evolution with a fluctuating population size given by Eqs. (2.20) and (2.21). When \( s_o \neq 0 \), fluctuations of population size becomes \( f \)-dependent with variance proportional to \((1 + s_o f)/N\), while \( f \) acquires an intriguing deterministic drift at \( c_T = 1 \) of the form \( v_R(f, c_T = 1) = -2s_o f (1 - f)/N \), which actually favors the fixation of the species with a reproductive disadvantage near the origin \((s_o < 0)\). The presence of non-vanishing deterministic drift is in stark contrast to the unbiased random walk behavior of neutral evolution along the equilibrium line \( c_T = 1 \) displayed in Eq. (2.27). For the generalization of Eqs. (2.34) and (2.35) that reveals the role of a non-vanishing selection function \( v_E(f) c_T \) in other competition scenarios, see Appendix A.1.

Eq. (2.35) drives small excursions from \( c_T = 1 \) which changes the dynamics of \( f \). To understand in more detail how quasi-neutral evolution with a reproductive advantage near the origin \( s_o \) differs from the classic Moran model, we seek an effective dynamics of \( f \) near the fixed equilibrium population size \( c_T = 1 \). Lin et al. addressed a similar problem in the context of ecology, and deduced an effective dynamics of \( z \equiv c_1 - c_2 \) at \( c_T = 1 \) to first order in \( 1/N \), using a combination of heuristic arguments and asymptotic expansions in powers of \( 1/N \) [91]. Here, we present an alternative argument based on adiabatic elimination of a fast variable which appears to be more direct. The method exploits an appropriate choice of coordinates and can be easily extended to other competition scenarios. With the effective dynamics in hand, we then calculate the fixation probability as well as the mean fixation time and compare to numerical simulations. Our stochastic simulations employ the Gillespie’s algorithm to efficiently simulate the discrete Master equation of Sec. 2.4.
Chapter 2: Population Genetics with Fluctuating Population Sizes

The simulated fixation probabilities and the mean fixation times for each initial condition are constructed from $10^4$ realizations of fixation events.

2.7 A Naive Approximation

In the limit $1/N \ll 1$, one may attempt to close Eq. (2.34) for $f$ by substituting $c_T = 1$ and ignoring weak population size fluctuations of order $1/N$, similar to the analysis in Sec. 2.2.5. This naive approximation yields

$$\frac{df}{dt} = -\frac{2s_o}{N}f(1-f) + \sqrt{\frac{2D_o(f)}{N}}(1+s_o(1-f))\Gamma_f(t). \quad (2.36)$$

Straightforward application of Eqs. (2.25) and (2.26) now leads to $s_o$-dependent corrections to the fixation probability and the mean fixation time,

$$u(f) = \frac{u_{\text{neutral}}(f)}{1 + s_o(1-f)}, \quad (2.37)$$

$$\tau(f) = \frac{\tau_{\text{neutral}}(f)}{1 + s_o(1-f)}, \quad (2.38)$$

where $u_{\text{neutral}}(f)$ and $\tau_{\text{neutral}}(f)$ are given by Eqs. (2.28) and (2.29). Although at $s_o = 0$ we recover the results of the Moran model, dashed lines in Fig. 2.5 show that these approximations are poor even when $O(s_o) \sim 1$. These errors stem from neglecting overall population size fluctuations of the term $s_o f(1-f)(1-c_T)$, which (as we shall see) actually contribute a deterministic drift of order $s_o/N$ comparable to the term kept in Eq. (2.36).
Chapter 2: Population Genetics with Fluctuating Population Sizes

Figure 2.5: The fixation probability (top) and the mean fixation time in the units of $\mu/N$ (bottom) as a function of the initial frequency $f$ with the initial overall population size $c_T = 1$. Predictions from adiabatic elimination of a fast variable shown in solid lines are in excellent agreement with simulations shown in symbols, while dashed lines are predictions from our “naive approximation.” The fixation probability is $N$-independent while the mean fixation time scales linearly with $N$, which are features of unbiased random walk (genetic drift.) Population size fluctuations, however, induce selection that disfavors a species that grows faster near the origin ($s_o > 0$), resulting in a decline in the fixation probability as well as a reduced mean fixation time.
2.8 Effective Evolutionary Dynamics from Adiabatic Elimination of the Fast Population Size Variable

We now employ adiabatic elimination of a fast variable to deduce an effective evolutionary dynamics for an approximately fixed population size close to $N$. Motivated by the approximate conservation of the composite variable $\rho$ in the mean field theory (see Eq. (2.11)), we calculate the stochastic dynamics of $\rho$ from Ito’s change of variable formula [24, 26], and find

$$\frac{d\rho}{dt} = \frac{v_\rho(f, c_T)}{N} + \sqrt{\frac{2D_\rho(f, c_T)}{N}} \Gamma_\rho(t), \tag{2.39}$$

where

$$v_\rho(f, c_T) = \frac{1}{2} \left( \frac{2 + s_o}{1 + s_o} \right) \left( \frac{1 - f}{f} \right) \left( \frac{1 + c_T}{c_T} \right) \left( \frac{c_T}{f} \right)^{1/(1 + s_o)} \tag{2.40}$$

and

$$D_\rho(f, c_T) = \frac{1}{2} \left( \frac{1 + s_o f}{1 + s_o} \right) \left( \frac{1 - f}{f} \right) \left( \frac{1 + c_T}{c_T} \right) \left( \frac{c_T}{f} \right)^{2/(1 + s_o)} \tag{2.41}$$

Eqs. (2.39)-(2.41) reveal that $\rho$ varies on a slow time scale of order $1/N \ll 1$ everywhere in our domain of interest. On the other hand, the dynamics of the overall population size given by Eq. (2.35) exhibits a fast relaxation toward $c_T = 1$, after which slow fluctuations of order $1/N$ take over. Since $\rho$ and $c_T$ together completely specify the state of the system, the dynamics of the system starts with a rapid quasi-deterministic relaxation toward $c_T = 1$ along a trajectory of fixed $\rho$; then the slow residual dynamics of $\rho$ take over. The slow dynamics of the coordinate $\rho$ generates an effective dynamics of $f$ when $c_T \approx 1$. Fig. 2.4 depicts the fluctuation-induced
selection emerging from the slow stochastic dynamics of $\rho$ near $c_T = 1$.

To eliminate the fast variable, we integrate out $c_T$ in the joint probability distribution of $c_T$ and $\rho$ at time $t$, $P(c_T, \rho, t)$, and obtain the marginal probability distribution $\tilde{P}(\rho, t) \equiv \int P(c_T, \rho, t) dc_T$. The Fokker-Planck equation for the marginal probability distribution dictates the effective dynamics of the remaining slow variable $\rho$. Motivated by the separation of timescales, we factorize $P(c_T, \rho, t) = P_{st}(c_T)P_{\rho}(\rho, t)$, assuming $c_T$ rapidly relaxes to $c_T = 1$ and forms a quasi-stationary distribution $P_{st}(c_T)$ before $\rho$ varies significantly. In other words, $c_T$ is slaved to $\rho$ [26]. Upon substituting this factorization into the Fokker-Planck equation associated with Eqs. (2.35) and (2.39), we find

$$\partial_t P(c_T, \rho, t) = -\nabla \cdot \vec{J}(c_T, \rho, t)$$

$$= - \left[ \frac{1}{N} \partial_\rho v_\rho(f, c_T)P_\rho(\rho, t) - \frac{1}{N} \partial_\rho^2 D_\rho(f, c_T)P_\rho(\rho, t) \right] P_{st}(c_T),$$

where the probabilistic current in the $c_T$ direction vanishes by the assumption of stationarity. Integrating out $c_T$ then leads to

$$\partial_t \tilde{P}(\rho, t) = -\frac{1}{N} \partial_\rho v_\rho(f, \langle c_T \rangle)\tilde{P}(\rho, t) + \frac{1}{N} \partial_\rho^2 D_\rho(f, \langle c_T \rangle)\tilde{P}(\rho, t), \quad (2.42)$$

where $\langle \cdot \rangle$ denotes an expectation value. The effective Langevin dynamics associated with Eq. (2.42) is precisely Eq. (2.39) with the substitution $c_T = \langle c_T \rangle$, which is here simply the equilibrium population size $\langle c_T \rangle = 1$. For a discussion of higher order corrections in powers of $1/N$ in adiabatic elimination of fast variables, see Refs [26] and [25].

We now determine the effective evolutionary dynamics when $c_T \approx 1$ by substituting $\rho(f)$ for $c_T \approx 1$, i.e. $\rho = (1 - f)/f^{1/(1+\kappa_0)}$. The Fokker-Planck equation for $\rho$ can be converted to the Fokker-Planck equation for $f$ along the line $c_T = 1$ via the
Chapter 2: Population Genetics with Fluctuating Population Sizes

chain rule \( d/d\rho = -[(1 + s_o)f(1 - f)/((1 + s_o)f)\rho]d/df \). The calculation is more easily carried out using the backward Kolmogorov equation, since derivatives only act on the probability distribution. A straightforward calculation leads to an effective Fokker-Planck equation of \( f \) for \( c_T \approx 1 \),

\[
\partial_t \tilde{P}(f, t) = -\partial_f \tilde{v}(f) \tilde{P}(f, t) + \frac{1}{N} \partial_f^2 \tilde{D}(f) \tilde{P}(f, t),
\]

(2.43)

where

\[
\tilde{v}(f) = -\left( \frac{1}{N} \right) s_o (1 + s_o) \frac{f(1 - f)}{(1 + s_o f)^2},
\]

(2.44)

and

\[
\tilde{D}(f) = D_g(f) \left( \frac{1 + s_o}{1 + s_o f} \right).
\]

(2.45)

Hence, the effective dynamics of \( f \) reads

\[
\frac{df}{dt} = \tilde{v}(f) + \sqrt{\frac{2\tilde{D}(f)}{N}} \Gamma_f(t),
\]

(2.46)

where \( \tilde{v}(f) \) is given by Eq. (2.44) and describes fluctuation-induced selection term displayed in Fig. 2.4, and \( \tilde{D}(f) \) is the effective genetic drift coefficient given by Eq. (2.45). Eq. (2.46) reduces to the Moran model for neutral evolution at \( s_o = 0 \). For \( s_o \neq 0 \), not only does fluctuation-induced selection appear, but we also obtain an effective genetic drift that differs from the Wright-Fisher sampling by a frequency-dependent factor \( (1 + s_o)/(1 + s_o f) \).

The fixation probability and the mean fixation time with an initial condition on the equilibrium line \( c_T = 1 \) follow immediately from Eq. (2.25) and (2.26):

\[
u(f) = \frac{2 + s_o f}{(2 + s_o)} u_{\text{neutral}}(f),
\]

(2.47)
\[ \tau(f) = -\left( \frac{N}{\mu} \right) \left[ \left( \frac{1 + \frac{s_o f}{2}}{1 + s_o} \right) f \ln f + \left( \frac{1 + \frac{s_o (1-f)}{2}}{1 + s_o} \right) (1-f) \ln(1-f) \right. \\
\left. + \frac{s_o^2}{2(1+s_o)(2+s_o)} f(1-f) \right]. \quad (2.48) \]

Eqs. (2.46)-(2.48) are in agreement with the results by Lin et al. [91] after an appropriate change of variable. At small \( s_o \), both Eq. (2.47) and Eq. (2.37) give \( u(f) = [1 - s_o (1 - f)] u_{\text{neutral}}(f) + O(s_o^2) \) while both Eq. (2.48) and Eq. (2.38) give \( \tau(f) = [1 - s_o (1 - f)] \tau_{\text{neutral}}(f) + O(s_o^2) \), reducing to the standard results of the Moran model when \( s_o = 0 \). Fig. 2.5 shows the predictions from Eq. (2.47) and (2.48) are in excellent agreement with the stochastic simulations.

Upon inoculating an equal mixture of each species and assuming species 1 has a reproductive advantage near the origin \( s_o \), the fixation probability of species 1 is \( 1/4 + 1/(4 + 2s_o) \) which monotonically decreases from 1/2 when \( s_o = 0 \) to 1/4 as \( s_o \rightarrow \infty \). Moreover, by defining \( \tilde{f} \) such that the fixation probability \( u(\tilde{f}) = 1/2 \), we find \( \tilde{f} = (2 + s_o)/(2 + \sqrt{4 + 2s_o(2 + s_o)}) \) which rises monotonically from \( \tilde{f} = 1/2 \) when \( s_o = 0 \) to \( \sqrt{2}/2 \) as \( s_o \rightarrow \infty \). Consequently, the faster growing species near the origin is only more likely to survive provided the initial fraction is biased in its favor, \( f \in [\sqrt{2}/2, 1] \approx [0.707, 1] \) for \( c_T = 1 \), confirming that population size fluctuations disfavor the ultimate survival of a species with a reproductive advantage near the origin.

### 2.9 Dimensional Reduction: the Fixation Probability and the Mean Fixation Time

Since the dynamics also contains the overall population size degree of freedom, the initial frequency \( f^0 \) and the initial population size \( c_T^0 \) will both in general enter
Figure 2.6: Dimensional reduction by adiabatic elimination of the fast population size variable for $s_o = 1$ and $N = 100$. In all figures, solid lines are analytical predictions constructed in Sec. 2.8 while symbols are simulation results. (a) and (c) show the fixation probability and the mean fixation time as a function of initial frequency $f$ in a dilute ($c_T = 0.5$), optimal ($c_T = 1$), and overcrowded ($c_T = 1.5$) initial population size. When plotted against the slow variable $\rho$, (b) and (d) show data collapse of the fixation probability and the mean fixation time onto $u_{s_o=1}(\rho)$ and $\tau_{s_o=1}(\rho)$. (a) demonstrates that population size degree of freedom plays a crucial role in determining the fate of competition; a wise strategy for the species with a reproductive advantage near the origin is to start in a dilute population size. On the other hand, a species with a reproductive disadvantage near the origin is better off starting in an overcrowded population size.

The fixation probability $u(f^0, c_T^0)$ and the mean fixation time $\tau(f^0, c_T^0)$. To keep the notation simple, we now continue with the practice of setting $f \equiv f^0$ and $c_T \equiv c_T^0$. Separation of dynamical timescales, in fact, implies the fixation probability and the mean fixation time are universal functions of the slow variable $\rho = (1 - f)f^{-1/(1+s_o)}c_T^{s_o/(1+s_o)}$, provided $1/N \ll 1$. In other words, $u(f, c_T) = u(f', c_T') = u(\rho)$.
and $\tau(f, c_T) = \tau(f', c'_T) = \tau(\rho)$ if $\rho(f, c_T) = \rho(f', c'_T) = \rho$. These simplifications arise from a rapid quasi-deterministic relaxation of the population size, with $\rho$ fixed, toward the line $c_T = 1$, after which the slow stochastic dynamics of $\rho$ dictates the outcome.

In principle, $u(\rho)$ and $\tau(\rho)$ follow from rewriting $f$ as a function of $\rho$ in Eqs. (2.47) and (2.48). For an arbitrary $s_o$, however, $f$ cannot easily be expressed as a function of $\rho$ at $c_T = 1$ because they are related by an $s_o$-dependent transcendental equation

$$\rho(f, c_T = 1) = (1 - f) f^{-1/(1+s_o)}.$$  \hspace{1cm} (2.49)

One can nevertheless extract $u(\rho)$ and $\tau(\rho)$ from Eqs. (2.47) and (2.48) by numerically solving Eq. (2.49).

Consider the particularly simple case $s_o = 1$, where the physically relevant closed-form solution associated with Eq. (2.49) is $f(\rho) = 1 + \frac{1}{2} \left( \rho^2 - \rho \sqrt{\rho^2 + 4} \right)$. Substituting $f(\rho)$ into Eqs. (2.47) and (2.48) now yields analytical results of $u_{s_o=1}(\rho)$ and $\tau_{s_o=1}(\rho)$. It is also possible to reconstruct $u_{s_o=1}(f, c_T)$ and $\tau_{s_o=1}(f, c_T)$ for arbitrary $f$ and $c_T$ from $u_{s_o=1}(\rho)$ and $\tau_{s_o=1}(\rho)$ by a direct substitution $\rho = (1 - f) f^{-1/2} c_T^{1/2}$.

To test these predictions, we simulated $10^4$ fixation events per each initial condition, with $N = 100$ and with $c_T = 0.5$, $c_T = 1$, $c_T = 1.5$ representing dilute, optimal, and overcrowded initial population sizes with $s_o = 1$. Figs. 2.6(a) and 2.6(c) show excellent agreement between $u_{s_o=1}(f, c_T)$ as well as $\tau_{s_o=1}(f, c_T)$ and the simulations. Figs. 2.6(b) and 2.6(d) show data collapse of the fixation probability and the mean fixation time onto $u_{s_o=1}(\rho)$ and $\tau_{s_o=1}(\rho)$ constructed above. These results demonstrate that population size degree of freedom plays a crucial role in determining the results of competition, here through the composite variable $\rho = (1 - f) f^{-1/(1+s_o)} c_T^{s_o/(1+s_o)}$. 

49
In this chapter, we study a strong mutualism scenario ($\beta_i \sim O(1)$ in Fig. 2.3(b)), where the replicator condition is no longer satisfied. In this limit, the coexistence fixed point shifts far away from the line $c_T = 1$ and becomes strongly attractive in all eigen-directions. The faint orange grid in Fig. 2.3(b) illustrates a typical fixation trajectory exhibiting a decline of overall population size as weak fluctuations about the strongly stable fixed point eventually drive one of the two species (in this case, species 1) to fixation.

In Sec. 3.1, we demonstrate the failure of replicator dynamics with genetic drift to describe strong mutualism with a varying population size. We then construct the fixation probability and the mean fixation time allowing an arbitrary initial population size and an initial frequency from the method of matched asymptotic expansions. The details of the matched asymptotic expansions to achieve the result of this chapter are provided in Appendix B.1 and B.2.

### 3.1 Failure of the Fixed Population Size Model near Boundary Layers

Suppose we accept Eq. (1.17) as a phenomenological model for mutualism and fit the resulting fixation probability in Eq. (2.33) to empirical data; how well would
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

this model with a strictly fixed population size do? To motivate the choice of fitting parameters, we first discuss the behavior of the fixation probability $u(f)$ predicted by Eq. (2.33). For $\tilde{\beta}N \ll 1$ (recall that $\tilde{\beta} \equiv \beta_1 + \beta_2$ in Eq. (2.33)), genetic drift dominates mutualistic selection and the fixation probability approaches the result of an unbiased random walk of neutral evolution, Eq. (2.28). For $\tilde{\beta}N \gg 1$, the coexistence fixed point is metastable and fixation driven by weak genetic drift becomes a rare event. Initial conditions in close proximity to $f^*$ almost surely visit $f^*$ before fixation occurs, giving rise to a plateau of equal fixation probability $u(f^*)$ in the neighborhood of $f^*$. Furthermore, the fixation probability $u(f)$ only varies rapidly within the boundary layers of width $\sim 1/N$ adjacent to each absorbing states $f = 0$ and $f = 1$, away from which $u(f)$ exhibits crossovers to a plateau value $u(f^*)$. The boundary layers near the absorbing states contain initial conditions that can be driven by genetic drift to fixation before being attracted toward $f^*$. For symmetric mutualism ($f^* = 1/2$), the plateau height $u(f^*)$ is 1/2 by symmetry from Eq. (2.33) and is independent of $N$. For asymmetric mutualism, the $N$-dependent behavior of $u(f^*)$ can be understood by studying rare event escape from a metastable state. For an evolutionary game with a stable coexistence fixed point, it can be shown that $u(f^*)$ is given by the ratio of the flux into the absorbing state $f = 1$ to the total flux into the absorbing states $f = 0$ and $f = 1$ whose $N$-dependent behavior in the limit $N \gg 1$ is given by [31, 32]

$$u(f^*) \approx \frac{1}{1 + e^{-N\Delta S_0 + \Delta S_1}},$$

(3.1)

where, from the perspective of a Feynman’s path integral formulation of the stochastic dynamics [94, 95], $\Delta S_0 \equiv S_0[\gamma_{f^* \rightarrow 1}] - S_0[\gamma_{f^* \rightarrow 0}]$ is the difference between the “action” $S_0[\gamma_{f^* \rightarrow x}]$ associated with the most probable escape path $\gamma_{f^* \rightarrow x}$ beginning
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

at $f^*$ and ending at an absorbing state $x$, and $\Delta S_1 \equiv \ln w[\gamma_{f^*\rightarrow 1}] - \ln w[\gamma_{f^*\rightarrow 0}]$ is the difference between fluctuations corrections to the action of the most probable escape path. The $N$-independent functions $\Delta S_0$ and $\Delta S_1$ are known analytically [31, 32] but are unnecessary for illustrating the failure of the fixed population size model.

Note that Eq. (3.1) resembles the Boltzmann weight in equilibrium statistical mechanics if $N$ is interpreted as inverse temperature while $S_0[\gamma_{f^*\rightarrow x}]$ and $\ln w[\gamma_{f^*\rightarrow x}]$ play the role of energy and entropy, respectively, as in the classical Kramers escape-over-a-barrier problem due to thermal fluctuations [25]. For $f^* > 1/2$, it is more likely for species 1 to be fixed and we can infer from Eq. (3.1) that $\Delta S_0 > 0$, resulting in $u(f^*) \to 1$ as $N \to \infty$. Similar arguments give $\Delta S_0 < 0$ if $f^* < 1/2$, implying that $u(f^*) \to 0$ as $N \to \infty$.

We now denote the two free parameters of Eq. (2.33) by $\tilde{\beta}N_{\text{eff}}$ and $f_{\text{eff}}^*$, and fit $u(f)$ to our numerically simulated fixation probability for strong asymmetric mutualism with $s_o = 0$, $\beta_1 = 0.75$, and $\beta_2 = 0.70$ whose actual coexistence fixed point is $(f^*, c_T^*) \approx (0.517, 1.568)$. The initial overall population size in our simulations is taken to be $c_T = 1 < c_T^*$, i.e., the initial overall population size is less than the fixed point value. The simulated results shown in Fig. 3.1 reveal a plateau of equal fixation probability even for relatively small $N \gtrsim 12$. To match the center of the plateau, we choose $f_{\text{eff}}^* = f^*$. The other free parameter $\tilde{\beta}N_{\text{eff}}$ controls both the plateau height and width. Because the plateau structure occupies most region, it is reasonable to adjust $\tilde{\beta}N_{\text{eff}}$ so that $u(f^*)$ matches the height of the simulated plateau. With this fitting procedure, the plateau in the fixed population size model is guaranteed to agree with the simulated plateau.

Although the fixed population size model can be adjusted to fit the elongated plateau in agreement with simulations, it fails to capture the behavior near the absorbing boundaries as revealed by Fig. 3.1, where the simulated points systemat-
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

Figure 3.1: Comparison between fixed population size predictions given by Eq. (2.33) (dashed lines) and the simulated fixation probabilities (symbols) for a strong asymmetric mutualism with a varying population size with $s_0 = 0$, $\beta_1 = 0.75$, and $\beta_2 = 0.70$. The plateau-fitting procedure yields the fitting parameters $f^*_{\text{eff}} = f^* = \beta_1/(\beta_1 + \beta_2) \approx 0.517$ and $\tilde{\beta}N_{\text{eff}} = 58.0$, 74.4, 91.0 for the simulated $N = 12$, 16, 20, respectively. This procedure always fits the plateau, but misses the boundary layer behavior.

ically fall away from the predicted dashed lines. In fact, it is the boundary behavior that distinguishes the fixation probability of mutualism with a fixed population size from mutualism with a varying population size. As we now show, the elongated plateaus also exist for strong mutualism with a varying population size, but the behavior near absorbing boundaries depends on the delicate interplay between the relative frequency and the overall population size.

3.2 The Fixation Probability and the Mean Fixation Time from Matched Asymptotic Expansions

We now study the fixation probability and the mean fixation time taking into account both the frequency and the population size degrees of freedom. Our results for the fixation probability are summarized in Fig. 3.2. A fixation event with
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

Figure 3.2: The fixation probability for strong symmetric mutualism with vanishing reproductive advantage near the origin \((s_0 = 0), \beta_1 = \beta_2 = 0.75\) and \(N = 20\) from the matched asymptotic expansions. In the boundary layers adjacent to the absorbing boundaries, the fixation probability shows crossovers with the characteristic width \(1/N\) from the boundary condition values to the plateau value \(P = 1/2\). Within the boundary layers, fixation at early times before approaching the vicinity of the coexistence fixed point is likely. In contrast, initial conditions in the plateau almost surely arrive at the coexistence fixed point before rare fluctuations eventually leads to fixation. The heat map provides an alternative representation of \(u(c)\).

Initial frequency \(f\) and initial population size \(c_T\) requires a two-dimensional escape to an absorbing boundary from the initial condition \((fc_T, (1-f)c_T)\) in the \((c_1, c_2)\) coordinates. In contrast to mutualism under the replicator condition \((0 < \beta_i \ll 1\) and \(c_T \approx 1)\), there is no dimensional reduction to an effectively one-dimensional dynamics with approximately fixed \(c_T \approx 1\) here. In fact, the fixation probability \(u(c)\) obeys a two-dimensional backward Kolmogorov equation, namely

\[
0 = \sum_{i=1}^{2} \left( v_i(c) \partial_{c_i} u(c) + \frac{1}{2N} D_i(c) \partial_{c_i}^2 u(c) \right),
\]

\[ (3.2) \]
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

with the deterministic drifts \( v_i(c) \) and diffusion coefficients \( D_i(c) \) given by Eqs. (2.15)-(2.18). The absorbing boundaries corresponding to the fixation of species 1 and of species 2 impose the boundary conditions \( u(c_1,0) = 1 \) and \( u(0,c_2) = 0 \), respectively. Eq. (3.2) does not admit an exact solution, and the standard technique of escape from a potential well can not be applied since \( v_i(c) \) is not a gradient of a potential function, i.e., \( \nabla \times \vec{v}(c) = [\mu_1(1 - \beta_1)c_1 + \mu_2(1 - \beta_2)c_2](\hat{c}_1 \times \hat{c}_2) \neq \vec{0} \).

Despite these complications, given an empirical data set with a plateau structure of fixation probability \textit{a priori}, we can solve for \( u(c) \) accurate to first order in \( 1/N \) by the method of matched asymptotic expansions. The strategy is to separately find asymptotic solutions of \( u(c) \) in the plateau region and in the boundary layers adjacent to the absorbing boundaries, and then perform asymptotic matching of the local solutions.

In the plateau region, the fixation probability \( u(c) \) near the coexistence fixed point \( c^* = (c_1^*, c_2^*) \) is approximately equal to \( P \equiv u(c^*) \). Similar to strong mutualism with a fixed population size, the dynamics in the plateau can be characterized by a rapid approach to the coexistence fixed point \( c^* \) before weak fluctuations slowly drive the system toward fixations by a large deviation. Eq. (3.2) guarantees the existence of the plateau structure if number fluctuations are sufficiently weak. Indeed, in the limit \( 1/N \to 0 \), Eq. (3.2) reduces to the simple advection equation

\[
0 = \sum_{i=1}^{2} \frac{\partial}{\partial c_i} u(c).
\]

The associated characteristics \( c(t) \) obey the mean field dynamics \( dc_i(t)/dt = v_i(c) \) on which \( du(c(t))/dt = 0 \), meaning that the fixation probability along each characteristic is constant. Because all the characteristics meet at the stable fixed point \( c^* \), we conclude \( u(c) = P \equiv u(c^*) \). This plateau value, however, can not extend over the entire domain or the boundary conditions will be violated, implying the existence of boundary layers adjacent to the absorbing boundaries.
Analogous to $u(f^*)$ in the previous section, the plateau fixation probability $P \equiv u(c^*)$ is the ratio of the flux into the absorbing boundary $f = 1$ to the total flux into both of the absorbing boundaries $f = 0$ and $f = 1$, see Appendix B.2. The flux peaks at the saddle fixed point of each absorbing boundary, which suggests that these saddle fixed points dominate the most probable escape routes for each absorbing boundary. Appendix B.2 discusses the derivation of the $N$-dependence of $P$ which takes the asymptotic form similar to Eq. (3.1):

\[ P \approx \frac{1}{1 + e^{-N\Delta S_0 + \Delta S_1}}, \]

(3.3)

where $\Delta S_0$ and $\Delta S_1$ are treated as fitting parameters. For strong symmetric mutualism with $s_o = 0$ and $\beta_1 = \beta_2$ the dynamics has a reflection symmetry with respect to the line $f = 1/2$ in the $c_1$-$c_2$ plane; hence, fixation of either species is equally likely and $P = 1/2$ independent of $N$. In the symmetric case, we can thus infer $\Delta S_0 = \Delta S_1 = 0$. For strong asymmetric mutualism, we expect that, in the limit $N \gg 1$, species 1 is more likely to be fixed if $f^* > 1/2$, and hence $\Delta S_0 > 0$. This assertion is confirmed by simulations in Fig. 3.3 where we find $\Delta S_0 > 0$ for $f^* > 1/2$ from fitting to Eq. (3.3). Extrapolations of fits to Eq. (3.3) imply that if $f^* > 1/2$, then species 1 will be fixed with probability 1 in the limit $N \to \infty$.

In the boundary layers (see Fig. 3.2), $u(c)$ crosses over from the boundary condition values to the plateau value. This crossover embodies two different types of dynamics: fixation at early times without falling into the basin of attraction of $c^*$ and fixation by a large deviation after captured by $c^*$. Fixation at early times is possible if the initial condition lies within the boundary layers, whose characteristic width is $1/N \ll 1$. The characteristic widths and the behavior of crossovers can be extracted by the method of matched asymptotic expansions. The details of matched
Figure 3.3: Extrapolations of the plateau fixation probability $P$ (top) based on Eq. (3.3), and a semilog plot of the plateau mean fixation time $T_p$ (bottom) based on Eq. (3.12), by best fits to simulations with $N = 12, 14, 16, 18, 20$ in different cases of slightly asymmetric strong mutualism. Symbols are simulation results and solid lines are best fitted curves based on Eqs. (3.3) and (3.12). The coexistence fixed point lies closer to the fixation of species 1, yielding $\Delta S_0 > 0$. The plateau fixation probability $P$ increases at increasing $N$ and saturates at 1 as $N \to \infty$. The plateau mean fixation time $T_p$ grows exponentially with $N$. Asymptotic expansions are given by Appendix B.1.
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

Figure 3.4: Comparisons between the predicted fixation probability (solid lines) and simulation results (symbols). The plateau fixation probabilities $P$ are determined by fitting, Eq. (3.3), to simulations with $N = 12, 14, 16, 18, 20$ with other parameters fixed. Top: the fixation probability for $s_o = 0, \beta_1 = 0.75, \beta_2 = 0.70$ at increasing $N$, indicated by the arrow. The matched asymptotics yield excellent estimates for the crossovers from the boundary values to the plateau value $P(N)$ even for $N$ as small as 12. As $N$ increases, the plateau region becomes more elongated (i.e., the boundary layers vanish as $1/N$) while the plateau value $P(N)$ increases and eventually saturates at 1 as $N \to \infty$, similar to Fig. 3.3(top). Bottom: the fixation probability at $N = 20$ for different cases of strong mutualism. Note the improved agreement between simulation and theory compared to Fig. 3.1.

In the boundary layer adjacent to the absorbing boundary $c_2 = 0$ and away the absorbing boundary $c_1 = 0$, the asymptotic large $N$ form of the fixation probability

58
Chapter 3: Strong Mutualism with Fluctuating Population Sizes

reads

\[ u(c) = P + (1 - P)e^{-Nc_2/\Phi_1(c_1)}, \]  
(3.4)

where the function \( \Phi_1(x) \) satisfies

\[ 0 = -(1 - x)\Phi'_1(x) + (1 + s_0)^{-1}[1 - (1 - \beta_2)x]\Phi_1(x) - \frac{1}{2}(1 + s_0)^{-1}[1 + (1 - \beta_2)x], \]  
(3.5)

with the matching condition \( \lim_{x \to 1} \Phi_1(x) = \frac{2 - \beta_2}{2\beta_1} \). For a fixed \( c_1 \), Eq. (3.4) implies that the fixation probability exhibits a crossover from 1 to \( P \) as \( c_2 \) increases from \( c_2 = 0 \) to \( c_2 \gg 1/N \). The details of the crossover depend on \( 1/\Phi_1(x) \), which is a monotonically decreasing function of \( x \) for \( \beta_1 < 1 \) and \( \beta_2 < 1 \), with \( 1/\Phi_1(0) = 2 \).

In the complementary boundary layer adjacent to the absorbing boundary \( c_1 = 0 \) and away from the absorbing boundary \( c_2 = 0 \), the asymptotic form of the fixation probability reads

\[ u(c) = P - Pe^{-Nc_1/\Phi_2(c_2)}, \]  
(3.6)

where the function \( \Phi_2(x) \) satisfies

\[ 0 = -(1 - x)\Phi'_2(x) + (1 + s_0)[1 - (1 - \beta_1)x]\Phi_2(x) - \frac{1}{2}(1 + s_0)[1 + (1 - \beta_1)x], \]  
(3.7)

with the matching condition \( \lim_{x \to 1} \Phi_2(x) = \frac{2 - \beta_1}{2\beta_2} \). For a fixed \( c_2 \), Eq. (3.6) implies that the fixation probability exhibits a crossover from 0 to \( P \) as \( c_1 \) increases from \( c_1 = 0 \) to \( c_1 \gg 1/N \). Similar to \( 1/\Phi_1(x) \), \( 1/\Phi_2(x) \) is a monotonically decreasing function of \( x \) in the parameter range of interest with \( 1/\Phi_2(0) = 2 \).

The above local behaviors of fixation probability can be combined into the global asymptotic solution

\[ u(c) = P + (1 - P)e^{-Nc_2/\Phi_1(c_1)} - Pe^{-Nc_1/\Phi_2(c_2)}, \]  
(3.8)
where the functions $\Phi_1(c_1)$ and $\Phi_2(c_2)$ obey Eq. (3.5) and Eq. (3.7) with the associated matching conditions. This global asymptotic solution valid everywhere on the domain except in the small box near the origin $[0, 1/N] \times [0, 1/N]$ where the two boundary layers overlap. Upon changing the coordinate to $(f, c_T)$ to emphasize the important population size degree of freedom, we obtain, finally,

$$u(f, c_T) = P + (1 - P)e^{-N(1-f)c_T/\Phi_1(fc_T)} - Pe^{-Nfc_T/\Phi_2((1-f)c_T)}.$$  \hfill (3.9)

Fig. 3.2 summarizes the fixation probability as a function of $(c_1, c_2)$ predicted by Eq. (3.8) for a strong symmetric mutualism. Fig. 3.4 shows excellent agreement between the prediction of Eq. (3.8) and the simulation results for a strong asymmetric mutualism. As expected, the improvements relative to Fig. 3.1 occur for $f$ near 0 and 1: the boundary behavior missed by the fixed population size model are now well captured even for $N$ as small as 12.

The mean fixation time, $\tau(c)$, can also be constructed by the method of matched asymptotic expansions. In this case, we need to solve $[24, 26]

$$-1 = \sum_{i=1}^{2} \left( v_i(c) \partial_c \tau(c) + \frac{1}{2N} D_i(c) \partial^2_c \tau(c) \right),$$  \hfill (3.10)

subject to the boundary conditions $\tau(c_1, 0) = 0$ and $\tau(0, c_2) = 0$. Asymptotic matching arguments similar to Appendix B.1 can be applied to Eq. (3.10), resulting in the global asymptotic solution for the mean fixation time, namely

$$\tau(c) = T_p \left( 1 - e^{-Nc_2/\Phi_1(c_1)} - e^{-Nc_1/\Phi_2(c_2)} \right),$$  \hfill (3.11)

where the functions $\Phi_1(c_1)$ and $\Phi_2(c_2)$ still obey Eq. (3.5) and Eq. (3.7), and $T_p$ is the plateau mean fixation time for the initial condition at the coexistence fixed
Figure 3.5: Comparisons between the predicted mean fixation time (solid lines) and simulation results (symbols). Parameters are the same as Fig. 3.4’s. The plateau mean fixation times are determined by fitting, Eq. (3.12), to simulations with \( N = 12, 14, 16, 18, 20 \) with other parameters fixed. Top: the mean fixation time for \( s_o = 0, \beta_1 = 0.75, \beta_2 = 0.70 \) at increasing \( N \). The matched asymptotic expansions yield excellent estimates for the crossovers from the boundary values to the plateau value \( T_p(N) \) even for \( N \) as small as 12. As \( N \) increases, the plateau region becomes more elongated (i.e., the boundary layers vanish as \( 1/N \)) while the plateau value \( T_p(N) \) grows exponentially similar to Fig. 3.3(bottom). Bottom: the mean fixation time at \( N = 20 \) for different cases of strong mutualism.

The function \( \tau(e) \) possesses a plateau structure in which \( \tau(e) \approx \tau(e^*) = T_p \), similar to the profile of \( u(e) \). Furthermore, crossovers from the boundary conditions to \( T_p \) are characterized by the same exponentials \( e^{-Nc_2/\Phi_1(c_1)} \) and \( e^{-Nc_1/\Phi_2(c_2)} \) as in
Eq. (3.8). In the limit $N \gg 1$, the behavior of $T_p$ is exponential in $N$,

$$T_p \approx \sigma_0 e^{N\sigma_1},$$

(3.12)

where we treat $\sigma_0$ and $\sigma_1$ as fitting parameters [96–98]. Fig. 3.3(bottom) confirms the exponential scaling of Eq. (3.12), while Fig. 3.5 reveals excellent agreement between Eq. (3.11) and the simulation results. Again, to emphasize the importance of the population size degree of freedom, we rewrite Eq. (3.11) in the coordinates $(f, c_T)$ as

$$\tau(f, c_T) = T_p\left(1 - e^{-N(1-f)c_T/\Phi_1(fc_T)} - e^{-Nfc_T/\Phi_2((1-f)c_T)}\right).$$

(3.13)

### 3.3 Conclusions and Outlooks

In chapter 2 and 3, we have explored the interplay between evolutionary dynamics and population dynamics in a well-mixed competitive Lotka-Volterra model in various limits. The model gives rise to 5 different scenarios, similar to evolutionary game theory, without however fixing the overall population size, thereby demonstrating an explicit microscopic system exhibiting the feedback between evolutionary dynamics and population dynamics phenomenologically studied in Refs. [99, 100].

The limit $|\beta_1| \ll 1$, $|\beta_2| \ll 1$, and $1/N \ll 1$, with an arbitrary reproductive advantage near the origin $s_o$, describes rapid relaxational dynamics of population size toward a fixed equilibrium size along a quasi-deterministic growth trajectory on which $\rho \equiv (1-f)f^{-1/(1+s_o)}c_T^{s_o/(1+s_o)}$ is constant. The variable $\rho$ relates the population frequency $f$ to the total population size $c_T$ as $c_T$ approaches the equilibrium at $c_T \approx 1$: The frequency of a reproductively advantageous species, on average, increases as dilute populations ($c_T < 1$) grow, and decreases as overcrowded popula-
tions \((c_T > 1)\) decline. For \(s_o = 0\), replicator dynamics with genetic drift is recovered when \(c_T \approx 1\), despite population size fluctuations away from \(c_T = 1\). Only in this limit is the dynamics near the equilibrium population size a simple generalization of conventional population genetics without mutation with an independently fluctuating population size. From the perspective of equilibrium statistical mechanics, this simple limit is analogous to the generalization from a canonical ensemble to a grand canonical ensemble in the thermodynamic limit. However, for \(s_o \neq 0\), population size fluctuations couple to evolutionary dynamics in a nontrivial fashion and replicator dynamics with genetic drift is no longer an appropriate description.

For a quasi-neutral evolution \((s_o \neq 0 \text{ and } \beta_1 = \beta_2 = 0)\), the effective evolutionary dynamics near a fixed equilibrium population size is deduced by means of adiabatic elimination of a fast variable, which leads to a fluctuation-induced selective advantage and unusual genetic drift of a non-Wright-Fisher (or non-Moran) type. We find that the fixation probability and the mean fixation time are universal functions of \(\rho\) due to a separation of dynamical timescales between a fast variable \(c_T\) and a slow variable \(\rho\). This separation allows the fate of competitions at an arbitrary initial population size to be deduced. Given a fixed initial relative frequency \(f\), a better strategy for a species with \(s_o > 0\) to ultimately fix is to begin in dilute populations \((c_T < 1)\), rather than overcrowded populations \((c_T > 1)\). These results demonstrate circumstances such that fixed effective population size models in population genetics are incomplete. The dynamics when \(s_o \neq 0\) illustrates the strong effect that population size fluctuations can have on the standard evolutionary game theory and on conventional population genetics more generally.

It would be interesting to study how population size fluctuations affect evolutionary dynamics near a fixed equilibrium population size in other selection scenarios when \(s_o \neq 0\). Particularly interesting is the Prisoner’s Dilemma briefly discussed at
the end of Appendix A.1. In this case, fluctuation-induced selection can actually oppose the usual selection bias in the Prisoner’s Dilemma, illustrating a fluctuation-driven mechanism other than genetic drift (or spatial segregation [101]) that can alleviate the dilemma of cooperation [30].

We also studied competitions that take place with a strongly varying population size (as opposed to competition with a nearly fixed population size), as in the strong mutualism limit. Fixation events can now arise via two distinct mechanisms: fixation at long times by rare escape from the strongly attractive coexistence fixed point, and fixation at early times before reaching the neighborhood of the coexistence fixed point. The former situation is typical for initial conditions away from the absorbing boundaries \((c_1, 0)\) and \((0, c_2)\), where the system initially falls toward the coexistence fixed point, resulting in a plateau of constant fixation probability and constant mean fixation time. The latter situation arises for initial conditions lying close to the absorbing boundaries where fluctuations can fix one species before falling into the coexistence fixed point. The crossovers from the absorbing boundaries to the plateau in the fixation probability and the mean fixation time can be studied by matched asymptotic expansions, accounting for number fluctuations, evolutionary dynamics, and population dynamics. As shown in Fig. 3.1, the fixed population size model underestimates the number of fixation events that can occur near the absorbing boundaries, thereby overestimating the probability and duration of species coexistence. This dynamics can be important in the context of range expansions of mutualists [3, 35] where populations at the expanding frontier are continuously subject to interaction in a growing population size, which may alter parameter values separating an active (mutualistic) phase from inactive (single species domination) phase.

It would be interesting to explore the plateau fixation probability and the mean
fixation time without resorting to fitting parameters that amount to parameterizing an optimal path toward fixation (instanton solution) [32, 96, 97, 102]. The fate of competitions as a function of both population size and the frequency in other competition scenarios with a varying population size under strong selection would also be worth investigating.
Chapter 4

Striated Populations in Disordered Environments with Advection

Growth in static and controlled environments such as a Petri dish can be used to study the spatial population dynamics of microorganisms. However, natural populations such as marine microbes experience fluid advection and often grow up in heterogeneous environments. In this chapter, we investigate a generalized Fisher-Kolmogorov-Petrovsky-Piscounov (FKPP) equation (1.22) describing single species population subject to a constant flow field and quenched random spatially inhomogeneous growth rates with a fertile overall growth condition. We analytically and numerically demonstrate that the non-equilibrium steady-state population density of Eq. (1.26) develops a flow-driven striation pattern. The striations are highly asymmetric with a longitudinal correlation length that diverges linearly with the flow speed and a transverse correlation length that approaches a finite velocity-independent value. Linear response theory is developed to study the statistics of the steady states. Theoretical predictions show excellent agreement with the numerical steady states of the generalized FKPP equation obtained from Lattice Boltzmann simulations. These findings suggest that, although the growth disorder can be spatially uncorrelated, correlated population structures with striations emerge naturally at sufficiently strong advection.

In Sec. 4.1, we show that the steady-state density fluctuations from the mean carrying capacity \( \bar{K} \equiv a_0/b \) can be regarded as a linear response to a small per-
Chapter 4: Striated Populations in Disordered Environments with Advection

turbation caused by the quenched random background growth rate. The Green’s function of the associated reaction-diffusion-advection operator, which we calculate in $d$ dimensions, describes the steady-state density response to a point-like growth hot spot. In the absence of advection, the response decays isotropically, with the diffusive localization length $\xi_D \equiv \sqrt{D/a_0}$, from the center of a growth hot spot. We then discuss how advection breaks isotropy: advection elongates (shortens) the downstream (upstream) longitudinal localization length, while symmetrically contracting the transverse localization length. Figs. 4.2 and 4.3 illustrate this effect, in $d = 1$ and $d = 2$, respectively. For strong advection such that $\tilde{v} \equiv v/v_F \gg 1$ where $v_F \equiv 2\sqrt{Da_0}$ is the characteristic Fisher speed associated with the average growth rate [45, 75], the downstream and upstream longitudinal localization lengths scale as $\xi_{\parallel}^+ \approx \tilde{v}\xi_D = v/2a_0$ and $\xi_{\parallel}^- \approx (1/\tilde{v})\xi_D = 2D/v$, while the transverse localization length scales as $\xi_{\perp} \approx (1/\tilde{v})\xi_D$. The asymmetric elongation and contraction occur simultaneously with the decay in the response amplitude, a consequence of the conservation of density fluctuations discussed at the end of Sec. 4.2.

In Sec. 4.3, we show that advection and spatially quenched uncorrelated random growth rates together lead to striated patterns of steady-state population density in two dimensions, as depicted in Fig. 4.1. Although the growth rate fluctuations are spatially uncorrelated and the transverse localization length of the response from an isolated growth hot spot shrinks to zero as advection becomes stronger, a finite transverse correlation length of the steady state density nevertheless emerges at strong advection. Theoretical analysis of the two-point correlation function leading to striated population structures is provided in Sec. 4.4. There, we also show that, for strong advection $v \gg v_F$, the longitudinal correlation length elongates without bound as $\xi_{\parallel} = v/a_0$ whereas the transverse correlation length approaches the limiting value $\xi_{\perp} = \xi_D = \sqrt{D/a_0}$. Consequently, the population structure becomes highly
Chapter 4: Striated Populations in Disordered Environments with Advection

Figure 4.1: Typical long-time steady states $c^*(x)$ with striations in two dimensions that arise from the generalized FKPP equation with periodic boundary conditions and variable strength advection in the $+\hat{x}_\parallel$ direction and with the same realization of a weak spatially quenched random growth rate. The black lines at the bottom indicate contours of $c^*(x)$ at the value of the mean carrying capacity $\bar{K} = 10^4$, separating the more favourable growth domains from the less favourable ones.

These plots result from the lattice Boltzmann simulation discussed in Sec. 4.5. Without advection, as in (a), steady state density exhibit mild fluctuations around the mean carrying capacity $\bar{K} = 10^4$. The correlations of the density fluctuations without advection are isotropic. However, advection breaks statistical isotropy of the steady state; correlation length is developed and is elongated in the longitudinal direction, as in (b), (c) and (d). Striated patterns emerge at strong flow $v \gg v_F$ with the longitudinal correlation length of order the system size while the transverse correlation length remains finite, see (d). Note also that strong advection lowers the amplitude of the density fluctuations.

anisotropic, with the ratio of correlation lengths given by $\xi_\parallel / \xi_\perp = 2\tilde{v}$. Sec. 4.5 compares the long distance correlations from theoretical analysis with those from the lattice Boltzmann simulations. Concluding remarks appear in Sec. 4.6, and
detailed calculations of Green’s functions, correlations and structure functions are contained in Appendix C.1.

4.1 Steady-State Density From The Linear Response Approximation

To determine how weak spatial perturbations in the growth rate alter a steady-state density that is otherwise homogeneous and equal to the mean carrying capacity $\bar{K} = a_0/b$, we introduce the density deviation from the steady state, normalized by the mean carrying capacity, as follows:

$$\phi(x) \equiv \frac{c^*(x) - \bar{K}}{\bar{K}}.$$  \hfill (4.1)

henceforth referred to as a (static) density fluctuation. Upon substituting the steady-state density of Eq. (4.1) into Eq. (1.26), we find that this (static) density fluctuation satisfies

$$(-D\nabla^2 + \mathbf{v} \cdot \nabla + a_0) \phi(x) = \delta a(x) + [\delta a(x)\phi(x) - a_0\phi^2(x)].$$  \hfill (4.2)

We now establish the linear response theory for the case of weak spatial fluctuations in the growth rate. First, we define $\varepsilon \equiv (\Delta/\Delta_0) \geq |\delta a(x)/a_0|$, a dimensionless measure of growth rate fluctuations. Observe that, in the absence of both advection and diffusion, non-zero populations at any point $x$ will grow and saturate according to the local logistic growth process; the steady state density is then given by the local carrying capacity:

$$c^*(x) = K(x) \equiv a(x)/b.$$  \hfill (4.3)
Eq. (4.1), Eq. (4.3), and the bound associated with uniform distribution \( \delta a(x) \in [-\Delta, \Delta] \), implies that the density fluctuations obey \( |\phi(x)| \lesssim \varepsilon \). In the presence of either diffusion or advection, populations traverse longer distances and sample a spatial average of local growth rates; the steady state density is then smoothed out, and the condition \( |\phi(x)| \lesssim \varepsilon \) should remain approximately valid. In fact, when diffusion or advection become strong, numerical simulations reveal the suppression of density fluctuations, as shown for strong advection in Figs. 4.1 and 4.6 in two dimensions and one dimension, respectively (the case of suppression by strong diffusion is similar.) Upon dividing Eq. (4.2) by \( a_0 \) and noting that \(|\phi(x)| = O(\varepsilon)\), we see that the terms in the square bracket scale as \( \varepsilon^2 \), whereas the other terms scale as \( \varepsilon \). Hence, in the limit of small \( \varepsilon \), we can linearize Eq. (4.2):

\[
(-D \nabla^2 + \mathbf{v} \cdot \nabla + a_0) \phi(x) = \delta a(x).
\]

(4.4)

Thus, in this linear approximation (used throughout this chapter), static density fluctuations are generated in response to the growth disorder \( \delta a(x) \) acting as a source term. As we show later, Eq. (4.4) becomes a better and better approximation upon increasing either the \( D \) or \( |\mathbf{v}| \).

It is convenient to introduce the non-dimensionalized linear response equation to simplify further calculations. Upon defining the diffusion length in a growth time

\[
\xi_D \equiv \sqrt{D/a_0},
\]

(4.5)
and rescaled quantities

\[ \tilde{x} \equiv x / \xi_D, \quad (4.6) \]
\[ \tilde{v} \equiv v / 2\sqrt{D} a_0 = v / v_F, \quad (4.7) \]
\[ U(x) \equiv \delta a(x) / a_0, \quad (4.8) \]

Eq. (4.4) takes the dimensionless form

\[ \left( -\tilde{\nabla}^2 + 2\tilde{v} \cdot \tilde{\nabla} + 1 \right) \phi(\tilde{x}) = U(\tilde{x}), \quad (4.9) \]

where \( \tilde{\nabla} \) denotes a gradient with respect to \( \tilde{x} \). The density fluctuations are then given by the convolution

\[ \phi(\tilde{x}) = \int G(\tilde{x} - \tilde{y}) U(\tilde{y}) \, d^d \tilde{y}. \quad (4.10) \]

where \( G(\tilde{x} - \tilde{y}) \) is a reaction-diffusion-advection Green’s function that satisfies

\[ \left( -\tilde{\nabla}^2 + 2\tilde{v} \cdot \tilde{\nabla} + 1 \right) G(\tilde{x} - \tilde{y}) \propto \delta^d(\tilde{x} - \tilde{y}). \quad (4.11) \]

The steady-state population density, corrected for diffusion and advection, then reads

\[ c^*(\tilde{x}) = \bar{K} \left[ 1 + \int G(\tilde{x} - \tilde{y}) U(\tilde{y}) \, d^d \tilde{y} \right]. \quad (4.12) \]

And we must now determine \( G(\tilde{x} - \tilde{y}) \).
4.2 Density Modulation Due to a Point-like Growth Hot Spot

The Green’s function of the reaction-diffusion-advection operator of Eq. (4.9) describes the response to a Dirac delta function source term, a point-like growth hot spot. In our rescaled coordinates, the Green’s function satisfies

\[
\left( -\nabla^2 + 2\tilde{v} \cdot \nabla + 1 \right) G(\tilde{x}) = \frac{\Delta}{a_0} \delta^d(\tilde{x}),
\]

which can be simplified via the substitution

\[
f_{\tilde{v}}(\tilde{x}) \equiv \exp(-\tilde{v} x_\parallel) f(\tilde{x}),
\]

where \(x_\parallel\) is the direction along the advective flow. The result is a Helmholtz equation with the minus sign in the Laplacian:

\[
\left[ -\nabla^2 + (1 + \tilde{v}^2) \right] G_{\tilde{v}}(\tilde{x}) = \frac{\Delta}{a_0} \delta^d(\tilde{x}).
\]

In \(d\) dimensions, the isotropic solution of Eq. (4.15), such that \(G_{\tilde{v}}(\tilde{x})\) vanishes as \(\tilde{x} \to \infty\), reads [103]

\[
G_{\tilde{v}}(\tilde{x}) = \frac{\Delta}{a_0} \left( \frac{1}{\pi} \right)^{d/2} \times \left[ \left( \frac{|\tilde{x}|}{\sqrt{1 + \tilde{v}^2}} \right)^{1-d/2} K_{1-d/2} \left( \sqrt{1 + \tilde{v}^2} |\tilde{x}| \right) \right].
\]
so that

\[ G(\tilde{x}) = \exp(\tilde{v}\tilde{x}_\parallel)G_\tilde{v}(\tilde{x}). \]  

(4.17)

Eqs. (4.16) and (4.17) are derived via direct Fourier transformation in Appendix C.1.

Eq. (4.17) encapsulates how advection breaks isotropy: by enhancing downstream response and suppressing upstream response in Eq. (4.16). Since \( K_\alpha(x) = \sqrt{\frac{\pi}{2x}} \exp(-x) \left[ 1 + \frac{4\alpha^2-1}{8x} + O(x^{-2}) \right] \) for large \( x \) [23], Eq.(4.17) implies that \( G(\tilde{x}) \) is exponentially localized in all transverse directions with the transverse localization length \( \tilde{\xi}_\perp = 1/\sqrt{1+\tilde{v}^2} \) similar to the localization length of \( G_\tilde{v}(\tilde{x}) \), i.e.,

\[ G(\tilde{x}_\parallel = 0, \tilde{x}_\perp) \sim \exp \left( -\sqrt{1+\tilde{v}^2}|\tilde{x}_\perp| \right). \]  

(4.18)

For \( \tilde{v} \gg 1 \), \( \tilde{\xi}_\perp \) contracts as \( 1/\tilde{v} \). However, the longitudinal localization lengths are asymmetric with an elongation in the downstream direction and a contraction in the upstream direction. For \( \tilde{v} \gg 1 \),

\[ G(\tilde{x}_\parallel, \tilde{x}_\perp) = 0 \sim \exp(\tilde{v}\tilde{x}_\parallel)\exp(-\sqrt{1+\tilde{v}^2}|\tilde{x}_\parallel|) \]

\[ = \exp \left[ (\tilde{x}_\parallel - |\tilde{x}_\parallel|)\tilde{v} - \frac{|\tilde{x}_\parallel|}{2\tilde{v}} + O \left( \frac{1}{\tilde{v}^3} \right) \right]; \]  

(4.19)

thus, the downstream \( (\tilde{x}_\parallel > 0) \) and the upstream \( (\tilde{x}_\parallel < 0) \) localization lengths are given by \( \tilde{\xi}_\parallel^+ \approx 2\tilde{v} \) and \( \tilde{\xi}_\parallel^- \approx 1/2\tilde{v} \), respectively. This steady state asymmetry, with a contraction in the transverse direction, differs from the early time growth dynamics, where diffusion spreads out the population superdiffusively in the transverse direction [61, 62, 65].

73
Figure 4.2: One-dimensional Green’s functions given by Eq. (4.21) which describe the density fluctuations as a response to a point-like growth hot spot, introduced at the origin and indicated as the dashed green line. The response becomes strongly asymmetric as \( \tilde{v} \) increases, with the downstream localization length and the upstream localization length that scale respectively as \( \tilde{\xi}_+ = 2\tilde{v} \) and \( \tilde{\xi}_- = 1/2\tilde{v} \) for \( \tilde{v} \gg 1 \). Note that the response amplitude decays to compensate for the elongated downstream propagation. In this plot, the strength of the hot spot \( \Delta / a_0 \) is set to unity.

The steady-state response to advection is, however, constrained by a conservation law. By integrating Eq. (4.13) over the whole domain with a periodic boundary condition, one finds that

\[
\int_{\Omega} G(\tilde{x})d^{d}\tilde{x} = \frac{\Delta}{a_0}.
\]

As a result, the overall response is suppressed at strong advection: the response amplitude decreases to compensate for elongated downstream amplitude. Suppres-
Chapter 4: Striated Populations in Disordered Environments with Advection

Figure 4.3: Two-dimensional Green’s function of Eq.(4.22). The response to a point-like growth hot spot is isotropic in the absence of flow, as represented in (a), but streaked out in the direction parallel to the flow with an elongated downstream localization length, as shown in (b)-(d). With increasing flow speed, the overall response amplitude decays. In addition, the transverse response, controlled by \( G(\tilde{x}_\parallel, \tilde{x}_\perp) \sim K_0(\sqrt{1+\tilde{v}^2|\tilde{x}|}) \), contracts symmetrically, decaying exponentially with the localization length \( \xi = 1/\sqrt{1+\tilde{v}^2} \). Similar to one dimension, \( \xi^+ = 2\tilde{v} \) and \( \xi^- = 1/2\tilde{v} \) in the flow direction for \( \tilde{v} \gg 1 \). Solid lines shown on the plane \( \tilde{x}_\parallel = -2.5 \) and \( \tilde{x}_\perp = 2.5 \) are cross-sections of the response along the plane \( \tilde{x}_\parallel = 0 \) and the plane \( \tilde{x}_\perp = 0 \), respectively. These cross-sections reveal the behavior of localization lengths at different velocities. The boundaries of the colored contour above each plot correspond to the Green’s function contours 0.0, 0.05, 0.1, 0.15, 0.2, 0.25 and 0.3, with the color ranging from yellow to dark green, respectively. The reduction in the area of these colored contours at stronger advection reflects the decay of the response amplitude. In this plot, the amplitude of the \( \delta \)-function hot spot \( \Delta/a_0 \) is set to 1.

The shrinkage of the response amplitude as a tradeoff for elongated downstream propagation is readily verified in one-dimension. In this case, the modified Bessel function takes
Chapter 4: Striated Populations in Disordered Environments with Advection

the simple form $K_{1/2}(x) = \sqrt{\frac{x}{2\pi}} \exp(-x)$ [23], and Eqs. (4.16)-(4.17) give

$$G(\tilde{x}) = \frac{\Delta}{a_0} \frac{1}{\sqrt{2}} \frac{\exp(\tilde{v}\tilde{x} - \sqrt{1 + \tilde{v}^2} |\tilde{x}|)}{\sqrt{1 + \tilde{v}^2}}.$$ (4.21)

Hence, advection leads to anisotropic response whose amplitude at $\tilde{x} = 0$ decays as $1/\sqrt{1 + \tilde{v}^2}$. Fig. 4.2 shows $G(\tilde{x})$ in one dimension for a variety of velocities. In two dimensions, Eqs. (4.16)-(4.17) give

$$G(\tilde{x}_\parallel, \tilde{x}_\perp) = \frac{\Delta}{a_0} \left( \frac{1}{\pi} \right) \exp(\tilde{v}\tilde{x}_\parallel) K_0 \left( \sqrt{1 + \tilde{v}^2} |\tilde{x}_\parallel| \right).$$ (4.22)

Although there is a logarithmic singularity at $|\tilde{x}| = 0$ (cut off by, say, the spacing between microorganisms), one can see from the profiles of the plots of Eq. (4.22) in Fig. 4.3 that the overall response amplitude is again suppressed at strong advection. Fig. 4.3 also illustrates the process of symmetric transverse contraction, in addition to the asymmetric downstream elongation and upstream contraction in the longitudinal direction.

4.3 Longitudinal Striations in the Steady State for Strong Advection for $d \geq 2$

Although the linear response theory developed in Sec. 4.1 suggests that the transverse localization length of an isolated growth hot spot shrinks to zero as the advection speed increases, we now show that superposition of the responses from uncorrelated growth disorder leads to correlations whose transverse correlation length approaches a finite velocity-independent value, while the longitudinal correlation length grows linearly with $\tilde{v}$. The result is striated population correlations in the
steady state.

Figure 4.4: How striations driven by flow and spatially quenched growth disorder arise in two dimensions. (Left) Without advection, density fluctuations arise from the superposition of the isotropic responses $G_{v=0}(x - x')$ centered around a source ($U(x') > 0$) or a sink ($U(x') < 0$), depicted in green and red respectively. The localization length is given by $\xi_D = \sqrt{D/a_0}$. (Right) For strong advection $\tilde{v} \gg 1$, each response is distorted asymmetrically: the downstream localization length grows as $\xi^+ = 2\tilde{v}\xi_D$, and the transverse localization length contracts as $\xi^- = \xi_D/\tilde{v}$. Although each response is subjected to transverse contraction that shrinks to zero at stronger flow, the random superposition of responses exhibit non-vanishing transverse correlations, as shown in Figs. 4.5 and 4.8.

4.4 Two-point Correlation Function and the Striation Pattern

To model rapid, statistically isotropic spatial variations in the growth rates, we assume each local growth rate $U(\tilde{x}) = \delta a(\tilde{x})/a_0$ is drawn from a uniform box distribution in the interval $[-\Delta/a_0, \Delta/a_0]$, which gives the two-point noise correlation function

$$\langle U(\tilde{x})U(\tilde{x}') \rangle = \frac{1}{3} \frac{\Delta^2}{a_0^2} \delta^d(\tilde{x} - \tilde{x}'),$$  \hspace{1cm} (4.23)
where \( \langle \cdot \rangle \) denotes ensemble averages over disordered growth rate realizations. In the Fourier domain, Eq. (4.23) gives white noise with magnitude

\[
\langle |U(\tilde{q})|^2 \rangle = \frac{1}{3} \frac{\Delta^2}{a_0^2}.
\]  

(4.24)

We expect statistical translational invariance,

\[
\langle \phi(\tilde{x})\phi(0) \rangle \equiv \langle \phi(\tilde{x} + \tilde{x}')\phi(\tilde{x}') \rangle,
\]

so the two-point correlation function of the density fluctuations is given by

\[
\langle \phi(\tilde{x})\phi(0) \rangle = \int \frac{d^d\tilde{q}}{(2\pi)^d} e^{i\tilde{q}\cdot\tilde{x}} S(\tilde{q}),
\]  

(4.25)

where the static structure factor in the steady state is related to the Fourier transformed Green’s function \( G(\tilde{q}) \),

\[
S(\tilde{q}) \equiv \langle |\phi(\tilde{q})|^2 \rangle = \langle |U(\tilde{q})|^2 |G(\tilde{q})|^2 \rangle = \frac{\langle |U(\tilde{q})|^2 \rangle \langle |G(\tilde{q})|^2 \rangle}{(\tilde{q}^2 + 1)^2 + 4\tilde{v}^2 \tilde{q}_\parallel^2}.
\]  

(4.26)

Upon substituting Eqs. (4.26) and (4.24) into Eq. (4.25), the two-point correlation function of the density fluctuations with spatially uncorrelated random growth rates reads

\[
\langle \phi(\tilde{x})\phi(0) \rangle = \frac{\Delta^2}{3a_0^2} \int \frac{d^d\tilde{q}}{(2\pi)^d} \frac{e^{i\tilde{q}\cdot\tilde{x}}}{(\tilde{q}^2 + 1)^2 + 4\tilde{v}^2 \tilde{q}_\parallel^2}.
\]  

(4.27)

We now evaluate Eq. (4.27) in various dimensions, with details relegated to Appendix C.1. For \( d = 1 \), Eq. (4.27) can be evaluated via contour integration, with
the result (setting $\tilde{x}_\parallel = \tilde{x}$)

$$
\langle \phi(\tilde{x})\phi(0) \rangle = \frac{\Delta^2}{24a_0^2 \tilde{v}\sqrt{1+\tilde{v}^2}} \times \left[ \frac{1}{g_-(\tilde{v})}e^{-g_-(\tilde{v})|\tilde{x}|} - \frac{1}{g_+(\tilde{v})}e^{-g_+(\tilde{v})|\tilde{x}|} \right],
$$

(4.28)

where the two exponential decays are controlled by $g_{\pm}(\tilde{v}) \equiv \sqrt{1+\tilde{v}^2} \pm \tilde{v}$. Although this result appears singular at $\tilde{v} = 0$, the limit $\tilde{v} \to 0$ is in fact well-defined, and given by

$$
\lim_{\tilde{v} \to 0} \langle \phi(\tilde{x})\phi(0) \rangle = \frac{\Delta^2}{12a_0^2} (1 + |\tilde{x}|) e^{-|\tilde{x}|}.
$$

(4.29)

The exponential localization associated with a single hot spot is broadened by a factor $(1 + |\tilde{x}|)$; the correlation length, however, is the same as the diffusive localization length $\xi_D$. For strong advection, $g_-(\tilde{v}) = 2\tilde{v} + 1/2\tilde{v} + O(1/\tilde{v}^3)$ and $g_+(\tilde{v}) = 1/2\tilde{v} + O(1/\tilde{v}^3)$; the first term in the square bracket of Eq. (4.28) dominates, resulting in

$$
\lim_{\tilde{v} \to \infty} \langle \phi(\tilde{x}_\parallel)\phi(0) \rangle \sim \frac{e^{-|\tilde{x}_\parallel|/2\tilde{v}}}{\tilde{v}}.
$$

(4.30)

For $d = 2$, the spatial structure of steady state is embodied in the longitudinal correlation function $\langle \phi(\tilde{x}_\parallel,0)\phi(0) \rangle$ and the transverse correlation function $\langle \phi(0,\tilde{x}_\perp)\phi(0) \rangle$ that are defined by Eq. (4.27). In the absence of advection, the correlation function is isotropic and is given by

$$
\lim_{\tilde{v} \to 0} \langle \phi(\tilde{x})\phi(0) \rangle \sim |\tilde{x}|K_1(|\tilde{x}|),
$$

(4.31)

where $K_1(x)$ is the modified Bessel function of a second kind that decays exponentially at large distance $x$ as $K_1(x) = \sqrt{\frac{\pi}{2x}} \exp(-x) [1 + 3/8x + O(x^{-2})]$ for $x \gg 1$ [23]. The correlation length is thus given by diffusive correlation length $\xi_D$ in this
Chapter 4: Striated Populations in Disordered Environments with Advection

In the strong advection limit, however, correlations become highly anisotropic:

\[
\lim_{\tilde{v} \to \infty} \langle \phi(\tilde{x}_\parallel, 0) \phi(0) \rangle \sim \left( \frac{|\tilde{x}_\parallel|}{2\tilde{v}} \right)^{1/4} K^{-1/4} \left( \frac{|\tilde{x}_\parallel|}{2\tilde{v}} \right), \tag{4.32}
\]

\[
\lim_{\tilde{v} \to \infty} \langle \phi(0, \tilde{x}_\perp) \phi(0) \rangle \sim e^{-|\tilde{x}_\perp|/\tilde{v}}, \tag{4.33}
\]

with the longitudinal and transverse correlation lengths that behave as follows,

\[
\lim_{\tilde{v} \to \infty} \xi_\parallel = \xi_v = \frac{v}{a_0}, \tag{4.34}
\]

\[
\lim_{\tilde{v} \to \infty} \xi_\perp = \xi_D = \sqrt{\frac{D}{a_0}}. \tag{4.35}
\]

Note the limiting transverse correlation length is given by a velocity-independent, diffusive localization length. However, the ratio of transverse to longitudinal correlation lengths becomes highly anisotropic:

\[
\lim_{\tilde{v} \to \infty} \left( \frac{\xi_\parallel}{\xi_\perp} \right) = \frac{v}{\sqrt{D}a_0} = 2\tilde{v}. \tag{4.36}
\]

These anisotropic patterns might arise in natural marine microbial populations, their Fisher wave spreading velocity is much smaller than oceanic flow speed; for a motile bacteria, a typical diffusion constant \(D\) is of order \(10^{-5} \text{ cm}^2 \text{ s}^{-1}\) and the typical doubling time \(a_0\) is of order \(10^{-3} \text{ s}^{-1}\) which gives \(v_F = 2\sqrt{Da_0}\) of order 1 \(\mu\text{m} \text{ s}^{-1}\) [75], which is indeed small compared to a typical oceanic current; see Ref. [71] and references therein.

Note that uncorrelated disorder (viewed as a superposition of responses from growth hot spots) leads to a non-zero transverse correlation length, as shown in Eq. (4.35) for two dimensions, in contrast to the result for a single point-like growth hot-spot of Sec. 4.2. In fact, a finite transverse correlation arises for all \(d \geq 2\). To
see this, consider the transverse correlation functions:

\[
\langle \phi(0, \tilde{x}_\perp) \phi(0) \rangle \sim \int \frac{d^{d-1}q_\perp e^{i\tilde{q}_\perp \cdot \tilde{x}_\perp}}{(2\pi)^{d-1}} \int \frac{d|q||}{2\pi} \frac{1}{(q^2 + 1)^2 + 4\tilde{v}^2 q^2_{\|}}
\]

\[
= \pi \int \frac{d^{d-1}q_\perp e^{i\tilde{q}_\perp \cdot \tilde{x}_\perp}}{(2\pi)^{d} (q^2_{\perp} + 1)(q^2_{\perp} + 1 + \tilde{v}^2)^{1/2}},
\]

(4.37)

Upon taking the limit \( \tilde{v} \gg 1 \), it follows that

\[
\lim_{\tilde{v} \to \infty} \langle \phi(0, \tilde{x}_\perp) \phi(0) \rangle \sim \frac{\pi}{\tilde{v}} \int \frac{d^{d-1}q_\perp e^{i\tilde{q}_\perp \cdot \tilde{x}_\perp}}{(2\pi)^{d} (q^2_{\perp} + 1)},
\]

(4.38)

\[
= e^{-|\tilde{x}_\perp|} \frac{1}{4\tilde{v}},
\]

(4.39)

which show that the transverse correlation length still obeys Eq. (4.35) for \( d \geq 2 \). Thus, the limiting transverse correlations in higher dimensions remains velocity-independent and is characterized by the diffusive localization length \( \xi_D \), while the amplitude of fluctuations decays at increasing advection as \( \lim_{\tilde{v} \to \infty} \langle \phi^2(0) \rangle \sim 1/\tilde{v} \).

### 4.5 Comparisons to Numerical Simulations

We now check theoretical predictions in Sec. 4.4.4 by comparing the theoretical structure factor \( S_{th}(\tilde{q}) \) given by Eq. (4.26) to the structure factor \( S_{num}(\tilde{q}) \) calculated from the numerical solution of Eq. (1.26). To obtain the numerical steady state satisfying Eq. (1.26), we inoculate the simulation domain with a uniform concentration equal to the mean carrying capacity \( \bar{K} \), and numerically evolve the time-dependent advective FKPP Eq. (1.22) until the steady state is reached using the Lattice Boltzmann method, a robust lattice discretization scheme for advective and diffusive transports [104] that can incorporate reactive agents [105]. This
method achieves high numerical accuracy for reaction-diffusion-advection problems, even in the strong advection limit [106]. Random growth rate \( U(\tilde{x}_i) = \delta a(\tilde{x}_i)/a_0 \) is introduced on each lattice site \( \tilde{x}_i \), and is independently drawn from a uniform box distribution in the interval \( [-\Delta/a_0, \Delta/a_0] \). The 9-speed 2-dimensional (D2Q9) lattice is adopted to evolve the density field according to Eq. (1.22) on a two-dimensional square lattice with \( N_x \times N_y \) sites and a periodic boundary condition implemented by adding the buffer sites at the boundaries [105]. In this scheme, time evolution on a one-dimensional lattice with a periodic boundary condition also follows immediately, provided \( N_x \) is set to 1. Once the numerical steady state \( c^*_{\text{num}}(x) \) is reached, one can calculate the squared modulus of the discrete Fourier transform of the steady state density fluctuations, denoted by \( |\phi_{\text{num}}(\tilde{q})|^2 \). After taking ensemble averages over the random growth rates, this results in the numerical structure factor \( S_{\text{num}}(\tilde{q}) \). This numerical structure factor is the discrete counterpart of the theoretical structure factor \( S_{\text{th}}(\tilde{q}) \) predicted by Eq. (4.26).

Figures 4.6 and 4.1 report numerical steady states obtained from the Lattice Boltzmann simulations with random growth rates in one and two dimensions. These simulations show the decrease in the amplitude of the density fluctuations as advection becomes stronger, confirming the assumption of the linear response theory in Sec. 4.1. In addition, unlike standard finite-difference schemes in which advection can lead to spurious alignment of population structure along lattice directions [62], the D2Q9 discretization, which ensures the fourth order isotropy of lattice tensors, avoids the artifact of preferred lattice orientations. The left column of Fig. 4.5 illustrates the steady state density fluctuations for the same realization of the random growth rate for different advection velocities: \( \mathbf{v} = \mathbf{0} \) in Fig. 5(a), \( \mathbf{v} = 8v_F \) along the \( x \)-axis in Fig. 4.5(c), and \( \mathbf{v} = 8v_F \) inclined at 37° relative to the directions associated with the periodic boundary conditions in Fig. 4.5(d). The steady state
Chapter 4: Striated Populations in Disordered Environments with Advection

Figure 4.5: Steady state density fluctuations $\phi_{\text{num}}(\mathbf{x}) \equiv |c_{\text{num}}(\mathbf{x}) - \bar{K}|/\bar{K}$ from Lattice Boltzmann simulations with $256^2$ sites (left) and the squared modulus of the discrete Fourier transform of the steady state density fluctuations $|\phi_{\text{num}}(\mathbf{q})|^2$ (right) for the same disorder realization but different flow velocities: $v = 0$ in (a) and (b), $v = 8v_F$ along the $x$-axis in (c) and (d), and $v = 8v_F$ inclined at a $37^\circ$ angle relative to the periodic boundary conditions in (e) and (f). The coordinate $\mathbf{x} = (x, y)$ and the wavevector $\mathbf{q} = (q_x, q_y)$ are displayed in the original (dimensional) scale with the lattice unit $l_0$ set to 1. The color codes for the structure factors ((b), (d), and (f)) are displayed in base-10 logarithmic scale. The parameters in this figure are $D = 5 \times 10^{-4}$, $a_0 = 10^{-5}$, $\Delta = 0.1a_0$, and $b = 10^{-9}$, which give $\xi_D = \sqrt{D/a_0} \approx 7$, $v_F = 2\sqrt{Da_0} \approx 1.4 \times 10^{-4}$, and $\bar{K} = a_0/b = 10^4$.

Long wavelength modes in the absence of advection, in (b), appear statistically isotropic as expected. The blue diamonds at large $q$ in (b), (d), and (f) are an artifact of the underlying lattice. Populations in Figs. 4.5(c) and (e) streak out along the advection direction, despite the bias imposed by an underlying square lattice. The elongation of steady state
populations along the advection direction also appears in the Fourier-transformed $|\phi_{\text{num}}(q)|^2$ shown on the right column of Fig. 4.5. There, the short wavelength modes in the longitudinal direction are noticeably suppressed, while the modes in the transverse direction are only mildly modified, leading to elliptical contours at small $q$’s with the minor and major axis aligned along the longitudinal and transverse direction, respectively. Notice the plots of $|\phi_{\text{num}}(q)|^2$ are approximately zero whenever $q \cdot r_{\text{lat}} = \pm \pi / l_0 = \pm \pi$, where $r_{\text{lat}}$ belongs to either the nearest neighbor basis $\{e_x, e_y\}$ or the next nearest neighbor basis $\{e_x + e_y, e_x - e_y\}$ of the square lattice, corresponding to the blue squares or the blue diamonds, respectively. These short-wavelength anisotropic lattice artifacts arise from the D2Q9 scheme that allows both nearest and next nearest neighbor hopping on the square lattice. However, the long-wavelength modes ($q$ near the origin), which characterize the macroscopic striations, are further away from the blue diamond and are orders of magnitude larger than zone-boundary modes with wavenumber $q \gtrsim \pi / \sqrt{2} l_0 = \pi / \sqrt{2}$. As discussed above, the long-wavelength physics is insensitive to these lattice artifacts, as can be confirmed by rotating Fig. 4.5(f) counterclockwise by an angle $37^\circ$ and comparing to Fig. 4.5(d). We thus expect that the long-range striation patterns are well-described by the behavior of $S_{\text{num}}(\tilde{q})$ near the origin, as determined by the Lattice Boltzmann simulations, averaged over ensembles of growth rates.

We now compare the long-wavelength modes of $S_{\text{num}}(\tilde{q})$ to $S_{\text{th}}(\tilde{q})$ predicted by Eq. (4.26). Although theoretical results of Sec. 4.4 assume spatially uncorrelated growth disorder whereas the Lattice Boltzmann simulation has an intrinsic short-range disorder correlation on the order of the lattice size $l_0 \equiv 1$, we expect that the long-distance statistics are insensitive to such microscopic details, provided the correlation length of the emerging pattern is larger than the lattice size. Because advection stretches out correlations in the longitudinal direction, the approximation
to a continuum model should become even more accurate at stronger advection. In fact, as shown in the comparisons between $S_{\text{num}}(\tilde{q})$ to $S_{\text{th}}(\tilde{q})$ in one and two dimensions in Figs 4.7 and 4.8 respectively, excellent agreement is obtained at long wavelengths even when $\xi_D$ is of order $2l_0 = 2$, which is only twice the correlation length of the simulated noise. The agreement is even better for stronger advection. Thus, the uncorrelated noise predictions of Sec. 4.4 are able to capture the long-range statistics of the striation pattern, despite the presence of inherent lattice-scale correlations of the simulated noise.

Figure 4.6: Typical steady-state densities $c^*_{\text{num}}(x_\parallel)$ in one dimension due to growth disorder and advection from the lattice Boltzmann simulation using a $D2Q9$ scheme of size $1 \times 1024$ sites. The parameters are $D = 5 \times 10^{-4}, a_0 = 1 \times 10^{-5}, \Delta = 0.4a_0,$ and $b = 10^{-9}$, resulting in $\xi_D = \sqrt{D/a_0} \approx 7.07, v_F = 2\sqrt{Da_0} \approx 1.4 \times 10^{-4},$ and $\bar{K} = a_0/b = 10^4$. The dashed lines denote the local carrying capacity $\bar{K}$. As advection becomes stronger, density fluctuations are suppressed and (longitudinal) correlation lengths stretch out. The details of correlations in one dimensional systems are elucidated in the structure factors of Fig.4.7.
Chapter 4: Striated Populations in Disordered Environments with Advection

Figure 4.7: Comparisons between the predicted one-dimensional structure factor $S_{\text{th}}(q)$ (solid lines), and numerical structure factor $S_{\text{num}}(q)$ (symbols) from the lattice Boltzmann simulation with $1 \times 256$ sites. The wavenumber $q$ is displayed in the original (dimensional) scale with the lattice unit $l_0$ set to 1. Simulation results are obtained from averaging the structure factors of density fluctuations over 200 growth rate realizations. The predicted long-wavelength modes of $S_{\text{th}}(q)$ are in excellent agreement with those of the simulated $S_{\text{num}}(q)$. Although the simulated structure factors are affected by inherent lattice-size disorder correlations near $q \sim \pi/l_0 = \pi$, the $q \to 0$ modes are much larger (note the logarithmic scale on the $S(q)$ axis). At strong advection when $\tilde{v} \gg 1$, long-wavelength modes are even more pronounced, and $S_{\text{th}}(q)$ is even closer to $S_{\text{num}}(q)$. The parameters in this figure are $D = 5 \times 10^{-4}$, $a_0 = 9 \times 10^{-5}$, $\Delta = 0.1a_0$, and $b = 10^{-9}$, giving $\xi_D = \sqrt{D/a_0} \approx 2.36$, $v_F = 2\sqrt{Da_0} \approx 4.2 \times 10^{-4}$, and $\bar{K} = a_0/b = 9 \times 10^4$.

4.6 Conclusions and Outlooks

Previous work has demonstrated the important role of spatially quenched disorder in the local growth rates on the long-time population structure that disperses through unidirectional advection and diffusion in hostile growth environments [63, 65, 66, 79]. Here, we study the role of growth disorder on the long-time population structure of populations in fertile random growth environments in which the local carrying capacity is strictly positive with weak disorder in the growth rates.
Chapter 4: Striated Populations in Disordered Environments with Advection

Figure 4.8: Comparisons between the predicted two-dimensional structure factor $S_{\text{th}}(q\parallel, q\perp)$ in the third row ((i)-(l)) and the simulated structure factor $S_{\text{num}}(q\parallel, q\perp)$ in the second row ((e)-(h)) from the Lattice Boltzmann (LB) simulations with $128^2$ sites. Typical simulated steady-state density fluctuations $\phi_{\text{num}}(x\parallel, x\perp)$ from the same disorder realization are shown in the first row ((a)-(e)). The coordinate $x = (x\parallel, x\perp)$ wavevector $q = (q\parallel, q\perp)$ are displayed in the original (dimensional) scale with the lattice unit $l_0$ set to 1. Simulation results are obtained from averaging the structure factors of density fluctuations over 200 disorder realizations. Color codes for the structure factors ((e)-(l)) are displayed in base-10 logarithmic scale. To highlight the agreements between the predicted long wavelength modes and the simulated ones, the values below $10^{-3}$-times the maximum are all represented in black; the predicted long-wavelength modes of $S_{\text{th}}(q)$ are in excellent agreement with those of the simulated $S_{\text{num}}(q)$. The next-nearest neighbor effect of the D2Q9 scheme represented by the dashed blue diamonds in $S_{\text{num}}(q\parallel, q\perp)$ is similar to that displayed in Fig. 4.5, and is many order of magnitude smaller than the long-wavelength modes. At strong advection (right columns), the longitudinal long-wavelength modes are even more pronounced while the transverse long-wavelength modes are not significantly affected; this longitudinal compression in the structure factors indicates the formation of striated population structure exemplified in (c) and (d). The parameters in this figure are $D = 5 \times 10^{-4}$, $a_0 = 9 \times 10^{-5}$, $\Delta = 0.1a_0$, and $b = 10^{-9}$, which give $\xi_D = \sqrt{D/a_0} \approx 2.36$, $v_F = 2\sqrt{Da_0} \approx 4.2 \times 10^{-4}$, and $\bar{K} = a_0/b = 9 \times 10^4$. 
Chapter 4: Striated Populations in Disordered Environments with Advection

To leading order in the strength of weak growth disorder, fluctuations from the homogeneous carrying capacity of steady-state populations can be perturbatively treated as a response to growth disorder.

Within the linear response theory, growth disorder and strong fluid advection leads to the emergence of striated steady-state population structure in two or higher dimensions with a periodic boundary condition. For large advection speeds \( v \gg v_F = 2\sqrt{Da_0} \), the longitudinal correlation length elongates as \( v/a_0 \) and the transverse correlation length approaches the finite velocity-independent value given by the diffusive localization length \( \xi_D = \sqrt{D/a_0} \). In contrast to steady-state populations with a point-like growth hot spot, whose transverse localization length contracts to zero as the advection speed increases, spatially quenched random growth disorder impedes the disappearance of transverse correlations. In our case, transverse correlations exhibit anomalous transverse diffusive spreading behavior of the early-time growth dynamics at strong advection as studied in Ref. [65]. The Lattice Boltzmann simulation of the generalized FKPP equation (1.22) confirms the emergence of striated population structure at strong advection, and verifies the predicted long-distance statistics from the linear response theory in one and two dimensions. Similar striated population structures might arise in natural marine microbial populations in spatially disordered growth environments and transported along closed recirculating flows, such as oceanic flows characterized by coherent Lagrangian vortices discussed in Refs. [107, 108].

In two-dimensional open flows, where populations do not actually recirculate, we numerically observe (not shown in this thesis) the pinning-depinning phenomena of the population frontier, the interface between occupied and unoccupied regions, analogous to those observed in autocatalytic chemical reaction front propagating in a porous media with a background fluid flow [109–111]. In particular, the population
Chapter 4: Striated Populations in Disordered Environments with Advection

frontier is pinned by growth rate disorder despite the presence of advection up to the critical advection speed, above which populations are completely washed away from the domain of interest. It would be interesting to explore, using the Lattice Boltzmann method, whether this pinning-depinning transition is a dynamic critical phenomena and, if so, whether it belongs to the universality class studied in Refs. [109–111].
Appendix A

Chapter 2 Appendices

A.1 Coupled Stochastic Dynamics of $f$ and $c_T$

Upon applying Ito’s change of variable to Eq. (2.19) and denoting $\mu = \mu_2$ and $(1 + s_o)\mu = \mu_1$ [24, 26], the coupled stochastic dynamics of $f$ and $c_T$ are described by

\begin{align}
\frac{d f}{dt} &= \mu v_R(f, c_T) + \mu \left( c_T + \frac{1}{N} \right) v_E(f) + \sqrt{\frac{\mu D_R(f, c_T) + \mu D_E(f)}{N}} \Gamma_f(t), \quad (A.1) \\
\frac{d c_T}{dt} &= \mu (1 + s_o f) v_G(c_T) + \mu (\alpha_1 + \alpha_2) c_T^2 f (1 - f) + \sqrt{\frac{\mu D_R(c_T, f) + \mu D_E(c_T, f)}{N}} \Gamma_{c_T}(t), \quad (A.2)
\end{align}

where the $N$-independent functions in the deterministic drifts and in the strength of an uncorrelated Gaussian white noise with $\langle \Gamma_a(t) \Gamma_b(t') \rangle = \delta_{ab} \delta(t - t')$ and $\langle \Gamma_a(t) \rangle = 0$
are given by

\[v_R(f,c_T) = \left(1 - c_T\right) - \frac{1}{N} \left(1 + \frac{c_T}{c_T}\right)\] so f(1 - f), \tag{A.3}\]

\[v_E(f) = f(1 - f)[\alpha_1 + (\alpha_1 + \alpha_2)f], \tag{A.4}\]

\[v_G(c_T) = c_T(1 - c_T), \tag{A.5}\]

\[D_R^{(f)}(f,c_T) = f(1 - f)[1 + s_o(1 - f)]\left(1 + \frac{c_T}{c_T}\right), \tag{A.6}\]

\[D_E^{(f)}(f) = -f(1 - f)[\alpha_1(1 - f)^2 + \alpha_2f^2], \tag{A.7}\]

\[D_R^{(c_T)}(f,c_T) = c_T(1 + c_T)(1 + s_o f), \tag{A.8}\]

\[D_E^{(c_T)}(f,c_T) = -(\alpha_1 + \alpha_2)c_T^2 f(1 - f). \tag{A.9}\]

Here, the subscript \(R\) denotes a contribution involving the reproductive advantage near the origin \(s_o\), whereas the subscript \(E\) denotes the contribution from evolutionary parameters defined in Eq. (2.10) \(\alpha_1 = (1 + s_o)\beta_1\) and \(\alpha_2 = \beta_2\), and \(v_E(f)\) and \(v_G(c_T)\) describe the deterministic replicator dynamics and the logistic growth dynamics respectively. The \(O(1/N)\) contributions to the deterministic drift induced by number fluctuations of \(c_1\) and \(c_2\) only appear in \(df/dt\) and originate from Ito’s change of variable formula [24, 26].

Under the replicator condition ( \(|\alpha_1| \ll 1, |\alpha_2| \ll 1, \text{ and } c_T \approx 1\) ), we have \(\left|D_E^{(f)}/D_R^{(f)}\right| \ll 1\) and \(\left|D_E^{(c_T)}/D_R^{(c_T)}\right| \ll 1\) so we can neglect the contributions from evolutionary parameters in the noise. Therefore, at \(s_o = 0\) the equations simplify,

\[
\frac{df}{dt} = \mu v_E(f) \left(c_T + \frac{1}{N}\right) + \sqrt{\frac{\mu}{N}} f(1 - f) \left(1 + \frac{c_T}{c_T}\right) \Gamma_f(t), \tag{A.10}\]

\[
\frac{dc_T}{dt} = \mu v_G(c_T) + \mu (\beta_1 + \beta_2)c_T^2 f(1 - f) + \sqrt{\frac{\mu}{N}} c_T(1 + c_T) \Gamma_{c_T}(t). \tag{A.11}\]
When $c_T \approx 1$, (A.10) reduces to Eq. (2.20) and (A.11) reduces to Eq. (2.21) in the limit $|\beta_1 + \beta_2| \ll 1/N \ll 1$. For $s_o \neq 0$ and $1/N \ll 1$, the coupled stochastic dynamics when $c_T \approx 1$ acquires contributions from the reproductive advantage near the origin $s_o$. The dynamics is now described by

$$\frac{df}{dt} = \mu v_R(f, c_T) + \mu c_T v_E(f) + \sqrt{\frac{\mu}{N} f(1-f) [1 + s_o (1-f)] \left(\frac{1 + c_T}{c_T}\right)} \Gamma_f(t),$$

(A.12)

$$\frac{dc_T}{dt} = \mu (1 + s_o f) v_G(c_T) + \mu (\alpha_1 + \alpha_2) c_T^2 f(1-f) + \sqrt{\frac{\mu}{N} c_T (1 + c_T)(1 + s_o f)} \Gamma_{c_T}(t).$$

(A.13)

When $\alpha_1 = \alpha_2 = 0$, (A.12) and (A.13) reduce to Eq. (2.34) and Eq. (2.35), respectively.

An interesting limit arises when $\alpha_1 = -\alpha_2 \sim O(s_o/N)$, where the fluctuation-induced selection term $\mu v_R(f, c_T = 1) = -(2\mu s_o/N) f(1-f)$ can compete with the usual fixed population size selection strength $\mu v_E(f) = \mu_1 f(1-f)$ and the genetic drift $D_R^{(d)}(f, c_T = 1) = (2\mu/N) f(1-f)[1+s_o(1-f)]$. In this case, fluctuation-induced selection can oppose the standard selection in population genetics, or equivalently in the Prisoner’s Dilemma of evolutionary game theory, and thus influence the dilemma of cooperation.
Appendix B

Chapter 3 Appendices

B.1 Matched Asymptotics for Strong Mutualism with a Varying Population Size

In this appendix, we construct the fixation probability for strong mutualism from the method of matched asymptotic expansions, or equivalently the boundary layer method, discussed in [98, 112–114]. First, consider the asymptotic large \( N \) solution \( u(c_1, c_2) \) of the backward Kolmogorov equation (3.2) near the saddle fixed points. In the neighborhood of the fixed point \((0, 1)\), we introduce the stretched coordinates \( \eta_1 = c_1 N \) and \( \eta_2 = (c_2 - 1)\sqrt{N} \), and rewrite the fixation probability in the new coordinates as \( U(\eta_1, \eta_2) = u(\eta_1/N, 1 + \eta_2/\sqrt{N}) \). Upon neglecting the terms of \( \mathcal{O}(1/\sqrt{N}) \), Eq. (3.2) in the new coordinates reads

\[
0 = \mu_1 \beta_1 \eta_1 \partial_{\eta_1} U + \left( \frac{\mu_1 + \mu_2}{2} - \frac{\mu_1 \beta_1}{2} \right) \eta_1 \partial_{\eta_1}^2 U - \mu_2 \eta_2 \partial_{\eta_2} U + \mu_2 \partial_{\eta_2}^2 U. \tag{B.1}
\]

Separation of variables \( U(\eta_1, \eta_2) = X_1(\eta_1)X_2(\eta_2) \) reduces Eq. (B.1) to an eigenvalue problem

\[
\lambda X_1 = \left( \frac{\mu_1 + \mu_2}{2} - \frac{\mu_1 \beta_1}{2} \right) \eta_1 X_1'' + \mu_1 \beta_1 \eta_1 X_1', \tag{B.2}
\]

\[
-\lambda X_2 = \mu_2 X_2'' - \mu_2 \eta_2 X_2'. \tag{B.3}
\]
Appendix B: Chapter 3 Appendices

where $\lambda$ is an eigenvalue. The general solution to (B.3) is

$$X_2(\eta_2) = C_1 H_{\lambda/\sqrt{2}}(\eta_2) + C_2 \, _1F_1\left(-\frac{\lambda}{2\mu_2}; \frac{1}{2}; \eta_2^2\right), \quad (B.4)$$

where $C_1$ and $C_2$ are constants, $H_n(z)$ is the Hermite polynomial, and $\, _1F_1(a;b;z)$ is the confluent hypergeometric function of the first kind. The matching condition to the plateau fixation probability $U(\eta_1 \to \infty, \eta_2) = P$ enforces $X_2(\eta_2)$ to be a constant. This is possible only when $C_2 = 0$ and $\lambda = 0$ so $X_2(\eta_2) = C_1$. Because zero is the only eigenvalue consistent with the matching condition, (B.2) reduces to

$$0 = \left(\frac{\mu_1 + \mu_2}{2} - \frac{\mu_1\beta_1}{2}\right) \eta_1 X_1'' + \mu_1\beta_1 \eta_1 X_1', \quad (B.5)$$

whose general solution is

$$X_1(\eta_1) = B_1 - B_2 \left(\frac{\mu_1 + \mu_2}{2\mu_1\beta_1} - \frac{1}{2}\right) e^{-\eta_1\mu_1\beta_1/(\mu_1 + \mu_2 - \mu_1\beta_1)}, \quad (B.6)$$

where $B_1$ and $B_2$ are constants. By imposing the boundary condition $U(0, \eta_2) = 0$ and the matching condition $U(\eta_1 \to \infty, \eta_2) = P$, it follows that the fixation probability in the original coordinates valid in the vicinity of the fixed point $(0, 1)$ is

$$u(c) = P + P e^{-N e_1[2\beta_1/(2 - \beta_1)]}. \quad (B.7)$$

A similar argument can be applied to the asymptotic solution near the fixed point $(1, 0)$. In this case, the stretched coordinates are $\eta_1 = (c_1 - 1)\sqrt{N}$ and $\eta_2 = c_2 N$, with the fixation probability in the new coordinates given by $U(\eta_1, \eta_2) = u(1 + \eta_1/\sqrt{N}, \eta_2/N)$. Eq. (3.2) in the new coordinates, with terms of $O(1/\sqrt{N})$ neglected,
reads

\[ 0 = -\mu_1 \eta_1 \partial_{\eta_1} U + \mu_1 \partial_{\eta_1}^2 U + \mu_2 \beta_2 \eta_2 \partial_{\eta_2} U + \left( \frac{\mu_1 + \mu_2}{2} - \frac{\mu_2 \beta_2}{2} \right) \eta_2 \partial_{\eta_2}^2 U, \quad (B.8) \]

which is equivalent to (B.1) with indices 1 and 2 interchanged. Following the method of separation of variables as above and imposing the boundary condition \( U(\eta_1, 0) = 1 \) as well as the matching condition \( U(\eta_1, \eta_2 \to \infty) = P \), we arrive at the fixation probability valid in the vicinity of the fixed point \((1, 0)\)

\[ u(c) = P + (1 - P)e^{-Nc_2[2\beta_2/(2-\beta_2)]}. \quad (B.9) \]

Now consider the asymptotic solutions away from the saddle fixed points but still in the boundary layers. In the boundary layer adjacent to the absorbing boundary \( c_1 = 0 \) but away from the saddle fixed point \((0, 1)\), we introduce the stretched coordinate \( \eta_1 = c_1 N \) and \( \eta_2 = c_2 \). Upon neglecting the contributions of \( O(1/N) \) and rewriting the fixation probability in the new coordinate as \( U(\eta_1, \eta_2) = u(\eta_1/N, \eta_2) \), Eq. (3.2) becomes

\[ 0 = 2\mu_2 \eta_2 (1 - \eta_2) \partial_{\eta_2} U + 2[\mu_1 - (\mu_1 - \mu_1 \beta_1) \eta_2] \eta_1 \partial_{\eta_1} U \]

\[ + [\mu_1 - (\mu_1 + \mu_1 \beta_1) \eta_2] \eta_1 \partial_{\eta_1}^2 U. \quad (B.10) \]

We can turn (B.10) into a separable PDE and solve the associated eigenvalue problem by transforming to the new coordinates \( x_1 = \eta_1/\Phi_2(\eta_2) \), and \( x_2 = \eta_2 \). Substituting the coordinate transformation \( V(x_1, x_2) = U(x_1 \Phi_2(x_2), x_2) \) into (B.10), we find \( V(x_1, x_2) \) satisfies a separable PDE

\[ 0 = x_1 \partial_{x_1}^2 V + x_1 \partial_{x_1} V + 2\frac{\mu_2}{\mu_1} \left[ \frac{x_2(1 - x_2)}{1 - (1 + \beta_1)x_2} \Phi_2(x_2) \right] \partial_{x_2} V. \quad (B.11) \]
with $\Phi_2(x)$ obeys the first order differential equation given in Eq (3.7). Separation of variables $V(x_1, x_2) = V_1(x_1)V_2(x_2)$ turns (B.11) into an eigenvalue problem

\begin{align*}
\lambda V_1 &= x_1 V_1'' + x_1 V_1', \quad (B.12) \\
-\lambda V_2 &= 2\frac{\mu_2}{\mu_1} \left[ -\frac{x_2(1 - x_2)}{1 - (1 + \beta_1)x_2} \Phi_2(x_2) \right] V_2', \quad (B.13)
\end{align*}

with $\lambda$ the eigenvalue. Again, matching to the fixation probability at the plateau $U(\eta_1 \to \infty, \eta_2) = P$ enforces $V_2(x_2)$ to be constant which is possible only if $\lambda = 0$. The general solution to (B.12) with $\lambda = 0$ is

\begin{equation}
V_1(x_1) = D_1 - D_2 e^{-x_1}. \quad (B.14)
\end{equation}

Upon imposing the boundary condition $U(0, \eta_2) = 0$ as well as the matching condition $U(\eta_1 \to \infty, \eta_2) = P$, we obtain the fixation probability in the original coordinate valid within the boundary layer adjacent to the absorbing boundary $c_1 = 0$, namely

\begin{equation}
u(c) = P - Pe^{-Nc_1/\Phi_2(c_2)}. \quad (B.15)\end{equation}

Upon matching (B.15) to the asymptotic solution in the vicinity of the saddle fixed point $(1, 0)$, (B.9), we find a first order differential equation governing $\Phi_2$, given by Eq. (3.7), with the matching condition $\lim_{x \to 1} \Phi_2(x) = \frac{2 - \beta_1}{2\beta_1}$.

A similar argument with index 1 and 2 interchanged determines the asymptotic solution within the boundary layer adjacent to the absorbing boundary $c_2 = 0$. We find that the fixation probability in this region is given by

\begin{equation}
u(c) = P + (1 - P)e^{-Nc_2/\Phi_1(c_1)}, \quad (B.16)\end{equation}
where $\Phi_1(x)$ obeys Eq. (3.5) subject to the matching condition $\lim_{x \to 1} \Phi_1(x) = \frac{2 - \beta_2}{2\beta_2}$.

Therefore, the global solution with smooth crossovers from the plateau $P$ to the the boundary layer behavior of (B.15) and (B.16) is given by Eq. (3.8).

### B.2 Fixation from a Quasi-Stationary Distribution

The analysis in this section follows the general discussion on the high dimensional exit problem by Grasman and Herwaarden [98]. We first argue that, near the absorbing boundaries, the quasi-stationary distribution (QSD) $p_{st}(c)$ is peaked at the saddle fixed point $(0,1)$ and $(1,0)$. To see this, consider the (stationary) Fokker-Planck equation

$$0 = \sum_{i=1}^{2} \left( - \partial_c v_i p_{st} + \frac{1}{2N} \partial_c^2 D_i p_{st} \right), \quad (B.17)$$

where $v_i$ and $D_i$ are given by Eqs. (2.15)-(2.18). When the problem can be regarded as a rare event escape from a metastable state, the asymptotic solution to (B.17) is solved by the WKB ansatz [98, 102, 115, 116]

$$p_{st}(c) = w(c)e^{-N\Psi(c)}. \quad (B.18)$$

Substituting (B.18) into (B.17) and collecting the leading order terms in $N$ leads to an eikonal equation,

$$0 = \sum_{i=1}^{2} \left( v_i(\partial_c \Psi) + \frac{1}{2} D_i(\partial_c \Psi)^2 \right). \quad (B.19)$$
Collecting terms of $O(1)$ results in

$$0 = \sum_{i=1}^{2} \left( D_i (\partial_{c_i} \Psi) + v_i \right) \partial_{c_i} w + \sum_{i=1}^{2} \left( \frac{1}{2} D_i \partial_{c_i}^2 \Psi + (\partial_{c_i} D_i) (\partial_{c_i} \Psi) \right) w. \quad (B.20)$$

In the neighborhood of the absorbing boundary $c_1 = 0$, we expand $\Psi$ around $c_1 = 0$ as

$$\Psi(c) = \Psi_2^{(0)}(c_2) + \Psi_2^{(1)}(c_2)c_1 + \frac{1}{2} \Psi_2^{(2)}(c_2)c_1^2 + \cdots, \quad (B.21)$$

where the subscript 2 of $\Psi$ denotes the expansion around the fixation of species 2 and the superscript labels the order of expansion. Upon substituting the expansion (B.21) into (B.19) and collecting terms of $O(c_0^1)$, we find

$$\Psi_2^{(1)}(c_2) = -\frac{2}{\beta_1} \left( \frac{c_2 - 1}{c_2 + 1} \right).$$

Therefore, in the limit $c_1 \to 0$, $\Psi(c)$ is minimal at $c_2 = 1$, implying that $p_{st}$ is peaked in the neighborhood of the fixed point $(0,1)$ provided $N \gg 1$.

For the behavior of $w(c)$ near $(0,1)$, it turns out the singular behavior of $w$ when $c_1 \to 0$ scales as $w \sim 1/c_1$. We refer to the discussion in Ref. [98] for the related problem of extinction probability in the predator-prey model. The singular behavior suggests the QSD is concentrated in the neighborhood of the saddle fixed point.

To extract the quantitative behavior of $p_{st}$ near the saddle fixed point $(0,1)$, we Taylor expand $\Psi$ around $(0,1)$

$$\Psi(c) = \Psi_2^{(0)} + \Psi_2^{(1)}c_1 + \Psi_2^{(2)}(c_2 - 1) + \frac{1}{2} \Psi_2^{(3)}(c_2 - 1)^2 + \cdots, \quad (B.22)$$

where we denote the $i$th expansion coefficient around the saddle fixed point of species 2 by $\Psi_2^{(i)}$. Upon substituting the expansion (B.22) into (B.19) and (B.20) we get $\Psi_2^{(1)} = -2\beta_1/(2 - \beta_1)$, $\Psi_2^{(2)} = 0$, and $\Psi_2^{(3)} = 1$. Therefore, in the neighborhood of the
fixed point \((0, 1)\), the QSD takes the form
\[
p_{st}(c) \approx \frac{\tilde{w}_2^{(0)} \exp(-N\tilde{\Psi}_2^{(0)})}{c_1} \exp \left[ N \left( \frac{2\beta_1}{2 - \beta_1} c_1 - \frac{(c_2 - 1)^2}{2} \right) \right]. \quad (B.23)
\]

Similar arguments lead to the behavior of the QSD in the neighborhood of the fixed point \((1, 0)\), which reads
\[
p_{st}(c) \approx \frac{\tilde{w}_1^{(0)} \exp(-N\tilde{\Psi}_1^{(0)})}{c_2} \exp \left[ N \left( \frac{2\beta_2}{2 - \beta_2} c_2 - \frac{(c_1 - 1)^2}{2} \right) \right]. \quad (B.24)
\]

We now relate the behavior of the QSD near the absorbing boundaries to the plateau fixation probability \(P\) in the bulk region by employing the identity resulting from the divergence theorem:
\[
\int_{\Omega} (p\hat{L}u - u\hat{M}p)dc_1 dc_2 = \int_{\partial\Omega} \left( \sum_{i=1}^{2} \frac{1}{2N} \left[ n_i D_i (p\partial_{c_i} u - u\partial_{c_i} p) - n_i (\partial_{c_i} D_i)pu \right] \right. \\
+ \left. \sum_{i=1}^{2} n_i v_i pu \right) dS, \quad (B.25)
\]
where \(\hat{L}\) is the backward-Kolmogorov operator, \(u\) is the solution to the backward-Kolmogorov equation, \(\hat{M}\) is the forward-Kolmogorov (Fokker-Planck) operator, \(p\) is the solution to the forward-Kolmogorov equation, \(\Omega\) is the domain of interest, and \(n_i\) is the \(i^{th}\) components of the normal vector at the boundary \(\partial\Omega\). In the long-time limit when the QSD has already developed, the volume integral (left hand side) of (B.25) vanishes since \(\hat{L}u = 0\) and \(\hat{M}p_{st} = 0\). To evaluate the surface integral in (B.25) and avoid the singularity of \(p_{st}\) on each absorbing boundary, we consider the domain \(\Omega = \{c | c_1 > \varepsilon, c_2 > \varepsilon\}\) and evaluate (B.25) in the limit \(\varepsilon \to 0\). In this
Appendix B: Chapter 3 Appendices

domain, (B.25) becomes

\[
0 = \int_{\epsilon}^{\infty} dc_2 \left( \frac{1}{2N} \left[ D_1 (p_{st} \partial_{c_1} u - u \partial_{c_1} p_{st}) - (\partial_{c_1} D_1) p_{st} u \right] + v_1 p_{st} u \right)_{c_1=\epsilon}
+ \int_{\epsilon}^{\infty} dc_1 \left( \frac{1}{2N} \left[ D_2 (p_{st} \partial_{c_2} u - u \partial_{c_2} p_{st}) - (\partial_{c_2} D_2) p_{st} u \right] + v_2 p_{st} u \right)_{c_2=\epsilon} \tag{B.26}
\]

(B.26) relates the plateau fixation probability \( P \) contained in \( u \) by (B.7) and (B.9) to the boundary behavior of \( p_{st} \). Substituting the asymptotic solutions of the QSD given by (B.23) and (B.24), the asymptotic solutions of \( u \) given by (B.7) and (B.9), and the deterministic drifts as well as diffusion coefficients given by Eqs. (2.15)-(2.18) into (B.26), we obtain after taking the limits \( \epsilon \to 0 \) and \( N \gg 1 \)

\[
0 = \frac{2\beta_1}{2 - \beta_1} \bar{w}_2^{(0)} \exp(-N \bar{\Psi}_2^{(0)}) \int_0^{\infty} dc_2 \left\{ \left[ -\frac{\mu_1 P}{2} - \frac{\mu_1 (1 - \beta_1) P}{2} \right] \right. \\
\times \exp \left[ -\frac{N(c_2 - 1)^2}{2} \right] \right\} \\
+ \frac{2\beta_2}{2 - \beta_2} \bar{w}_1^{(0)} \exp(-N \bar{\Psi}_1^{(0)}) \int_0^{\infty} dc_1 \left\{ \left[ \mu_2 \left( 1 - \frac{2 - \beta_2}{2\beta_2} - \frac{P}{2} \right) \right. \\
+ \mu_2(1 - \beta_2) \left( 1 + \frac{2 - \beta_2}{2\beta_2} - \frac{P}{2} \right) \right\} c_1 \times \exp \left[ -\frac{N(c_1 - 1)^2}{2} \right] \right\} \tag{B.27}
\]

The integrals can be evaluated by the standard method of Laplace integration when \( N \gg 1 \). The result reads

\[
0 = \frac{2\beta_1}{2 - \beta_1} \bar{w}_2^{(0)} \exp(-N \bar{\Psi}_2^{(0)}) \sqrt{\frac{\pi}{N}} \left[ -\frac{\mu_1 P}{2} - \frac{\mu_1 (1 - \beta_1) P}{2} \right] \\
+ \frac{2\beta_2}{2 - \beta_2} \bar{w}_1^{(0)} \exp(-N \bar{\Psi}_1^{(0)}) \sqrt{\frac{\pi}{N}} \times \left[ \mu_2 \left( 1 - \frac{2 - \beta_2}{2\beta_2} - \frac{P}{2} \right) \right. \\
\left. + \mu_2(1 - \beta_2) \left( 1 + \frac{2 - \beta_2}{2\beta_2} - \frac{P}{2} \right) \right] \tag{B.28}
\]

Upon rewriting \( \beta_1 = \mu_1 - \lambda_{12} N \), and \( \beta_2 = \mu_2 - \lambda_{21} N \) and keeping only the
leading order term in $1/N$, we obtain the plateau fixation probability

$$P \approx \frac{\lambda_{21} \bar{w}_1(0) e^{-N \bar{\Psi}_1(0)}}{\lambda_{21} \bar{w}_1(0) e^{-N \bar{\Psi}_1(0)} + \lambda_{12} \bar{w}_2(0) e^{-N \bar{\Psi}_2(0)}}. \quad (B.29)$$

Recall that $\bar{w}_1(0) e^{-N \bar{\Psi}_1(0)}$ and $\bar{w}_2(0) e^{-N \bar{\Psi}_2(0)}$ are $p_{st}(c)$ evaluated at $(1,0^+)$ and $(0^+,1)$. Consequently, (B.29) is the ratio of the flux into $(1,0)$ to the total flux into $(1,0)$ and $(0,1)$. In the limit $N \gg 1$, (B.23) and (B.24) imply that, on each absorbing boundary, the QSD peaks up at the saddle fixed point while the width around the peak becomes vanishingly narrow; accordingly, the flux into the saddle fixed point well approximates the flux into the corresponding absorbing boundary. Hence, (B.29) describes the ratio of flux into the absorbing boundary at $f=1$ to the total flux into both the absorbing boundaries at $f=0$ and $f=1$.

Note that (B.29) can be rewritten in the form similar to Eq. (3.1) as

$$P \approx \frac{1}{1 + e^{-N \Delta S_0 + \Delta S_1}}, \quad (B.30)$$

where $\Delta S_0 \equiv \bar{\Psi}_2(0) - \bar{\Psi}_1(0)$ and $\Delta S_1 \equiv \ln(\lambda_{12} \bar{w}_2(0)) - \ln(\lambda_{21} \bar{w}_1(0))$. Since $\bar{w}_1(0)^{(0)}$ and $\bar{\Psi}_1(0)^{(0)}$ are independent of $N$, we can vary $N$ while fixing $s_o$, $\beta_1$ and $\beta_2$ to infer $\Delta S_0$ and $\Delta S_1$ by fitting the plateau fixation probability $P$ to simulations. In principle, the exact values of $\Delta S_0$ and $\Delta S_1$ may be obtained numerically by simultaneously solving $\bar{w}_1(0)^{(0)}$ and $\bar{\Psi}_1(0)^{(0)}$ from (B.19) and (B.20) [32, 96, 97, 102], but these are beyond the scope of this work.
Chapter 4 Appendices

C.1 Green’s Functions, Correlation Functions, and Structure Factors

This appendix contains calculation details of analytical results obtained in chapter 4. First, we solve Eq. (4.13) by a direct Fourier transformation, as a check on Eq. (4.17). Upon setting $(\Delta/a_0)$ to unity for convenience, the Green’s function of Eq. (4.13) in Fourier-space reads

$$G(\tilde{q}) = \frac{1}{\tilde{q}^2 - 2i\tilde{q}|| + 1}. \quad (C.1)$$

Inverse Fourier transformation gives the real-space Green’s function in $d$-dimensions:

$$G(\tilde{x}) = \int \frac{d^d\tilde{q}}{(2\pi)^d} e^{i\tilde{q} \cdot \tilde{x}} \frac{1}{\tilde{q}^2 - 2i\tilde{q}|| + 1}. \quad (C.2)$$

To evaluate (C.2), we use a trick, often employed to calculate correlation functions in field theory [117], to convert the denominator into an exponential integral:

$$\frac{1}{X^n} = \int_0^\infty ds \frac{s^{n-1}e^{-sX}}{(n-1)!}. \quad (C.3)$$
With $n = 1$, (C.2) becomes

$$G(\tilde{x}) = \int_0^\infty ds \int \frac{d^d \tilde{q}}{(2\pi)^d} e^{i \tilde{q} \cdot \tilde{x}} e^{-s(\tilde{q}^2 - 2i\tilde{q} \cdot \tilde{v} + 1)},$$

$$= \int_0^\infty ds \ e^{-s} \left( \int \frac{d^{d-1} \tilde{q}}{(2\pi)^{d-1}} e^{i \tilde{q} \cdot \tilde{x} - s\tilde{q}^2} \right) \frac{1}{(2\pi)^{d-1}} \sqrt{\frac{\pi}{s}} \exp(-\tilde{x}^2/4s) \times \left( \int \frac{d\tilde{q}_\parallel}{(2\pi)} e^{i \tilde{q}_\parallel (\tilde{x}_\parallel - 2\tilde{v}) - s\tilde{q}_\parallel^2} \right) \frac{1}{(2\pi)^{d-1}} \sqrt{\frac{\pi}{s}} \exp(-\tilde{x}_\parallel^2/4s),$$

$$= \exp(\tilde{v} \cdot \tilde{x}) \frac{(2\pi)^{d/2}}{} \int_0^\infty ds \ s^{-d/2} e^{-s(1+\tilde{v}^2) - \tilde{x}^2/4s}. \quad (C.4)$$

Upon recalling the integral representation of the modified Bessel function of a second kind [118]

$$K_\nu(z) = \frac{1}{2} \left( \frac{1}{2} z \right)^\nu \int_0^\infty dt \ t^{-\nu-1} e^{-t-z^2/4t}, \quad (C.5)$$

and using the symmetry $K_{-\nu}(z) = K_\nu(z)$, (C.4) becomes

$$G(\tilde{x}) = \frac{\exp(\tilde{v} \cdot \tilde{x})}{(\pi)^{d/2}} \left[ \left( \frac{\tilde{x}}{\sqrt{1+\tilde{v}^2}} \right)^{1-d/2} K_{1-d/2} \left( \sqrt{1+\tilde{v}^2} \right) \right], \quad (C.6)$$

which is identical to Eq. (4.17) with the gauge transformed Green’s function $G_\tilde{v}(\tilde{x})$ of Eq. (4.16) by the square bracket (after restoring the factor $\Delta/a_0$).

We now outline the details of calculations to achieve analytical results in Sec. 4.4 that follow from the general prescription of the two-point correlation function of Eq. (4.27), where again we set $\Delta/a_0 = 1$ for short. For $d = 1$, the correlation function reads

$$\langle \phi(\tilde{x})\phi(0) \rangle = \frac{1}{6\pi} \int_{-\infty}^\infty d\tilde{q} \ e^{i \tilde{q} \cdot \tilde{x}} \frac{1}{4} \prod_{k=1}^4 \frac{1}{(\tilde{q} - \tilde{q}_k)}, \quad (C.7)$$

where the four poles, that arise from the factorization of $[(\tilde{q}^2 + 1)^2 + 4\tilde{v}^2\tilde{q}^2]$, are
Appendix C: Chapter 4 Appendices

located at

\[
\tilde{q}_1 = -\tilde{q}_2 = i(\sqrt{1 + \tilde{v}^2} + \tilde{v}),
\]
\[
\tilde{q}_3 = -\tilde{q}_4 = i(\sqrt{1 + \tilde{v}^2} - \tilde{v}).
\]

Closing the contour in the upper (lower) half plane for \( \tilde{x} > 0 \) (\( \tilde{x} < 0 \)), then using the residue theorem yields the exact result of the one-dimensional correlation function given by Eq. (4.28)

For \( d = 2 \), we first calculate the transverse two-point correlation function:

\[
\langle \phi(0, \tilde{x}_\perp)\phi(0) \rangle \sim \int \frac{d\tilde{q}_\perp}{2\pi} e^{i\tilde{q}_\perp \tilde{x}_\perp} \int \frac{d\tilde{q}_\parallel}{2\pi} \frac{1}{(\tilde{q}_\parallel^2 + 1)^2 + 4\tilde{v}^2\tilde{q}_\parallel^2} = \frac{1}{4\pi} \int d\tilde{q}_\perp \frac{e^{i\tilde{q}_\perp \tilde{x}_\perp}}{(\tilde{q}_\perp^2 + 1)^{3/2}}.
\]

(C.8)

In the limit \( \tilde{v} \to \infty \), we obtain

\[
\lim_{\tilde{v} \to \infty} \langle \phi(0, \tilde{x}_\perp)\phi(0) \rangle \sim \frac{1}{4\pi\tilde{v}} \int d\tilde{q}_\perp \frac{e^{i\tilde{q}_\perp \tilde{x}_\perp}}{(\tilde{q}_\perp^2 + 1)}
\]

= \[ \frac{\exp(-|\tilde{x}_\perp|\tilde{v})}{4\tilde{v}}. \]

(C.9)

(C.10)

Now consider the longitudinal two-point correlation function.

\[
\langle \phi(\tilde{x}_\parallel, 0)\phi(0) \rangle \sim \int \frac{d\tilde{q}_\parallel}{2\pi} e^{i\tilde{q}_\parallel \tilde{x}_\parallel} \int \frac{d\tilde{q}_\perp}{2\pi} \frac{1}{(\tilde{q}_\parallel^2 + 1)^2 + 4\tilde{v}^2\tilde{q}_\parallel^2} = \int \frac{d\tilde{q}_\parallel}{2\pi} e^{i\tilde{q}_\parallel \tilde{x}_\parallel} \left\{ \frac{1}{8i\tilde{v} |\tilde{q}_\parallel|} \left[ \frac{1}{\sqrt{f^+(\tilde{q}_\parallel)}} - \frac{1}{\sqrt{f^-(\tilde{q}_\parallel)}} \right] \right\},
\]

(C.11)

(C.12)
where the term in the curly bracket is the result of contour integration with respect to $\tilde{q}_\perp$ with the contour in the upper (lower) half plane for $\tilde{q}_\parallel > 0$ ($\tilde{q}_\parallel < 0$), and $f_\pm(\tilde{q}_\parallel) \equiv (\tilde{q}_\parallel^2 + 1) \pm 2i\tilde{v}\tilde{q}_\parallel$. Because the integrand is real, the resulting integral must be real although it manifestly contains an imaginary part. To rewrite the squared bracket explicitly as a real-function, consider the following change of variable: $re^{i\theta} \equiv f_\pm(\tilde{q}_\parallel)$. Then, $r$ and $\theta$ are given by

$$r = f_+ + (\tilde{q}_\parallel^2 + 1) - f_-(\tilde{q}_\parallel^2)$$

and

$$\theta = -\theta_+$$

with $\tan(\theta_+) = 2\tilde{v}\tilde{q}_\parallel/(\tilde{q}_\parallel^2 + 1)$. After some algebra and trigonometric identities, one finds the term in the curly bracket of (C.11) becomes

$$\left[ \frac{1}{4ir|\sin(\theta_+)|} \frac{1}{\sqrt{r}} \left( e^{-i\theta_+/2} - e^{-i\theta_-/2} \right) \right] = \frac{1}{2\sqrt{2}} \frac{r^{-3/4}}{\sqrt{1 + \cos(\theta_+)}}$$

Substituting the expression into (C.11), we obtain

$$\langle \phi(\tilde{x}_\parallel, 0)\phi(0) \rangle = \int \frac{d\tilde{q}_\parallel}{4\sqrt{2}\pi} e^{i\tilde{q}_\parallel\tilde{x}_\parallel}$$

$$\times \frac{1}{[\tilde{q}_\parallel^2 + 1 + 4\tilde{v}^2\tilde{q}_\parallel^2]^{3/2}} \left\{ 1 + \frac{(\tilde{q}_\parallel^2 + 1)^2}{[\tilde{q}_\parallel^2 + 1 + 4\tilde{v}^2\tilde{q}_\parallel^2]^{1/2}} \right\}^{1/2}. \tag{C.15}$$

In the limit $\tilde{v} \to \infty$, one obtains the longitudinal correlation function

$$\lim_{\tilde{v} \to \infty} \langle \phi(\tilde{x}_\parallel, 0)\phi(0) \rangle \sim \frac{1}{4\sqrt{2}\pi} \int d\tilde{q}_\parallel e^{i\tilde{q}_\parallel\tilde{x}_\parallel}$$

$$\times \frac{1}{(1 + 4\tilde{v}^2\tilde{q}_\parallel^2)^{1/4}}.$$ \tag{C.16}

$$= \frac{-2^{5/4}}{\tilde{v}\pi\Gamma(-1/4)} \left( \frac{\|\tilde{x}_\parallel\|}{2\tilde{v}} \right)^{1/4} K_{-1/4} \left( \frac{\|\tilde{x}_\parallel\|}{2\tilde{v}} \right), \tag{C.17}$$

which gives (4.32).
Bibliography


Bibliography


[52] Q.-X. Liu, P. M. Herman, W. M. Mooij, J. Huisman, M. Scheffer, H. Olff, and


