DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Diffusion, Absorbing States, and Nonequilibrium Phase Transitions in Range Expansions and Evolution

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation	No citation.
Accessed	February 17, 2015 12:52:14 AM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:13070032
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

(Article begins on next page)

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Physics

have examined a dissertation entitled

Diffusion, Absorbing States, and Nonequilibrium Phase Transitions in Range Expansions and Evolution

presented by Maxim Olegovich Lavrentovich

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________________

Typed name: Professor David Nelson, Chair

Signature ______________________________

Typed name: Professor Erel Levine

Signature ______________________________

Typed name: Professor Subir Sachdev

Date: August 21, 2014

Diffusion, Absorbing States, and Nonequilibrium Phase Transitions in Range Expansions and Evolution

A dissertation presented

by

Maxim Olegovich Lavrentovich

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

August 2014

© 2014 — Maxim Olegovich Lavrentovich

All rights reserved.

Thesis advisor
David Robert Nelson

Author
Maxim Olegovich Lavrentovich

Diffusion, Absorbing States, and Nonequilibrium Phase Transitions in Range Expansions and Evolution

Abstract

The spatial organization of a population plays a key role in its evolutionary dynamics and growth. In this thesis, we study the dynamics of range expansions, in which populations expand into new territory. Focussing on microbes, we first consider how nutrients diffuse and are absorbed in a population, allowing it to grow. These nutrients may be absorbed before reaching the population interior, and this "nutrient shielding" can confine the growth to a thin region on the population periphery. A thin population front implies a small local effective population size and enhanced number fluctuations (or genetic drift). We then study evolutionary dynamics under these growth conditions. In particular, we calculate the survival probability of mutations with a selective advantage occurring at the population front for two-dimensional expansions (e.g., along the surface of an agar plate), and three-dimensional expansions (e.g., an avascular tumor). We also consider the effects of irreversible, deleterious mutations which can lead to the loss of the advantageous mutation in the population via a "mutational meltdown," or non-equilibrium phase transition. We examine the effects of an inflating population frontier on the phase transition. Finally, we discuss how spatial dimension and frontier roughness influence range expansions of mutualistic, cross-feeding variants. We find here universal features of the phase diagram describing the onset of a mutualistic phase in which the variants remain mixed at long times.

Contents

Title Page **i**

Abstract **iii**

Citations to Previously Published Work **vii**

Acknowledgments **viii**

Dedication **x**

1 Introduction **1**

1.1 Motivation and Random Walks 1

1.2 Nutrient Shielding 11

1.3 Evolution of Range Expansions with Inflating Frontiers 19

1.4 Nonequilibrium Phase Transitions and Range Expansions with Mutualism 37

2 Nutrient Shielding in Clusters of Cells **43**

2.1 Introduction 43

2.2 Theoretical Model 48

2.2.1 Single Cell Nutrient Uptake 48

2.2.2 Nutrient Uptake in Cell Clusters 55

2.2.3 The Macroscopic Screening Length ξ 61

2.2.4 The Dilute Cell Limit 64

2.2.5 Effective Medium Theory for Dense Cell Clusters 66

2.3 Numerical Simulations 71

2.3.1 Cell Cluster Simulation 71

2.3.2 Comparison with Theoretical Results 74

2.4 Experiments 77

2.4.1 Experimental Setup 78
2.4.2 Data Analysis and Results 80
2.5 Conclusions 84

3 Radial Domany-Kinzel Models with Mutation and Selection 87
3.1 Introduction 87
3.2 Simulations 94
3.2.1 A Generalized Domany-Kinzel Model 96
3.2.2 The Bennett Model 98
3.3 Voter Models with Inflation 102
3.3.1 Neutral Evolution 108
3.3.2 Heterozygosity in Radial Range Expansions 113
3.3.3 Single Seed Dynamics and Selective Advantage 117
3.3.4 Directed Percolation with Inflation 126
3.4 Radial Range Expansions with Deflation 142
3.5 Conclusions 146

4 Survival Probabilities at Spherical Frontiers 148
4.1 Introduction 148
4.2 Simulations 153
4.3 Population Genetics at Compact Population Fronts 157
4.4 Two-Dimensional Range Expansions 163
4.5 Three-Dimensional Range Expansions 169
4.5.1 Treadmilling Fronts ($\Theta = 0$) 172
4.5.2 Linearly Inflating Fronts ($\Theta = 1$) 174
4.5.3 Marginally Inflating Fronts ($\Theta = 1/2$) 177
4.5.4 Rough Fronts 179
4.6 Conclusions 182

5 Asymmetric Mutualism in Two- and Three-Dimensional Range Expansions 184
5.1 Introduction 184
5.2 Flat Front Models 186
5.3 Rough Front Models 193
5.4 Conclusions 195

Appendices 196

A Chapter 2 Appendices 197

A.1 Detailed Effective Medium Calculation 197

B Chapter 3 Appendices 205

B.1 Quasi-species and Mutation-Selection Balance 205

B.2 Coarse-Graining and Langevin Description 208

B.3 Single Sector Dynamics 211

C Chapter 4 Appendices 216

C.1 Response Functional Formalism 216

C.2 Survival Probability Calculation in Three Dimensions 220

C.3 Survival Probability for $s < 0$ 222

D Chapter 5 Appendices 224

D.1 Flat Front Models and Phase Boundaries 224

D.2 Undulating Front Models 228

D.3 The Rough DP2 Transition in $d = 1 + 1$ Dimensions 231

Bibliography 235

Citations to Previously Published Work

Chapter 2 appears in its entirety as

Nutrient shielding in clusters of cells
M. O. Lavrentovich, J. H. Koschwanez, and D. R. Nelson
Physical Review E **87**, 062703 (2013)

Chapter 3 appears in its entirety as

Radial Domany-Kinzel models with mutation and selection
M. O. Lavrentovich, K. S. Korolev, and D. R. Nelson
Physical Review E **87**, 012103 (2013)

Chapter 4 will be submitted in its entirety as

Survival probabilities at spherical frontiers
M. O. Lavrentovich and D. R. Nelson
to be submitted for publication (2014)

Chapter 5 appears in its entirety as

Asymmetric mutualism in two- and three-dimensional range expansions
M. O. Lavrentovich and D. R. Nelson
Physical Review Letters **112**, 138102 (2014)

Acknowledgments

I am grateful to the many people who have helped and supported me throughout my graduate career. First, I would like to thank my advisor, David Nelson, for his attentive guidance and mentorship. His keen physical intuition and careful thinking have been both an inspiration and a tremendous help. I cherish our discussions, which have been deeply instructive and a source of excitement throughout my research experience as a graduate student. I most often left his office bursting with renewed enthusiasm to pursue new avenues in our projects. His comments provided deep insights which I would sometimes only fully appreciate much later. I expect that what I have learned from our interactions will resonate frequently and for a long time into the future as I continue my physics career. I will also miss his thorough comments on my scientific writing. I'd also like to thank Andrew Murray and his group, especially John Koschwanez and Mary Wahl, for interesting discussions and for helping me learn and appreciate biology and various lab techniques. I'd also like to thank Kirill Korolev for his collaboration and helpful discussions.

I would also like to thank Subir Sachdev and Erel Levine for serving on my committee and for our fruitful interactions. In particular, Subir's class on quantum and classical phase transitions sparked my interest in directed percolation. My final project in that class directly influenced a significant portion of my thesis. I also fondly remember his presentation of a simulation of a quenched two-dimensional Ising model in Physics 262. His clear, enthusiastic lecturing style was an inspiration during my own experiences teaching Physics 262 with Erel Levine. The two semesters of teaching Statistical Physics with Erel helped cement my understanding of the subject, while letting me develop my teaching abilities. I am also grateful to all of the students who took the class. Our discussions, especially during office hours and sections, rekindled the excitement and joy I found when learning statistical mechanics for the first time. Also, this thesis would not have been completed without the kind support of Lisa Cacciabaudo, Jennifer Bastin, Carol Davis, and Sheila Ferguson.

A happy graduate school life would have been impossible without my friends. Thank you all! I would especially like to mention Tarek Anous, Jonathan Bittner, Rebecca Blum, Laura Celidonio, Thiparat Chotibut, Katie Deutsch, Brendan Enrick, Amy Gilson, Christy Han, Nick Hutzler, Greg Kestin, Andrej Kosmrlj, Kevin Mercurio, Wolfram Mobius, Melanie Müller, Lauren Nicolaisen, Jayson Paulose, Sarah Pekdemir, Cody

Polack, Doug Rubin, Katya Sergan, David Slochower, Nick van Meter, Eli Vizbal, Bryan Weinstein, and Stephan Zheng. A special thanks goes out to Elizabeth Petrik. Elizabeth and I navigated together through many late-night problem set sessions, practice talks, shared meals, and countless discussions of physics, philosophy, and life in general. She has greatly influenced my life in and outside of physics. Any successes I have had in either area are, to one extent or another, to her credit. This thesis has greatly benefited from her tireless editing and insightful comments.

I would also like to thank my teachers. I first realized I wanted to study theoretical physics after my AP Physics course with David Killius. At Kenyon College I had the opportunity to develop a strong base of physics knowledge. I would especially like to thank the professors with whom I worked on my first research projects: Sheryl Hemkin, Timothy Sullivan, Benjamin Schumacher, Frank Peiris. I would like to thank Royce Zia for introducing me to nonequilibrium statistical mechanics, for his friendly, inspiring mentorship and collaboration, and for two happy summers of research at Virginia Tech.

Finally, I am grateful to my family. They have encouraged and supported me throughout my education. I have been especially thankful for my mother's careful editing of my writing and for the many conversations about physics with my father. I thank my two siblings for keeping me grounded. I would also like to thank my grandparents in Ukraine. I cannot adequately express my gratitude for their caring support. They have enthusiastically followed my progress and have even flown across the Atlantic for each of my graduations, starting with the eighth grade.

To Sasha

Chapter 1

Introduction

1.1 Motivation and Random Walks

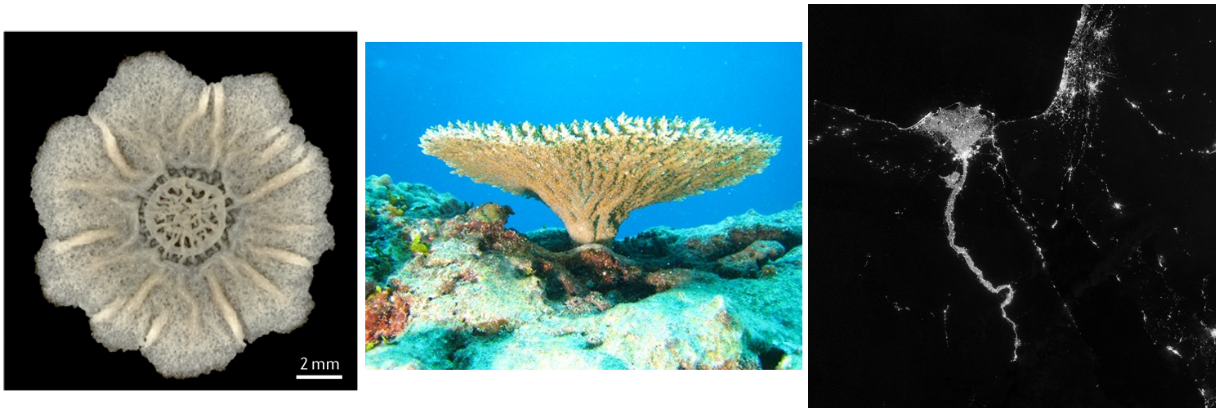


Figure 1.1.1: Complex spatial organization occurs in many diverse populations of organisms. In the left panel, a biofilm of *Bacillus subtilis* [1]. In the center, a coral flares out its branches in a complex, fractal arrangement. In the right panel, we see a satellite image of the Nile river valley taken at night (image credit: NASA Earth Observatory/Suomi NPP). The city lights show the heavily-populated areas.

Populations of living organisms exhibit a remarkable variety of complex spatial organization. Some particularly striking examples are shown in Fig. 1.1.1. In certain cases, such as the bacteria in the left panel of Fig. 1.1.1, the organization is generated dynamically. In other cases, such as the distribution of human life in the right panel, the organization is heavily influenced by aspects of the environment, such as its geographical features. In either case, the organization can profoundly impact how the population evolves and grows.

In this thesis, we will focus on simple, asexual organisms, and attempt to uncover general features of the interaction between spatial organization, growth, and evolution. Since we will be interested in the collective behavior of many organisms, the tools of statistical physics will be indispensable. In the spirit of the theory of phase transitions and critical phenomena, we will try to distill out the features of these interactions which are insensitive to the microscopic details of either the spatial organization or the organisms. This coarse-grained approach reveals the effects of macroscopic properties such as the spatial dimension and geometry of the population. It is also worth noting that living organisms, by consuming energy, reproducing, etc., are far from equilibrium. So far, we do not know of a universal description of non-equilibrium processes that would serve as the analogue of the Boltzmann and Gibbs treatment of equilibrium statistical mechanics. Hence, beyond the biological implications, our analyses provide important examples and test cases for developing fundamental theories of statistical states out of equilibrium.

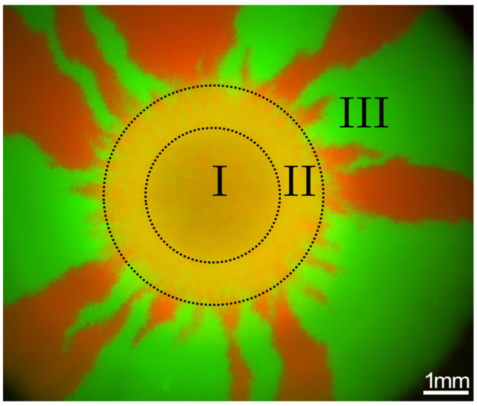


Figure 1.1.2: An image (adapted from Ref. [2]) of a range expansion of an initial disk-like inoculation of a 50-50 mixture of green-fluorescing and red-fluorescing (and otherwise identical) *E. coli* on an agar plate. The initial inoculation, or "homeland" is region I in the image. In region II, the red and green strains remain mixed, but start to fix locally. After some time, as the population grows outward in region III, genetic sectors of the green and red strains develop. This thesis is concerned with understanding the dynamics of such sectors, especially under the influence of an inflating circular (or spherical) population frontier.

Microbial populations are particularly well-suited for our purposes because they are easily cultured in the laboratory, and we may study their evolution as their generation time can be quite small ($\lesssim 1$ hour). Such populations require nutrients and space in order to grow, and this thesis will focus on populations expanding into new territory while consuming nutrients from the surrounding medium. Such growth is called a *range expansion*. These range expansions are ubiquitous, and describe the growth of cancerous tissue, biofilms, microbial colonies, the migrations of early humans, and many other populations. It is also easy to visualize the evolutionary dynamics in these populations through the use of fluorescent markers. In Fig. 1.1.2 we show a range expansion of *Escherichia coli* in which a mixture of two bacterial strains labelled with different fluorescent markers (but otherwise identical) is inoculated at the center of an agar plate. The population then grows outward with circular population fronts. Most of this thesis focuses on the dynamics of the colors when the strains can mutate into each other or have selective advantages. We also study the generalization of these dynamics to three dimensions.

Diffusion and absorbing states will serve as important, overarching mathematical themes in this thesis. First, the diffusion of a particle toward an absorbing location is a canonical example of a Markovian, non-equilibrium process. A Markovian process is one in which the system has no memory of its previous states, so transitions between system configurations $\mathcal{C}$ and $\mathcal{C}'$ are independent of the history of the system. Such a process is often most easily described via a master equation, which is the equation describing the evolution of the probability $P(\mathcal{C},t)$ of observing the system in a particular configuration $\mathcal{C}$ at time t. This equation reads

$$\partial_t P(\mathcal{C},t) = \sum_{\mathcal{C}'} \left[w(\mathcal{C}' \to \mathcal{C}) P(\mathcal{C}',t) - w(\mathcal{C} \to \mathcal{C}') P(\mathcal{C},t) \right], \tag{1.1}$$

where $w(\mathcal{C} \to \mathcal{C}')$ is the probability per unit time that a system in configuration $\mathcal{C}$ transitions into configuration $\mathcal{C}'$ and ∂_t is a partial derivative with respect to the time t.

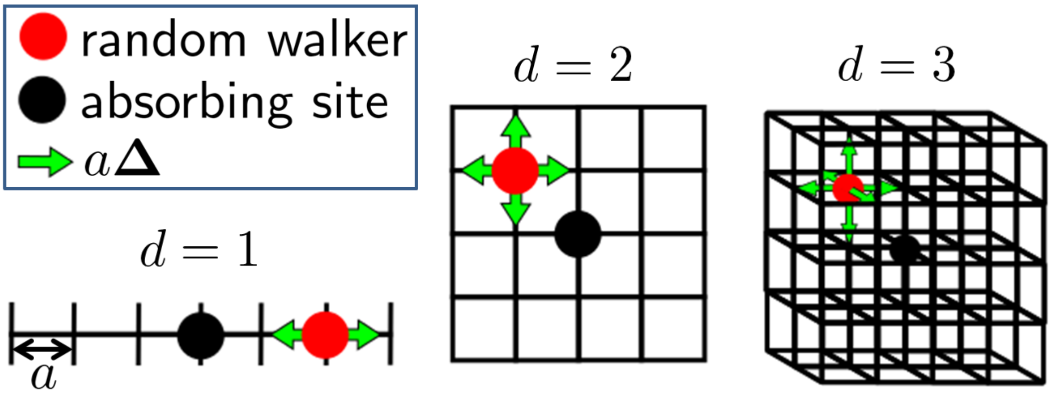


Figure 1.1.3: A random walk on the lattice $\mathbb{Z}^d$ with lattice spacing a and spatial dimension d. A small section of the lattice is shown for $d = 1, 2, 3$. We place an absorbing site at the lattice origin $\mathbf{r} = \mathbf{0}$.

Let's consider the master equation for a simple random walk model, where a walker moves around on a cubic lattice with lattice spacing a. We'll suppose that the walker steps to one of the nearest neighbor sites (chosen with equal probability) with a rate q. The stepping will be independent of the walker's history, so this is a Markovian processes describable by Eq. (1.1). Our configurations $\mathcal{C}$ will just be the walker position $\mathbf{r}$, which we take to be some $\mathbf{r} = a\,\mathbf{n}$, where $\mathbf{n} \in \mathbb{Z}^d$ is an n-tuple of integers in d dimensions. On this lattice there are $2d$ nearest neighbors for each lattice site. Hence, the rate at which the random walker hops to a particular nearest neighbor site must be $q/2d$. Thus, the master equation transition rates are

$$w(\mathbf{r} \to \mathbf{r} + a\boldsymbol{\Delta}) = \frac{q}{2d} \in (0,1), \tag{1.2}$$

where $\boldsymbol{\Delta} = (1,0,0,\ldots),(0,1,0,\ldots)$ are n-tuples pointing to the direction of one of the $2d$ nearest neighbor of the lattice site at $\mathbf{r}$. The master equation is

$$\partial_t P(\mathbf{r},t) = \frac{q}{2d} \sum_{\boldsymbol{\Delta}} \left[P(\mathbf{r} + a\boldsymbol{\Delta}) - P(\mathbf{r},t) \right]. \tag{1.3}$$

We also suppose that at the origin of the lattice, $\mathbf{r} = \mathbf{0}$, we have an absorber. Any random walker that reaches the origin will be absorbed, which introduces the boundary condition $P(\mathbf{0},t) = 0$. This discrete random walk model is illustrated in Fig. 1.1.3. We also recognize that the left-hand side of Eq. (1.3) is a finite-difference formula for the Laplacian operator ∇^2 in the limit $a \to 0$. Taking $a \to 0$ such that $D \equiv a^2 q/2d$ is finite, we find the probability obeys

$$\partial_t P(\mathbf{r},t) = D\nabla^2 P(\mathbf{r},t). \tag{1.4}$$

Equation (1.4) is the *diffusion* equation and describes how the probability density $P(\mathbf{r}, t)$ of finding a walker at position $\mathbf{r}$ and time t evolves over time. This equation and various generalizations will show up over and over again in this thesis. It is of interest to look at some solutions.

We now consider the concentration $\psi(\mathbf{r}, t)$ of N random walkers, defined by $\psi(\mathbf{r}, t) = NP(\mathbf{r}, t)$. This concentration clearly also satisfies Eq. (1.4). An illuminating initial condition is a uniform concentration $\psi(\mathbf{r}, t = 0) = \psi_0$. We may now solve Eq. (1.4) in the presence of the absorber. An invaluable mathematical tool in solving diffusion equations is the Laplace transform, which takes a function $f(x)$ to

$$F[s] \equiv \int_0^\infty f(x) e^{-sx} \, \mathrm{d}x. \tag{1.5}$$

The function $f(x)$ can be recovered by integrating s over a line parallel to the imaginary axis in the complex plane:

$$f(x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} F(x) e^{sx} \, \mathrm{d}x. \tag{1.6}$$

Taking a Laplace transform of Eq. (1.4) and assuming a spherically symmetric solution [so that $\psi(\mathbf{r}, t) = \psi(r = |\mathbf{r}|, t)$], we find that the diffusion equation reduces to

$$s\psi(r, s) - D \left[\frac{\partial^2}{\partial r^2} + \frac{d-1}{r} \frac{\partial}{\partial r} \right] \psi(r, s) = \psi(r, t = 0) = \psi_0. \tag{1.7}$$

Equation (1.7) is now just an ordinary differential equation in the variable r. We now solve it with an absorbing boundary condition. Our absorber is at the origin, and we'll assume it is an absorbing d-dimensional sphere with radius a. Then, our boundary condition is $\psi(r = a, s) = 0$. The solution to Eq. (1.7) is [3]

$$\psi(r,s) = \frac{\psi_0}{s}\left[1 - \left(\frac{r}{a}\right)^{1-d/2} \frac{K_{1-d/2}(r\sqrt{s/D})}{K_{1-d/2}(a\sqrt{s/D})}\right], \tag{1.8}$$

where $K_\nu(z)$ is a modified Bessel function of the second kind [4]. For $d = 1$ and $d = 3$, Eq. (1.8) is easily invertible because the Bessel functions reduce to $K_{\pm 1/2}(z) = e^{-x}\sqrt{\pi/2x}$. For $d = 2$, the inversion is more difficult and we'll have to make some approximations.

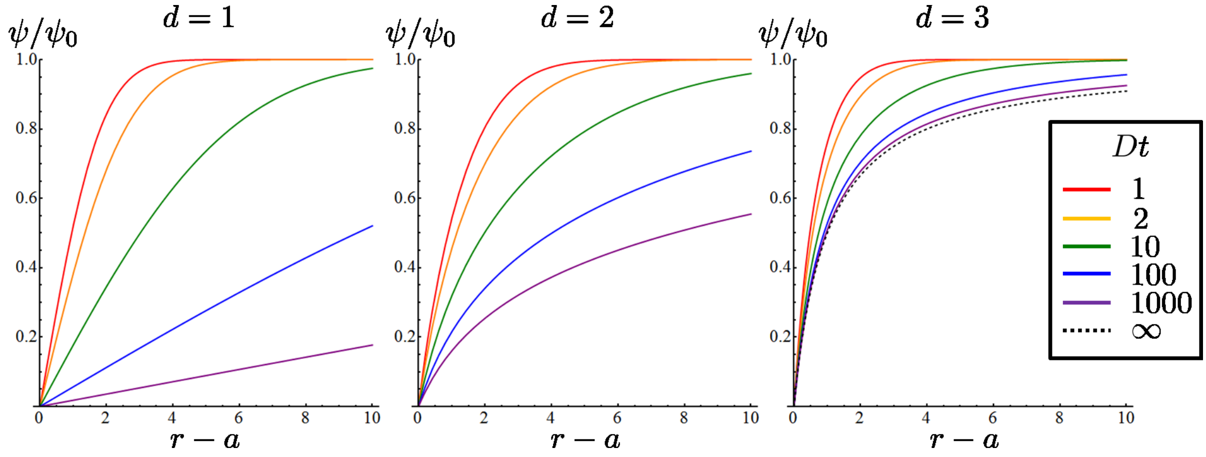


Figure 1.1.4: Solutions for the concentration of random walkers in a d-dimensional space with a spherical absorber of radius $a = 1$ at the origin for various Dt (measured in units of a^2). For $d = 1, 2$, the concentration eventually becomes completely depleted, as each random walker will eventually reach the absorber. For $d = 3$, the solutions approach a non-trivial steady-state satisfying Laplace's equation as $Dt \to \infty$.

Inverting Eq. (1.8) yields the full time-dependent solutions

$$\psi(r,t) = \begin{cases} \psi_0 \operatorname{erf}\left[\dfrac{r-a}{\sqrt{4Dt}}\right] & d = 1 \\ \dfrac{\psi_0}{2\pi i}\displaystyle\int_{\gamma - i\infty}^{\gamma + i\infty} \dfrac{e^{st}}{s}\left[1 - \dfrac{K_0(r\sqrt{s/D})}{K_0(a\sqrt{s/D})}\right] \mathrm{d}s \approx \dfrac{\psi_0 \ln(r/a)}{\ln\left(2e^{-\gamma_E}\sqrt{Dt}/a\right)} & d = 2\,, \\ \psi_0\left(1 - \dfrac{a}{r}\right) + \dfrac{a\psi_0}{r}\operatorname{erf}\left[\dfrac{r-a}{\sqrt{4Dt}}\right] & d = 3 \end{cases} \tag{1.9}$$

where we expanded the solution for $d = 2$ to leading order in $1/\ln[Dt/a^2] \ll 1$, an approximation valid for long times. In the $d = 2$ solution, $\gamma_E \approx 0.577$ is Euler's constant. These solutions are plotted in Fig. 1.1.4 for different times (time and space can be rescaled so that $D = a = 1$). Notice some crucial features of these solutions: For $d = 1$ and $d = 2$, we have $\psi(r, t \to \infty) \to 0$. This means that in these dimensions, the concentration around the absorber becomes depleted completely over time. This behavior is a manifestation of the fact that random walks are *recurrent* in $d = 1$ and $d = 2$. That is, random walkers will return to their initial position with probability 1 in one and two dimensions. However, in three dimensions, random walkers have a chance of escaping and never returning! In our problem, this means that any random walker that is some finite distance away from the absorber will eventually get absorbed for $d = 1, 2$. Conversely, for $d = 3$, the random walker may escape and never reach the absorber. Another important fact is that we may only take the $a \to 0$ limit without disturbing the concentration profile too much in the $d = 1$ case. Note that when we send $a \to 0$ in Eq. (1.9) for $d = 2, 3$, the absorbing point disappears completely and we find $\psi(\mathbf{x}, t) = \psi_0$! This is an important point: The problem of diffusion to an absorber requires a short distance (ultra-violet) cutoff for $d = 2, 3$. In other words, the absorber must be of non-zero size in order to find a meaningful solution to the problem.

Note that, over time, for $d = 3$ in Eq. (1.9), the concentration field $\psi(r, t)$ approaches a non-trivial *steady-state* form $\psi(r, t \to \infty) = \psi_0(1 - a/r)$ (dashed black line in right panel of Fig. 1.1.4). This form is independent of the diffusion constant D and is simply the solution to *Laplace's* equation

$$\nabla^2 \psi(\mathbf{r}) = 0. \tag{1.10}$$

If $\psi(\mathbf{r}, t)$ represents the concentration of a nutrient outside a metabolizing cell of radius a, the steady-state may be reached quite rapidly relative to the time scales of interest, such as the division time of the cell. We will exploit this fact in Chapter 2 where we study solutions to the simpler, time-independent Eq. (1.10) in three dimensions in order to understand how nutrients diffuse and are absorbed in clusters of cells. The two-dimensional case, $d = 2$, is marginal. It is the dimension at which a random walkers all eventually reach the absorber, but this process may take an extremely long time. Note that the depletion of the concentration field around the absorber grows only logarithmically in time in this case.

The $d = 3$ solution in Eq. (1.9) is clearly important for nutrient diffusion and absorption (see Chapter 2). However, the problem of diffusing walkers with an absorbing site will actually turn out to be relevant for our analysis of evolutionary dynamics, as well! It's already possible to see an aspect of this in region III of Fig. 1.1.2. It turns out that the boundaries between the differently colored sectors behave like random walkers. In the absence of natural selection, the random walks are unbiased. We shall see that the extinction of a genetic sector is related to a pair annihilation event between adjacent boundaries. This is how the absorbing events come into the picture and, eventually, the $d = 1$ solution of the diffusion equation in Eq. (1.9) (see Chapter 3). In fact, it turns out that the $d = 2$ solution will be relevant for *three*-dimensional range expansions, such as the growth of spherical microbial colonies in soft agar (see Chapter 4). Thus, the solutions in *all* dimensions $d = 1, 2, 3$ will be important, and the diffusion equation with an absorbing boundary condition is an overarching mathematical concept in this thesis. We discuss these ideas in more detail below.

This thesis is organized as follows: In Chapter 2 we analyze nutrient diffusion and absorption in clusters of cells, generalizing the simple diffusion problem discussed above to multiple, partially absorbing spherical absorbers. We use effective medium theory to calculate the characteristic screening length $\xi(\phi, \nu)$ of the nutrients inside the cluster, as a function of the cell packing fraction ϕ and the single-cell dimensionless absorption strength ν. In Chapter 3 we study two-dimensional range expansions, as might occur on the surface of a Petri dish. In Chapter 4 we extend our analysis to three dimensions, focussing on the survival probability of selectively advantageous mutations at the frontier. Finally, in Chapter 5, we consider range expansions of two species which provide either a growth enhancing or inhibiting effect to each other when they are in close proximity (via, e.g., a secreted nutrient or poison, respectively). We study how the spatial dimension and geometry of the population front influences the onset of the mutualistic phase in which the two species remain locally mixed during the range expansion. In the remainder of this chapter, we provide introductory material for each of these topics.

1.2 Nutrient Shielding

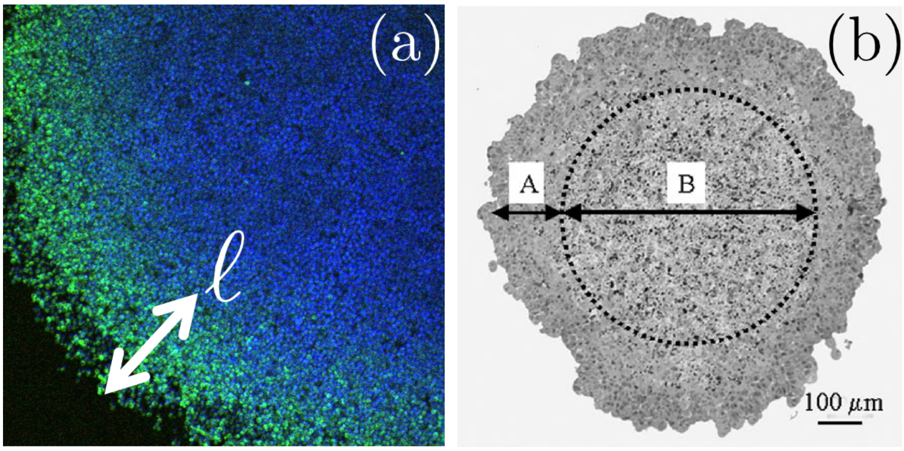


Figure 1.2.1: (a) A confocal microscope image (courtesy of J. H. Koschwanez) of the bottom of a roughly hemispherical yeast colony growing on the bottom of a glass well inside a nutrient-rich liquid medium with a glucose concentration $\psi_\infty \approx 4$ mM. The green fluorescent color indicates active cell divisions, in contrast to the blue colors, indicating regions where cell divisions have ceased. We see a screening length ℓ is present, indicating the fraction of the population that is receiving enough nutrients to grow. (b) A slice of a multicellular tumor spheroid of rat osteogenic sarcoma cells with a region A of viable cells and a region B of necrotic cells, which do not receive enough nutrients due to nutrient shielding. This image is taken from Ref. [5].

We study nutrient diffusion and absorption in clusters of cells in Chapter 2. These processes are important because many living organisms require a diffusive transport of a nutrient. For example, in human tissue, oxygen must diffuse to the cells. In yeast cells growing in a nutrient broth, sugar molecules must diffuse to the cell surface before this fuel can be taken up and metabolized. However, when the cells are closely packed in a cluster, the cells on the periphery of the cluster will “shield” the cells on the interior from the nutrients. This “nutrient shielding” can sometimes arrest the growth of the interior cells, as illustrated for yeast cells in Fig. 1.2.1(a). Nutrient shielding is also important in cancerous tissue which has not yet developed a vascular system (avascular tissue), which will only grow as long as enough nutrients are able to reach it from the surrounding medium via diffusion. When nutrient can no longer penetrate into the cancer interior, a necrotic region of dead cells develops, as illustrated in Fig. 1.2.1(b). In order to grow into a larger tumor, the cancerous tissue may circumvent the nutrient shielding by developing a vascular system to transport nutrients into the tissue interior.

Our simple analysis of random walkers in the previous subsection already tells us something about the length ℓ. The steady-state solution for the concentration of diffusing particles in the presence of the absorber for $d = 3$ [$t \to \infty$ in Eq. (1.9)] tells us that at long times, a non-trivial concentration profile develops around the cell, which we assume here is perfectly absorbing. This concentration profile will have a gradient, which according to Fick's law means that there will be a steady-state flux of particles into the absorber. The resulting flux density $j(r)$ into the cell may be integrated over the spherical absorber surface S to yield a total flux

$$J = \int_S j(r)\,\mathrm{d}\mathbf{r} = 4\pi a^2 j(a) = 4\pi a^2 D\psi_0 \left.\frac{\partial\psi}{\partial r}\right|_{r=a} = 4\pi a D\psi_0. \tag{1.11}$$

This result, first calculated by Smoluchowski in 1916 [6], has the interesting feature that the flux, perhaps counter-intuitively, grows with the radius a of the absorber, instead of a^2 as one might expect if it were to scale with the surface area.

If there is a collection of such absorbers, then, if they are sufficiently far apart, each one will have the same steady-state flux. We can then replace ψ_0 in Eq. (1.11) by the local steady-state concentration $\psi(\mathbf{r})$. Then, if the packing fraction ϕ of cells is dilute, so $\phi \ll 1$, these fluxes will deplete the concentration. This depletion can be interpreted as a local *chemical reaction.* At each point $\mathbf{r}$ in space, there will be a reaction between nutrient particles and the absorbing cells. The reaction will have a chemical kinetic equation, which in general looks like

$$\partial_t \psi(\mathbf{r}, t) = -R[\psi(\mathbf{r}, t)], \tag{1.12}$$

where $R[\psi(\mathbf{r}, t)]$ is the reaction rate at which the concentration is depleted. If the cells have a uniform number density $n = 3\phi/(4\pi a^3)$:

$$R[\psi(\mathbf{r}, t)] \approx 4\pi a D n \psi(\mathbf{r}, t) = \frac{3\phi D}{a^2}\,\psi(\mathbf{r}, t) \equiv k(\phi)\psi(\mathbf{r}, t), \tag{1.13}$$

where we have identified a *rate coefficient* $k(\phi) = 3\phi D/a^2$. We can extract a characteristic length ξ by recognizing that the reaction dynamics in Eq. (1.12) which deplete the nutrient will be balanced by the diffusive transport of the nutrient into the depletion zone. This balance will establish an exponentially decaying steady-state nutrient concentration inside the cell cluster. The characteristic exponential decay length can be found by simple dimensional analysis. The only combination of k and D that yields a length is

$$\xi = \sqrt{\frac{D}{k}} = \frac{a}{\sqrt{3\phi}}. \tag{1.14}$$

As expected, as the cluster becomes more and more dilute and $\phi \to 0$, the nutrient will be able to penetrate further into the cluster and the screening length $\xi \to \infty$.

The analysis sketched here is extended and its major deficiencies are discussed and remedied in Chapter 2. First, as mentioned above, we have also assumed the cells are perfectly absorbing. This is not realistic for many kinds of cells, and yeast cells in particular. Most glucose particles that hit a yeast cell as they random walk in the liquid medium are reflected back. Second, when the cells are crowded, like in the yeast colonies in Fig. 1.2.1(a), the deformations of the concentration field by one cell will clearly change the concentration field that nearby cells experience. We take these cell correlations into account in Chapter 2 by introducing an *effective medium theory*, in which we model the cell cluster as a homogeneous, reactive medium. Nutrients in this medium will experience some effective diffusion coefficient D_{eff} and an effective reaction rate $R_{\text{eff}}[\psi(\mathbf{r})]$, which will depend on the local nutrient concentration $\psi(\mathbf{r})$. We argue using a linear response theory that

$$R_{\text{eff}}[\psi(\mathbf{r})] \approx k_{\text{eff}}\psi(\mathbf{r}), \tag{1.15}$$

where k_{eff} is an effective rate coefficient. Hence, inside the cluster of cells, the nutrient concentration will no longer obey the simple diffusion equation. Instead, we will have an diffusion equation with a reaction term:

$$\partial_t\psi(\mathbf{r},t) = D_{\text{eff}}\nabla^2\psi(\mathbf{r},t) - R_{\text{eff}}[\psi(\mathbf{r},t)] \approx D_{\text{eff}}\nabla^2\psi(\mathbf{r},t) - k_{\text{eff}}\psi(\mathbf{r}). \tag{1.16}$$

The effective medium procedure discussed in detail in Chapter 2 then provides a self-consistent method of determining $D_{\text{eff}} = D_{\text{eff}}(\nu,\phi)$ and $k_{\text{eff}} = k_{\text{eff}}(\nu,\phi)$ as functions of the cell packing fraction ϕ and the absorption strength ν of each cell (discussed below). The effective medium approximation procedure is outlined in Fig. 1.2.2. Note that our simplified derivation with perfectly absorbing cells, interpreted in this effective medium language, yields $D_{\text{eff}} = D$ and $k_{\text{eff}} = 3\phi D/a^2$.

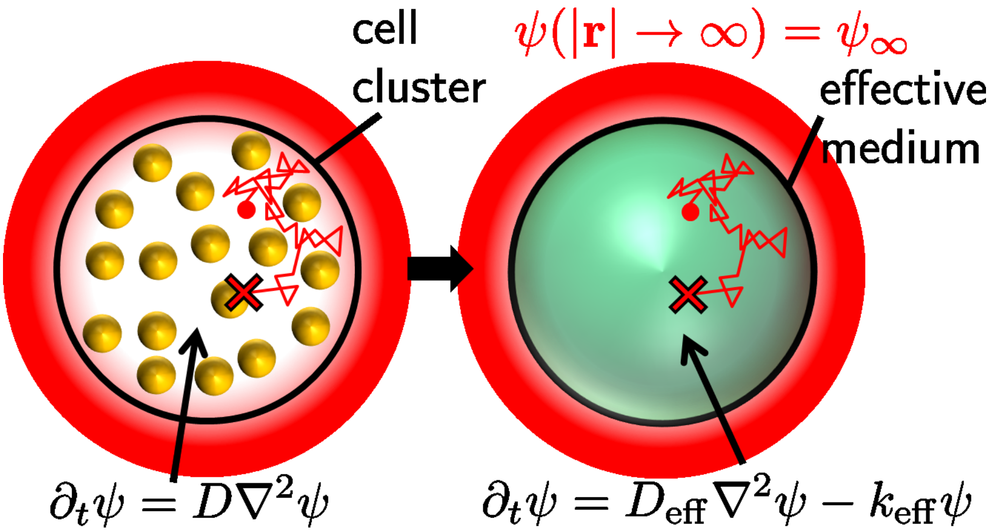

cell
cluster
$\psi(|\mathbf{r}| \to \infty) = \psi_\infty$
effective
medium
$\partial_t \psi = D \nabla^2 \psi$
$\partial_t \psi = D_{\text{eff}} \nabla^2 \psi - k_{\text{eff}} \psi$

Figure 1.2.2: A schematic of the effective medium approximation developed in Chapter 2 for nutrient absorption in a cluster, or colony, of cells immersed in a nutrient bath with some steady concentration ψ_∞ (red halos) far from the cluster. The cluster is approximated as a homogeneous “effective medium” in green. In the interstitial parts of the cluster, nutrient particles (red dot) perform random walks (red line) until they react with a cell (orange spheres) and get absorbed (red cross). The nutrient concentration field $\psi = \psi(\mathbf{r}, t)$ then obeys a simple diffusion equation with absorbing boundary conditions at the cell surfaces. In the effective medium approximation, the concentration obeys a reaction-diffusion equation instead. The nutrient particles diffuse with an effective diffusion constant D_{eff} and spontaneously disappear (react) with a rate coefficient k_{eff}, which will depend on the arrangement of cells in the cluster and the individual cell absorption properties.

In Chapter 2, we will allow for cells to be partially absorbing by generalizing the boundary condition at the cell surface S. As discussed above, perfect absorbers have $\psi(\mathbf{r}) = 0$ for all $\mathbf{r} \in S$. A more general condition, called a *radiation* boundary condition, states that the concentration is proportional to the diffusive flux of nutrients at the surface, which itself must be proportional to the concentration gradient:

$$D\, \hat{\mathbf{n}} \cdot \nabla \psi(\mathbf{r}) = \kappa\, \psi(\mathbf{r}) \quad \text{for all } \mathbf{r} \in S, \tag{1.17}$$

where $\hat{\mathbf{n}}$ is a unit-vector normal to the cell surface at $\mathbf{r}$ and κ is an absorption strength.

We can define a dimensionless absorption strength ν by introducing the cell radius a:

$$\nu \equiv \frac{\kappa a}{D}. \tag{1.18}$$

Note that as $\nu \to 0$, $\hat{\mathbf{n}} \cdot \nabla\psi(\mathbf{r}) \to 0$ and no nutrients will enter the cell. Conversely, as $\nu \to \infty$, we recover the perfect absorber condition $\psi(\mathbf{r}) = 0$. We may now use these partially absorbing cells to find the appropriate rate coefficient $k_{\text{eff}}(\phi, \nu)$, which will now depend on ν as well as the cell packing fraction ϕ. In the effective medium theory, we also assume that the nutrient diffusion is modified inside the cell cluster, due to the confinement of cells in the cell interstices prior to absorbtion. Hence, we will have a "renormalized" diffusion coefficient $D_{\text{eff}}(\phi, \nu)$ which will also depend on the packing fraction and the absorption strength.

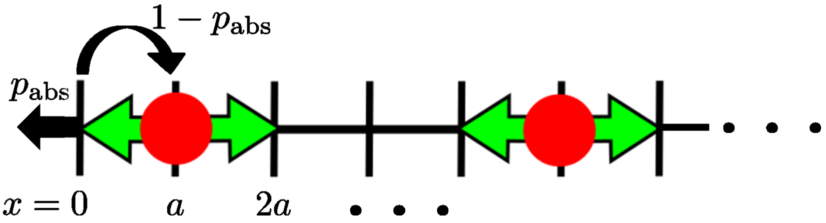


Figure 1.2.3: A model of a partially absorbing boundary condition in one dimension. The red nutrient particles perform random walks until they reach the set at $x = a$. Then, if the particle hopes left to $x = 0$, it gets immediately absorbed with probability p_{abs} or reflected with probability $1 - p_{\text{abs}}$.

What is the meaning of ν and how might the radiation boundary condition in Eq. (1.17) be derived? The master equation formalism in Eq. (1.1) provides a simple, illuminating derivation. Suppose for simplicity we have a random walk in one dimension. Then, suppose there is a partially absorbing cell wall at location $x = a$. We will assume that as soon as a nutrient particle at this location attempts to penetrate the barrier to move to $x = 0$, it will get absorbed only with probability p_{abs}. Otherwise, it gets reflected back (which must happen with probability $1 - p_{\mathrm{abs}}$), as illustrated in Fig. 1.2.3. This model is also analyzed in detail in Refs. [3, 7].

We can re-write the master equation in Eq. (1.3) for $d = 1$ by recognizing the left-hand side as a difference between fluxes:

$$\partial_t P(x,t) = \frac{q}{2}\left[P(x+a,t) + P(x-a,t) - 2P(x,t)\right] = J(x-a,t) - J(x,t), \tag{1.19}$$

where we have identified a nutrient flux $J(x,t)$ into site x at time t:

$$J(x,t) = -\frac{q}{2}\left[P(x+a,t) - P(x,t)\right]. \tag{1.20}$$

Equations (1.19) and (1.20) are discrete versions of the continuity equation ($\partial_t P + \nabla \cdot \mathbf{J} = 0$) and Fick's first law ($\mathbf{J} = -D\nabla P$), respectively. Let's now look at what happens to the flux to the origin, $J(x = 0, t)$. We know that particles will enter this site only if they hopped from $x = a$ and were absorbed. Hence,

$$J(x=0,t) = -\frac{p_{\mathrm{abs}} q}{2} P(x=a,t), \tag{1.21}$$

where the minus sign means that the flux is coming from the right (toward the negative x axis in Fig. 1.2.3). We now equate this result for $J(x = 0, t)$ to the general expression from the discrete Fick's law in Eq. (1.20). Upon taking the small a limit and Taylor expanding, we have

$$p_{\text{abs}} P(a,t) = P(a,t) - P(0,t) \tag{1.22}$$

$$\frac{P(a,t) - P(0,t)}{a} \approx \left.\frac{\partial P(x,t)}{\partial x}\right|_{x=0^+} = \frac{p_{\text{abs}}}{a(1-p_{\text{abs}})} P(0,t) \equiv \frac{\kappa}{D} P(0,t). \tag{1.23}$$

Equation (1.23) is the desired radiation boundary condition! Hence, in this discrete model, the dimensionless absorption strength ν is very simply related to the probability of absorption p_{abs}:

$$\nu = \frac{\kappa a}{D} = \frac{p_{\text{abs}}}{1-p_{\text{abs}}}. \tag{1.24}$$

In fact, there are many ways of deriving this partially absorbing radiation and it is quite general. We give another derivation in Chapter 2 where we model the cell wall as a potential barrier.

Nutrient shielding will not only be important for the population growth dynamics, but also for evolution. In particular, only the viable cells near a frontier will be able to propagate their genetic information to future generations. For example, only the yeast cells in the green rim in Fig. 1.2.1(a) and the cancer cells in the region "A" in Fig. 1.2.1(b) will form lineages by cell division. Hence, nutrient shielding will determine the effective population size for the evolutionary dynamics. It might also completely arrest the spread of a population. For example, when combined with pressure from the external, healthy tissue, nutrient shielding may arrest the spread of cancer cells. The necrotic region "B" in Fig. 1.2.1(b) will result in a volume loss as cells undergo apoptosis and are flushed out of the cluster. This volume loss can be balanced by cell division and growth in the viable region, resulting in a kind of "treadmilling" effect where cells at the tumor surface constantly divides while remaining the same radius. We consider the evolutionary implications of these growth regimes in detail in Chapters 3 and 4.

1.3 Evolution of Range Expansions with Inflating Frontiers

In Chapter 3 and 4 of this thesis, we will study range expansions of populations with some characterstic size, or radius, $R(t)$ at time t. In general, $R(t)$ can have a complicated time dependence. If all cells in the population are dividing, then we generally expect that the range expansion size grows exponentially like $R(t) \sim e^{\lambda t}$, where λ is the cell growth rate. However, effects such as nutrient shielding, discussed in Chapter 2, can slow down this growth. For example, small avascular tumors will often grow approximately linearly in time, with $R(t) \sim t$ until they reach a maximum size R_{max}. After this point they will "treadmill" as discussed above. We will be able to capture many biologically growth regimes by studying a general power-law growth

$$R(t) = R_0 \left[1 + \frac{t}{t^*}\right]^{\Theta}, \tag{1.25}$$

where t^* is a cross-over time and Θ is a growth exponent. In general, both of these paramters would have to be fit to experiments. However, we already know that the linearly inflating case $\Theta = 1$ and the "treadmilling" regime $\Theta = 0$ will be relevant. These regimes are indicated for tumor growth in Fig. 1.3.1(a). Also, since Θ decreases from 1 to 0 over the growth history of the tumor, intermediate cases with $0 < \Theta < 1$ are also of interest.

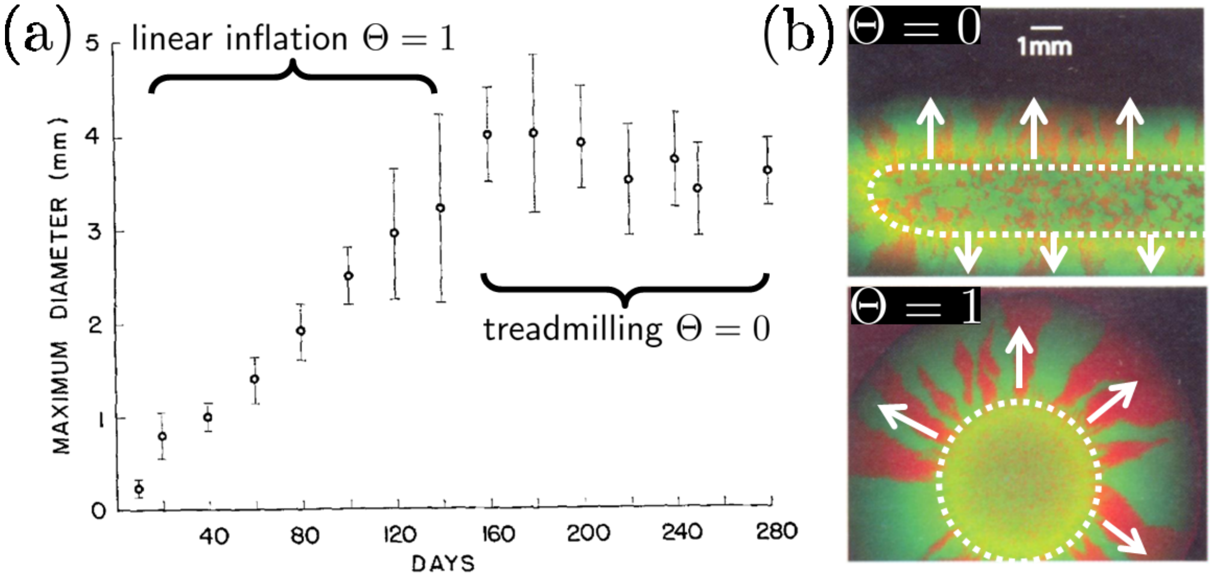


Figure 1.3.1: (a) Experimental results from Ref. [8] for the average diameter and standard deviation (error bars) of 70 tumor spheroids cultured in agar. There is a clear linear inflation regime in which the tumor spheroid diameter grows approximately linearly in time, $R(t) \sim t^{\Theta}$ with $\Theta = 1$. The tumors eventually saturate at some constant maximum diameter, corresponding to a "treadmilling," $\Theta = 0$, growth regime. (b) Two-dimensional expansions of *E. coli* on agar plates, taken from Ref. [9]. In the top panel, an expansion off of a razor blade inoculation is shown. This corresponds to a $\Theta = 0$ growth regime because the population front along the horizontal direction in the micrograph remains the same size. Conversely, in the bottom panel, a circular expansion has an inflating frontier which increases linearly in time, so that $\Theta = 1$.

The $\Theta = 0$ and $\Theta = 1$ regimes are also important in range expansions of microbes. A $\Theta = 0$ expansion may be realized by inoculating the microbes in a straight line on an agar plate with a razor blade, for example. Then, as the microbial colony grows out, the long edge of the razor blade inoculation will stay roughly the same size (provided that the population front does not undulate too much), similar to treadmilling. Conversely, if we inoculate the microbes in a disk-shaped "homeland" using a pipette, say, the resulting colony will grow out radially with circular population fronts with radii that grow approximately linearly in time, i.e., $R(t) \sim t$. These two cases are illustrated in Fig. 1.3.1. We study two-dimensional range expansions like this in Chapter 3 using simulations and analytic techniques. A realization of a circular expansion in a simulation is shown in Fig. 1.3.2(a). For these two-dimensional expansions, we focus in Chapter 3 on populations that grow out as *monolayers* of cells with a population front that is a single cell wide. This limit is interesting because the local effective population size for the evolutionary dynamics is extremely small and small number fluctuations, or "genetic drift" will play an important role.

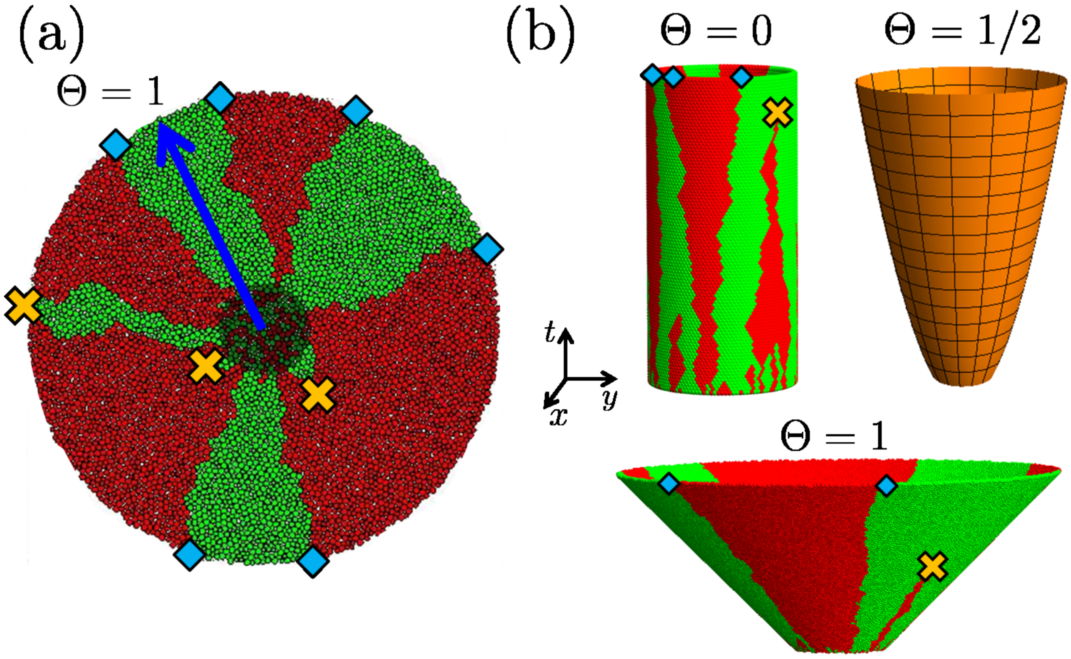


Figure 1.3.2: (a) A simulation of a circular range expansion with a radius $R(t) \sim t$ increasing linearly in time ($\Theta = 1$). The "homeland" consists of mixed, selectively neutral red and green species. We indicate boundaries between red and green genetic sectors by blue diamonds. Yellow crosses indicate pair annihilation events between these boundaries. (b) Range expansions at different inflation exponents Θ can be realized by growing cells on surfaces with different curvatures. These surfaces are illustrated in the figure, along with some example range expansion evolutions for the "treadmilling" $\Theta = 0$ case and the linearly inflating $\Theta = 1$ case. Note that the cylindrical $\Theta = 0$ is similar to a razor blade inoculation, but with periodic boundary conditions.

Another way of understanding the different growth exponents Θ is to consider a population, such as a biofilm, growing over a surface. Then, provided that the cells grow at a uniform surface density, each exponent Θ will correspond to a different kind of surface. We show some examples in Fig. 1.3.1(b). The $\Theta = 0$ and $\Theta = 1$ cases are realized in simulations in Chapter 3. However, our analytic theory can be done for arbitrary Θ, such as $\Theta = 1/2$. Chapters 3 and 4 focus on the evolutionary dynamics in populations with these different growth regimes in two and three dimensions, respectively. In two dimensions, we already know that random walkers will again play an important role. Note from Fig. 1.3.2(a) that a mixture of two neutral strains will form genetic sectors. The genetic sector boundaries will wander as the green and red cells divide and push each other at the frontier. If two such sector boundaries meet, they will annihilate as illustrated in the figure. Hence, we again have diffusing random walkers with absorption! For $\Theta > 0$, however, we have the additional complication that the random walkers live on a population frontier that is increasing in size. Hence, they are random walkers in an "inflationary universe." Note that these random walkers live in an effectively one-dimensional space. This illustrates the concept of *dimensional reduction.* Although our population expansion is two-dimensional, the population frontier at which all of the evolutionary dynamics happens is one-dimensional. The radial direction perpendicular to the front, then, may be treated as a time-like direction: as the population frontier sweeps out radially, the population that is left behind represents an evolutionary *history* of the population.

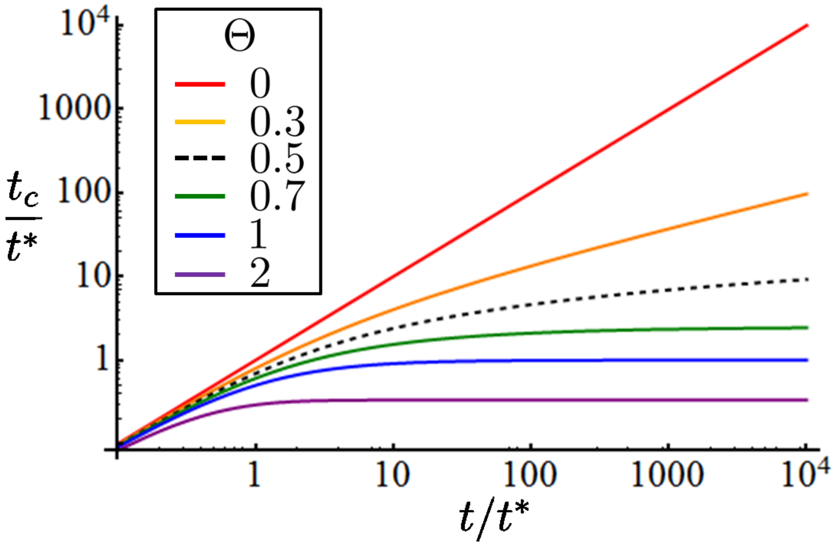


Figure 1.3.3: A plot of the transformed time coordinate t_c, relevant for diffusion in an inflationary environment that grows according to Eq. (1.25) for various growth exponents Θ. When $\Theta = 0$, $t_c = t$. Note that for $t \ll t^*$, $t_c \approx t$. However, for $t \gg t^*$, t_c approaches a finite value for all $\Theta > 1/2$. Otherwise, if $\Theta \leq 1/2$, $t_c \to \infty$ as $t \to \infty$. This means that a single genetic sector will take over for all $\Theta \leq 1/2$. However, for $\Theta > 1/2$, there is a non-zero probability that multiple genetic sectors survive at long times.

The effects of the inflation can be understood by considering the angular distance ϕ between two genetic boundaries along the frontier of a circular expansion with radius $R(t)$. We'll ignore the periodicity in ϕ for simplicity by analyzing small genetic sectors for which $\phi \ll 2\pi$, so we can extend our range to $\phi \in (-\infty, \infty)$, with negligible corrections. For monolayer growth, after a generation time τ_g a green or red cell near the boundary will divide and may push its daughter cell into the opposite cell color territory. Hence, after time τ_g, ϕ may change by an amount $\Delta\phi \sim \pm a/R(t)$. If the circular expansion front remains locally flat due to, say, surface tension effects, this boundary jostling (on either side of the sector) will happen every generation and ϕ will perform a random walk! Also, note that if at any time the angle ϕ goes to zero, it will stay zero because the sector will have gone extinct. Thus, $\phi = 0$ is an absorbing state. Hence, we may apply the master equation approach of Eq. (1.1) to find a diffusion equation for the probability $P(\phi, t)$ of observing a particular ϕ at time t. The effects of the inflating boundary are then captured by a time-dependent diffusion coefficient, whose time dependence may be transformed away with a variable change:

$$\partial_t P(\phi, t) = \frac{D}{[R(t)]^2} \frac{\partial^2 P(\phi, t)}{\partial \phi^2} \quad \Rightarrow \quad \partial_{t_c} P(\phi, t_c) = \frac{D}{R_0^2} \frac{\partial^2 P(\phi, t_c)}{\partial \phi^2}, \tag{1.26}$$

where $D \sim a^2/\tau_g$ and we have introduced a new time variable

$$t_c \equiv t_c(t) = \frac{(1 + t/t^*)^{1-2\Theta} \left[(1 + t/t^*)^{2\Theta - 1} - 1\right]}{2\Theta - 1} t^*, \tag{1.27}$$

where t^* is the crossover time appearing in Eq. (1.25). This expression for t_c at various Θ is plotted in Fig. 1.3.3.

The expression for t_c in Eq. (1.27) already reveals quite a bit about the effects of the inflating frontier. We already mentioned that random walks are recurrent in one dimension. This means that if $R(t) = R_0$ is a constant, we expect ϕ to eventually reach zero. Thus, all genetic sector boundaries pair annihilate at long times and a single sector reaches fixation. However, for this to happen with probability 1, we have to wait an infinitely long time ($t \to \infty$). In an inflationary scenario, ϕ will perform a random walk, but in the transformed time coordinate t_c. Hence, to ensure that ϕ eventually vanishes, we require that $t_c \to \infty$ in the diffusion equation in Eq. (1.26) as $t \to \infty$. For populations with $\Theta = 0$ (razor blade inoculation, or range expansion on a cylinder), $t_c = t$ and everything is consistent. However, note from Fig. 1.3.3 that $t_c \to \infty$ *for all* $\Theta \leq 1/2$. This means that a single sector must eventually reach fixation for all $\Theta \leq 1/2$, as well. What happens for $\Theta > 1/2$? In this case, as $t \to \infty$, $t_c \to t^*/(2\Theta - 1)$

approaches a finite limit! Thus, there will always be some residual probability that ϕ never reaches zero as $t \to \infty$. Hence, the population may have multiple genetic sectors at long times, as illustrated for $\Theta = 1$ in the circular expansions in Fig. 1.3.1(b) and Fig. 1.3.2(a). Our model for $R(t)$ in Eq. (1.25) can also describe an exponentially expanding population, which may occur when nutrient shielding is absent. We simply take $\Theta, t^* \to \infty$ such that $t_R = t^*/\Theta$ is a finite characteristic time. Then, $R(t) = R_0 e^{t/t_R}$ and $t_c(t) = t_R(1 - e^{-2t/t_R})/2$. This time approaches the finite limit $t_c \to t_R/2$ as $t \to \infty$. The average sector number [for a slightly different general power-law growth of $R(t)$] is calculated in Chapter 2.

What happens in three-dimensions? This case is considered in detail in Chapter 4. Unlike the two-dimensional case, a mapping of sector boundaries to random walks is not available. Indeed, the population front will be effectively two-dimensional so that the genetic sectors now form "patches." These patches will have complicated shapes. Hence, we need a different tool to analyze the evolutionary dynamics. To understand this, let us generalize our discussion a bit and allow the green strain to have a selective advantage over the red strain. Then, if the growth rates of the green and red strains are Γ_g and Γ_r, respectively, then we can define a selection coefficient via

$$s = \frac{\Gamma_g - \Gamma_r}{(\Gamma_g + \Gamma_r)/2} \geq 0. \tag{1.28}$$

Then, to further generalize our dynamics, we also assume that green cells can mutate to red cells with some rate μ (see next section for a biological justification).

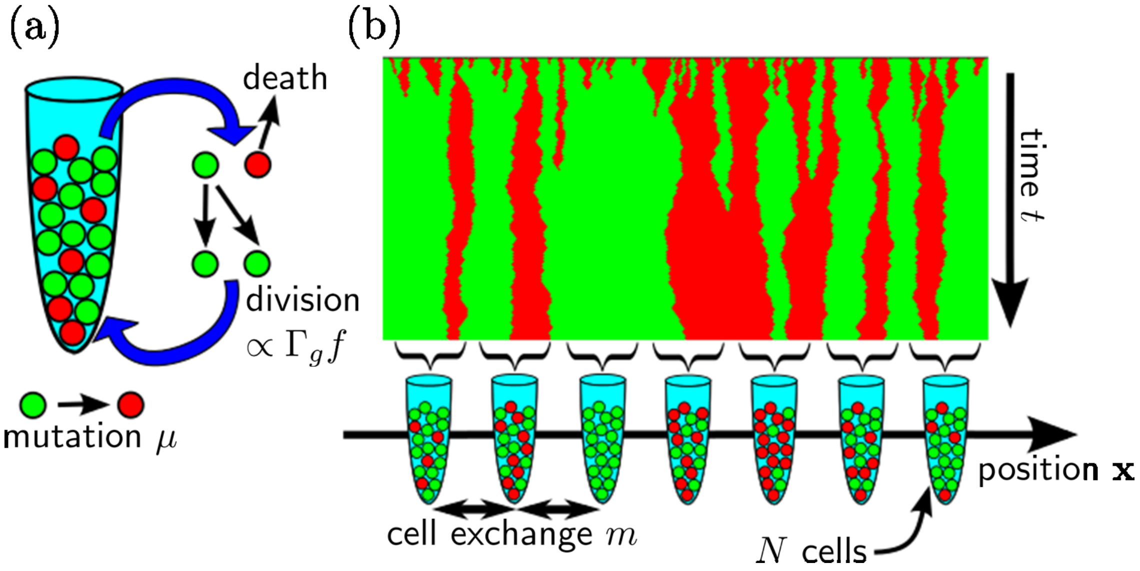


Figure 1.3.4: (a) We can model a well-mixed population of N asexual cells with the Moran model of population genetics [10]. At each time step, we pick one cell to divide with a probability proportional to its growth rate ($\Gamma_{g,r}$ for green and red cells, respectively) and another one to die. Cells can then mutate with some rate μ from green to red. (b) A simulation of a range expansion with a uniform, flat front that is just a single cell wide. The evolutionary dynamics can be related to a stepping-stone model of a series of connected well-mixed populations via a coarse-graining. In addition to the Moran model dynamics, each "stepping stone" population exchanges a fraction m of cells with their nearest neighbors every generation.

The main tool we use to study the evolutionary dynamics of the competing red and green species is *stochastic calculus*. In particular, we derive an evolution equation for the coarse-grained fraction $f(\mathbf{x}, t)$ of the green species at a point $\mathbf{x}$ along the population frontier at time t. The derivation involves mapping the range expansion to a stepping-stone model [11] in which adjacent well-mixed populations at each site interact by exchanging organisms. Upon coarse-graining, each group of N adjacent cells along the population front is regarded as a well-mixed population. Each well-mixed population [see Fig. 1.3.4(a)] will have its own evolutionary dynamics. We then allow for exchange of a fraction m of the N cells between neighboring populations, as shown in Fig. 1.3.4(b).

The exchange can be performed in many different ways [12], but always yields a diffusion term $D\nabla^2 f(\mathbf{x},t)$ in the evolution equation for $f(\mathbf{x},t)$, where $D \approx ma^2/\tau_g$, where a is the cell diameter and τ_g a generation time [12, 13]. To match our coarse-grained results to simulations of range expansions with population fronts of one cell width, we will have to take $N \to 1$, which maximizes the importance of number fluctuations at the frontier. The exchange dynamics are also fixed in this case with $m = 1$ (see Chapter 3 for more details). In this case, we also have $D \approx a^2/\tau_g$. We model the evolutionary dynamics in each well-mixed population using a *Moran* model [10] illustrated in Fig. 1.3.4(a). In this model, we pick two cells at random from the N cells in a test tube (or "deme") at every time step $\Delta t = \tau_g/N$ (τ_g is a generation time). One of the two cells divides with a probability proportional to its growth rate Γ_g or Γ_r if it is green or red, respectively. (For example, a green cell is selected to reproduce with probability $\Gamma_r/(\Gamma_r + \Gamma_g)$ if we pick a green and red cell pair.) The other cell is then thrown out of the population. If a green cell divides, one daughter cell can mutate to red with some probability μ. We will motivate the latter rule in the next section. For now, note that these rules provide a plausible model of the stochastic birth-death dynamics inside the test-tube that conserves the total population number N. (We will eventually allow organisms to exchange between neighboring test tubes.)

Consider the dynamics of $f(t)$, the fraction of green cells in a single well-mixed population. We may once again appeal to the powerful master equation formalism in Eq. (1.1) to write down an evolution equation for $P(f,t)$, i.e., the probability of observing a fraction f of green cells in a particular test-tube at time t. The master equation is particularly simple because the number of green cells in the population changes by at most *one* cell per unit time. So, in this *one-step* Master equation [14] there will be two relevant rates, call them u_n and d_n, which correspond to the increase ("step up") and decrease ("step down") of n by one cell, respectively. Then, our configurations are just the number of green cells $n \in \{1, 2, \ldots, N\}$, or, equivalently the fractions $f = n/N$. The Master equation for $P_n \equiv P(n,t)$ reads

$$\partial_t P_n = u_{n-1} P_{n-1} + d_{n+1} P_{n+1} - (u_n + d_n) P_n. \tag{1.29}$$

This one step Master equation is more conveniently written in terms of a step operator $\mathbb{E}$ that takes any function ψ_n and steps it forward via $\mathbb{E}\psi_n \equiv \psi_{n+1}$. Equation (1.29) becomes

$$\partial_t P_n = (\mathbb{E} - 1)\, d_n P_n + \left(\mathbb{E}^{-1} - 1\right) u_n P_n. \tag{1.30}$$

We now promote u_n, d_n and P_n to continuous functions of both n and t and consider the continuous version of the $\mathbb{E}$ operator, defined via the following Taylor expansion:

$$\psi(n+1) = \psi(n) + \frac{\partial \psi}{\partial n} + \frac{1}{2}\frac{\partial^2 \psi}{\partial n^2} + \ldots = e^{\frac{\partial}{\partial n}} \psi(n) = \mathbb{E}\psi(n). \tag{1.31}$$

Upon expanding $\mathbb{E}$ to second order in the derivatives, we have

$$\begin{aligned} \partial_t P_n &= \left(-\frac{\partial}{\partial n} + \frac{1}{2}\frac{\partial^2}{\partial n^2}\right) u_n P_n + \left(\frac{\partial}{\partial n} + \frac{1}{2}\frac{\partial^2}{\partial n^2}\right) d_n P_n \\ &= \frac{\partial}{\partial n}\left[\left(d_n - u_n\right) P_n\right] + \frac{1}{2}\frac{\partial^2}{\partial n^2}\left[\left(u_n + d_n\right) P_n\right]. \end{aligned} \tag{1.32}$$

Equation (1.32) is a *Fokker-Planck* equation. For the Moran model dynamics described above (see Fig. 1.3.4),

$$\begin{cases} u_n = \tau_g^{-1} \dfrac{\Gamma_g f(1-f)}{\Gamma_g f + \Gamma_r(1-f)} \\ d_n = \tau_g^{-1} \dfrac{\Gamma_r f(1-f)}{\Gamma_g f + \Gamma_r(1-f)} + \dfrac{\mu}{\tau_g} f \end{cases}, \tag{1.33}$$

where $f = n/N$. We can now move to a stochastic differential equation description of the dynamics.

Although the evolution equation for $P_n(t)$ (Eq. (1.32)) is useful, it is difficult to generalize to the spatial stepping-stone model illustrated in Fig. 1.3.4(b), including a cell exchange term. As mentioned above, we would like to include cell exchange as a diffusion term $D\nabla^2 f(\mathbf{x}, t)$. To do this, we need an equation that describes the dynamics of n *itself*. Consider first $n(t)$ for a single population. The evolution of $n(t)$ is *noisy* and it must follow the probability distribution that satisfies the Fokker-Planck equation [Eq. (1.32)]. It turns out that the appropriate evolution equation is

$$\partial_t n(t) = u_n - d_n + \sqrt{u_n + d_n}\,\xi(t) \equiv -R[n(t)] + \sqrt{V[n(t)]}\xi(t), \tag{1.34}$$

where we have emphasized that u_n and d_n may depend on $n(t)$ by introducing new functions $R[n(t)] = d_n - u_n$ and $V[n(t)] = u_n + d_n$. The Gaussian white noise $\xi(t)$ is constructed by drawing a random real number from a Gaussian distribution (with zero mean and variance equal to 1) at each instant of time. These random variables are drawn independently at each time, so $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$. Thus, $\sqrt{V[n(t)]}$ is the magnitude of the noise and since $\langle \xi(t) \rangle = 0$, it may be interpreted as a standard deviation, or square root of the variance. Written in this way, Eq. (1.34) has the form of a noisy chemical kinetic equation. Indeed, note that the only difference between Eq. (1.34) and Eq. (1.12) is the presence of the term proportional to $\xi(t)$ in Eq. (1.34). This noise comes from the discrete nature of the organisms or "chemical species," and has important consequences for the dynamics of $n(t)$.

The noise $\xi(t)$ is a highly stochastic function. In fact, its integral is a Wiener process $W(t)$, which is everywhere continuous and nowhere differentiable:

$$W(t) \equiv \int_0^t \xi(t')\,\mathrm{d}t' \tag{1.35}$$

The violently stochastic nature of $\xi(t)$ introduces an ambiguity when we try to integrate the noise term in Eq. (1.34). To see this, suppose we partition the interval $[0,t]$ into n slices, and try to define the integral of the noise term as follows:

$$\int \sqrt{V[n(t)]}\,\xi(t)\,\mathrm{d}t \overset{?}{=} \text{“}\lim_{n\to\infty}\text{”} \sum_{i=1}^{n} \sqrt{V[n(\tau_i)]}\,[W(t_i) - W(t_{i-1})], \tag{1.36}$$

where $\tau_i \in [t_{i-1}, t_i]$. We run into two issues here. First, $W(t_i)$ and $W(t_{i-1})$ are random variables and we have to specify how exactly we take the $n \to \infty$ limit of a sequence of such random variables (hence the quotes around the limit). Second, we need to decide which particular τ_i to pick. It turns out that the limiting procedure can be defined unambiguously [15]. The choice of τ_i, however, turns out to be quite subtle. In fact, different choices lead to different statistical properties of the process $n(t)$! Different evolution equations for $P_n(t)$ arise depending on which τ_i we pick. Two popular choices are

$$\tau_i = \begin{cases} t_{i-1} & \text{Îto interpretation} \\ \dfrac{t_i + t_{i-1}}{2} & \text{Stratanovich interpretation} \end{cases} . \tag{1.37}$$

The interpretation that yields the correct equation for population genetics [Eq. (1.32)] is the Îto interpretation, where $\tau_i = t_{i-1}$. This choice is plausible because we know that the noise is ultimately caused by the discrete nature of the dividing organisms. Hence, one expects that the noise contribution when we evolve from $n(t_{i-1})$ to $n(t_i)$ during some discrete division step will depend on the state of the system before any division occurs, i.e., $n(t_{i-1})$. Note, however, that it is possible to introduce an additional drift term to Eq. (1.32) and to use the Stratanovich (or some other choice for τ_i), instead [15].

Armed with the evolution equation for $n(t)$ for a single population, we now allow for an $n(t)$ variable to live at all points $\mathbf{x}$ along the population front and introduce a population *number density* $n(\mathbf{x}, t)$. We introduce a stochastic noise $\xi(t)$ at every position, which will clearly be uncorrelated between the populations at various $\mathbf{x}$. Hence, we define a new spatio-temporal noise $\xi(\mathbf{x}, t)$ such that

$$\langle \xi(\mathbf{x}, t) \rangle = 0 \quad \text{and} \quad \langle \xi(\mathbf{x}, t)\xi(\mathbf{x}', t') \rangle = \delta(\mathbf{x}' - \mathbf{x})\delta(t - t'). \tag{1.38}$$

Upon moving to the reduced variable $f(\mathbf{x}, t) = n(\mathbf{x}, t)/N$ (the deme size N is assumed fixed), we then add the cell exchange term and find the evolution equation

$$\partial_t f(\mathbf{x}, t) = D\nabla^2 f(\mathbf{x}, t) - R[f(\mathbf{x}, t)] + \sqrt{V[f(\mathbf{x}, t)]}\, \xi(\mathbf{x}, t), \tag{1.39}$$

where we find from Eq. (1.33):

$$R[f(\mathbf{x},t)] = \tau_g^{-1}\frac{(\Gamma_r - \Gamma_g)f(1-f)}{\Gamma_g f + \Gamma_r(1-f)} + \mu\tau_g^{-1}f \approx -s\tau_g^{-1}f(1-f) + \mu\tau_g^{-1}f \tag{1.40}$$

for $s \ll 1$ in Eq. (1.28) and

$$V[f(\mathbf{x},t)] = \frac{\tau_g^{-1}(\Gamma_r + \Gamma_g)f(1-f)}{\Gamma_g f + \Gamma_r(1-f)} + \mu\tau_g^{-1}f \approx 2\tau_g^{-1}f(1-f) \tag{1.41}$$

for $\mu, s \ll 1$. Equation (1.39) is our desired description! It is an evolution equation for the green cell fraction $f(\mathbf{x},t)$ at points $\mathbf{x}$ and times t along the frontier. If the frontier is inflating, we can use angular coordinates and a time-dependent diffusion coefficient, much like in Eqs. (1.26) and (1.27). However, as described in detail in Chapter 4, we also have to properly generalize the noise $\xi(\mathbf{x},t)$ to an inflating geometry. The reaction term $R[f(\mathbf{x},t)]$ remains the same because it represents the local birth-death dynamics that occur within each test-tube or deme in Fig. 1.3.4. As it stands, though, Eq. (1.39) can be used for any range expansion with a uniform, non-inflating population front. For example $\mathbf{x}$ could represent a point on the surface of a treadmilling tumor, or a point on a planar population front in a reference frame that moves with the front velocity..

What are the properties of Eq. (1.39)? It is a *noisy* reaction-diffusion equation with a non-linear reaction term. Without the noise, i.e. $V = 0$, such equations exhibit travelling wave solutions called Fisher-Kolmogorov-Petrov-Piscounov waves which describe the movement of an interface from a “stable” uniform state into an “unstable” uniform state: $f(\mathbf{x},t) = f^{(s)}$ and $f(\mathbf{x},t) = f^{(u)}$, respectively. Both states satisfy $R[f(\mathbf{x},t) = f^{(s),(u)}] = 0$. The derivative of the reaction term satisfies $R'[f(\mathbf{x}) = f^{(s)}] > 0$ for the stable state and $R'[f(\mathbf{x},t) = f^{(u)}] < 0$ for the unstable one (see Ref. [16] for more details). Under fairly broad conditions, the speed of such a wave is given by [16]

$$v \approx 2\sqrt{-DR'[f(\mathbf{x},t) = f^{(u)}]}. \tag{1.42}$$

In our case, from Eq. (1.40) in the $\mu, s \ll 1$ limit, we find the fixed points $f^{(u)} = 0$ and $f^{(s)} = 1 - \mu/s$ for $s > \mu$. The $f^{(s)}$ fixed point corresponds to a mixed region in which the green cells have enough of a selective advantage s to overcome the irreversible mutation rate μ. (The fixed point $f^{(u)} = 0$ is an all red absorbing state.) The Fisher wave speed in this case is $v \approx 2\sqrt{D(s-\mu)/\tau_g}$. Note that $\mu = s$ is a special! In this case, we only have a single fixed point $f(\mathbf{x},t) = 0$. This threshold is our *mutational meltdown* transition at which the green cells can no longer be sustained in the population. When we include the noise term, however, the properties of these Fisher waves are strongly modified [9, 17]. Indeed, the mutational meltdown no longer occurs at $\mu = s$ and is instead governed by scaling laws of the directed percolation universality class, studied in detail in Chapter 3 (also see the next section).

Let us briefly mention how solutions to the two-dimensional diffusion equation with an absorbing boundary condition come into play here [$d = 2$ case in Eq. (1.9)]. One useful tool in analyzing stochastic equations like Eq. (1.39) is the Îto calculus. This calculus allows us to calculate an evolution equation for an arbitrary functional $\mathcal{G}[f(\mathbf{x},t)]$ of the green cell density field $f(\mathbf{x},t)$. This important formula reads (see, e.g., Ref. [13]):

$$\partial_t \mathcal{G}[f(\mathbf{x},t)] = \mathcal{G}'[f(\mathbf{x},t)]\partial_t f(\mathbf{x},t) + \frac{1}{2}\mathcal{G}''[f(\mathbf{x},t)]V[\mathbf{x},t], \tag{1.43}$$

where the primes indicate (functional) derivatives with respect to $f(\mathbf{x},t)$ and $V[f(\mathbf{x},t)]$ is the noise variance term, as in Eq. (1.39). Equation (1.43) is a kind of *chain rule.* For the usual chain rule, the variance term is absent, as it would be in the Stratanovich interpretation. However, because of our choice of the Îto interpretation, we pick up this extra contribution. We can now average Eq. (1.43) over all noise realizations $\xi(\mathbf{x}, t)$ and use Eq. (1.38) to find an evolution equation for $\langle \mathcal{G}[f(\mathbf{x}, t)] \rangle$.

Important averages of functionals $\langle \mathcal{G}[f(\mathbf{x}, t)] \rangle$ include the various moments and correlation functions of the field, such as $\langle f(\mathbf{x}, t) \rangle$, $\langle f(\mathbf{x}, t) f(\mathbf{x}', t') \rangle$, etc. A particularly useful quantity is the *heterozygosity*. This is the probability $H(\mathbf{x}, t; \mathbf{x}', t')$ that if we take two cells at the points $\mathbf{x}$ and $\mathbf{x}'$ along the population frontier, at possibly two different times t and t', that the cells are different strains (colors). The heterozygosity is related to the two-point, two-time correlation function as follows:

$$H(\mathbf{x}, t; \mathbf{x}', t') = \langle f(\mathbf{x}, t)[1 - f(\mathbf{x}', t')] \rangle + \langle f(\mathbf{x}', t')[1 - f(\mathbf{x}, t)] \rangle. \tag{1.44}$$

Let's find the evolution equation for the heterozygosity for two neutral green and red strains, so that $s = \mu = 0$ in Eq. (1.40) and we take $V[f] \approx 2\tau_g^{-1} f(1 - f)$ in Eq. (1.39). We set $\mathcal{G} = f(\mathbf{x}, t)[1 - f(\mathbf{x}', t')] + f(\mathbf{x}', t')[1 - f(\mathbf{x}, t)]$ in Eq. (1.43) and average both sides. We assume that we seeded a range expansion with a translationally-invariant population frontier, so that $H(\mathbf{x}, t; \mathbf{x}', t') = H(|\mathbf{x} - \mathbf{x}'|, |t - t'|)$. The computation is a bit tedious, but we eventually find that the heterozygosity obeys the simple diffusion equation:

$$\partial_t H(\mathbf{x}, t) = 2D \nabla^2 H(\mathbf{x}, t), \tag{1.45}$$

where D is the same diffusion coefficient as in Eq. (1.39) and we have set $\mathbf{x}' = \mathbf{0}$ and $t' = 0$ without loss of generality. What boundary condition should we impose on $H(\mathbf{x}, t)$? For a population front with a single cell width of actively dividing cells growing outward in a monolayer, we know that if $|\mathbf{x}-\mathbf{x}'| \leq \ell$ for an average separation ℓ between cells ($\ell \sim a$ where a is the cell diameter, because of excluded volume interactions) and $t = t'$, then $\mathbf{x}$ and $\mathbf{x}'$ will represent a single species along the population frontier. This means that

$$H(|\mathbf{x}| \leq \ell, t) = 0. \tag{1.46}$$

This is precisely the absorbing boundary condition we considered previously! (We can clearly impose the condition at $|\mathbf{x}| = \ell$ and maintain a zero concentration for $|\mathbf{x}| \leq \ell$.) We already know the solution to Eq. (1.45) for $d = 1, 2, 3$ from Eq. (1.9).

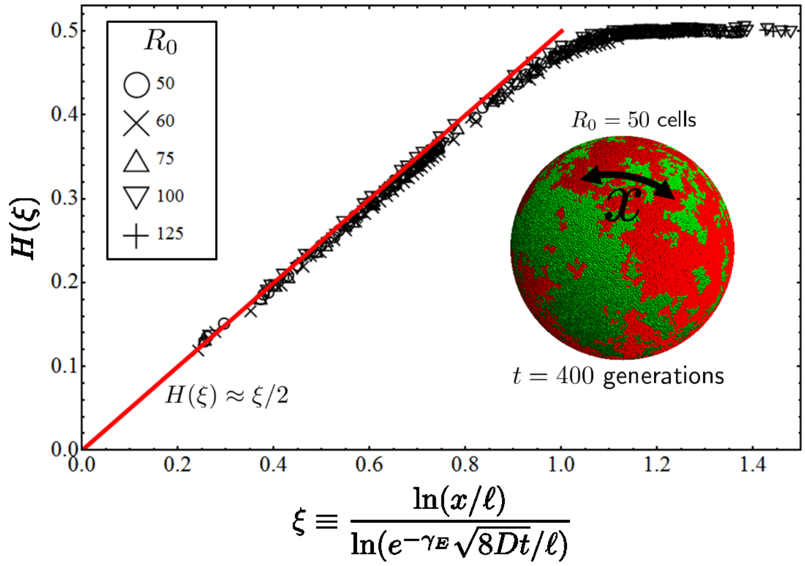


Figure 1.3.5: The heterozygosity for treadmilling expansions on a sphere, collapsed for different $R_0 = 50, 60, 75, 100, 125$ and at different times t. The collapsed data is used to fit the effective diffusion constant $D \approx 0.8$ and cell separation parameter $\ell \approx 0.2$. In the inset we show a snapshot of the unusual genetic domain coarsening, which proceeds logarithmically slowly in time for fixed $\mathbf{x}$. We also show a separation x measured as a geodesic distance along the population frontier.

For an initially well-mixed population of equal portions red and green cells, the initial condition reads $H(\mathbf{x}, t = 0) = 1/2$, since we're equally likely to find a red or green cell at any position along the initial frontier. The evolution of $H(\mathbf{x}, t)$ over time then tells us a lot about the formation of genetic sectors. In particular, we know that the absorbing boundary condition $H(\mathbf{x} = \ell, t) = 0$ will lead to a *depletion* of $H(\mathbf{x}, t)$ for small $\mathbf{x}$, like the solutions in Fig. 1.1.4. We can use Eq. (1.9) to get nice data collapses of the heterozygosity correlation function as determined from our computer simulations. For example, consider the dynamics on a treadmilling, spherical tumor. We can define a separation $x = |\mathbf{x}|$ between two points along the frontier as a geodesic distance, as illustrated in the inset of Fig. 1.3.5. Then, we would have a two-dimensional population front and the heterozygosity $H(\mathbf{x}, t)$ can be written as a function of a single scaling variable. From Eq. (1.9),

$$H(\mathbf{x}, t) = H\left[\xi = \frac{\ln(x/\ell)}{\ln(e^{-\gamma_E}\sqrt{8Dt}/\ell)}\right] \approx \frac{\xi}{2} \quad \text{for} \quad \xi \ll 1, \tag{1.47}$$

where $x \equiv |\mathbf{x}|$ and $\gamma_E \approx 0.577$ is Euler's constant. The data collapse associated with this solution is shown in Fig. 1.3.5. The depletion of $H(\mathbf{x}, t)$ for some $|\mathbf{x}|$ means that for those cell separations $|\mathbf{x}|$, we are more likely to pick two cells inside a single genetic sector. In fact, as shown in Refs. [13, 16], the density $\rho(t)$ of interfaces between red and green domains for three-dimensional expansions with two-dimensional frontiers(without inflation) is given approximately by

$$\rho(t) \approx \ell \left[\frac{\partial H(|\mathbf{x}|, t)}{\partial |\mathbf{x}|}\right]\Bigg|_{|\mathbf{x}|=\ell} \sim \frac{1}{\ln t}. \tag{1.48}$$

Equation (1.48) implies that the genetic domains coarsen *logarithmically slowly* in time. This slow coarsening comes from the special dynamics that occur at the population frontier. The genetic sectors have a wide size distribution and have no sharp boundaries due to a lack of line tension between domains. A typical configuration is shown in the inset of Fig. 1.3.5.

1.4 Nonequilibrium Phase Transitions and Range Expansions with Mutualism

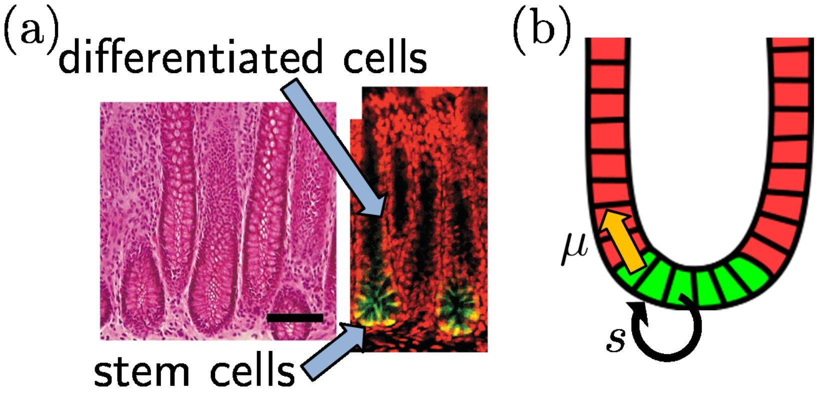


Figure 1.4.1: (a) Two micrographs of colonic crypts. The left micrograph, taken from Ref. [18], shows a 50 μm scalebar. The crypts are the elongated, cylindrical invaginations open at the top and closed at the bottom. In the right micrograph (Ref. [19]), the stem cells which replenish the other, differentiated cells in the crypt are shown in the lighter green color. The differentiated cells are in red, running along the walls of the crypt. (b) A schematic of the colonic crypt. Green stem cells grow at a more rapid rate, with some characteristic selection advantage s over the differentiated red cells which grow more slowly. The stem cells irreversibly convert to red cells at some rate μ.

The irreversible mutation dynamics from a fit strain to a less fit one have a natural, important biological implication in stem cell dynamics. For example, in colonic crypts, stem cells grow at a rapid rate and replenish the rest of the crypt. Some micrographs of these crypts are shown in Fig. 1.4.1. They have a cylindrical shape and a distinct spatial structure. Hence, our spatial analysis with $\Theta = 0$ and periodic boundary conditions around the cylindrical axis might be relevant for this system, as already noticed in previous studies (Refs. [20, 21]). These previous studies have focussed on the neutral competition dynamics occurring between stem cell clones. However, selection and mutation might also play an important role. For example, as the stem cells divide, they can also irreversibly “mutate” into terminally differentiated cells at some rate μ. Hence, they must divide fast enough (i.e., have a large enough selective advantage s) to overcome this differentiation rate, or they will die out. The mutation-selection balance that keeps the stem cell population constant is shown schematically in Fig. 1.4.1(b). In Chapter 3, we model range expansions with these dynamics and study the mutation-selection balance.

Interactions between variants are of course not limited to selective advantages and mutations. More complicated, interesting interactions can occur, for example, if the various cell types secrete a poison or nutrient which is either harmful or beneficial to another cell type. These interactions can be genetically engineered in the laboratory. For example, two yeast strains have been designed (see Refs. [22, 23]) to cross-feed each other as they expand across an agar plate. Each strain requires an amino acid that is over-produced by the other strain. The mutualistic interaction is suppressed when all necessary amino acids are already present in the agar. However, if the amino acids required by each strain are absent (or present in reduced concentrations), an obligate mutualistic interaction which prevents the strains from demixing is observed (see Fig. 1.4.2).

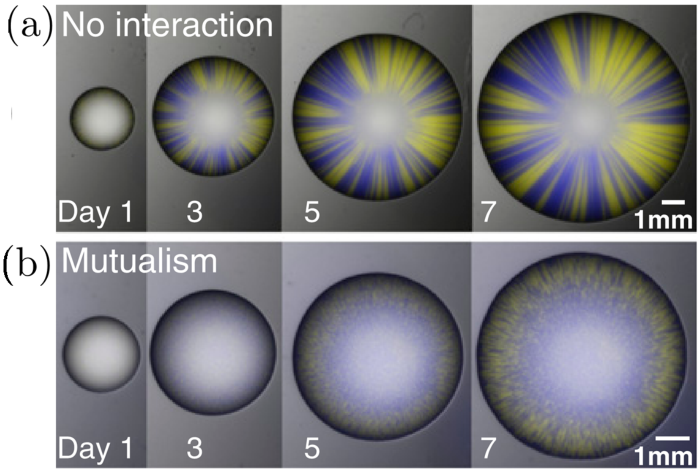


Figure 1.4.2: Figure adapted from Ref. [22]. (a) A yeast range expansion on a Petri dish of two strains, blue and yellow. Each one secretes an extra amount of an essential amino acid (tryptophan or leucine), while being unable to produce another kind of amino acid. The Petri dish in this case contains all amino acids, however, so that the two strains do not interact. (b) The Petri dish in this expansion lacks the two essential amino acids that each strain requires. Hence, the strains have to cross-feed each other in order to grow. This obligate mutualism prevents the demixing of the strains into distinct genetic sectors.

Both the mutational meltdown dynamics that we will study in Chapter 3 and the onset of mutualism discussed in Chapter 5 involve non-equilibrium phase transitions with absorbing states. In these transitions, there is typically a fluctuating, active phase which can transition to an inactive, absorbing state from which the system can no longer escape. It is clear that in the mutational meltdown dynamics for stem cells described above, the absorbing state is just a population of terminally differentiated cells. The active state is any region with stem cells. In the mutualistic dynamics, we have two absorbing states corresponding to populations of just one of the two species. The active phase is then a mixture of the two, with active cross-feeding between neighbors. We show in Chapters 3 and 5 that these dynamics are both describable by a reaction-diffusion equation of the general form of Eq. (1.39) (see also Ref. [24]). The only difference is in the form of the reaction and noise terms $R[f(\mathbf{x},t)]$ and $V[f(\mathbf{x},t)]$, respectively. In a well-mixed population or in a mean-field approximation, the phases can be found by studying uniform states or fixed points $f(\mathbf{x}, t) = f^*$ which satisfy $R[f(\mathbf{x}, t) = f^*] = 0$ in Eq. (1.39), as discussed in the previous section. Since we have no spatial dependence $\mathbf{x}$ in a well-mixed population, we can approximate the dynamics of $f(\mathbf{x}, t) = f(t)$ with just

$$\frac{df}{dt} = -R[f(t)], \tag{1.49}$$

where we have also ignored the noise term for simplicity. We already saw that for the mutational meltdown, there is a phase transition when $\mu = s$. A similar condition can be derived for an appropriate model of mutalistic dynamics for the onset of mutualism.

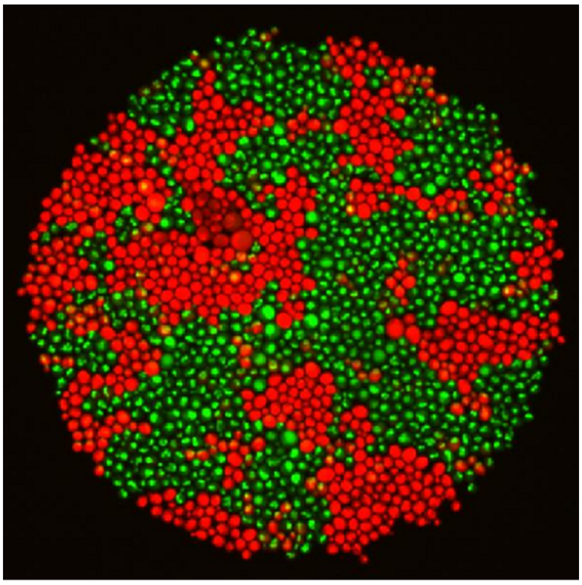


Figure 1.4.3: A small monolayer colony of two specially engineered yeast strains grown in a microfluidic device, falsely colored red and green using the two fluorescent markers exhibited by the strains (image courtesy of M. Wahl). The green cell daughters irreversibly mutate to the red cells with some tunable probability μ each time they divide. In the presence of cycloheximide, the green cells enjoy a concentration-dependent selective advantage s over the red cells. In this image, no cycloheximide is present.

When we include the noise term proportional to $\xi(\mathbf{x}, t)$ and the additional spatial dynamics in the diffusion term in Eq. (1.39) for a range expansion, the phase transition is typically strongly modified, especially for lower-dimensional population fronts (i.e., for $d = 1, 2$). These modifications are studied in detail in Chapters 3 and 5. However, to motivate that discussion, let us briefly mention a striking experimental realization of the irreversible mutations in the yeast *Saccharomyces cerevisiae.* Mary Wahl and Andrew Murray at Harvard University have developed a strain of fit "green" yeast (expressing a particular fluorescent protein) whose daughters can irreversibly "mutate" to a less-fit "red" strain (expressing another fluorescent protein) with some rate μ each time the yeast divides. The mutation is induced by a particular recombinase that splices out a region of the "green" yeast DNA. This region codes for a cycloheximide resistant ribosome, which provides the "green" yeast with a selective advantage s over the strain with the region spliced out (and only non-resistant ribosomes) when both are grown in the presence of cycloheximide (a poison that inhibits yeast growth). The selective advantage can be tuned by varying the concentration of cycloheximide in the nutrient medium. The mutation rate can be tuned experimentally as well. These dynamics mimic the stem-cell termination process in colonic crypts illustrated in Fig. 1.4.1.

The "green" strain expresses a particular fluorescent protein in the special region of DNA that gets spliced out in the red strain. This allows for easy visualization of the two strains. A monolayer of these strains, artificially colored green and red, is shown in Fig. 1.4.3. In Fig. 1.4.4 Wahl et al. (unpublished) compare the fraction of the green strain in both well-mixed populations and two-dimensional range expansions on agar plates. The spatial dynamics in the range expansion dramatically change the phase diagram. Instead of occurring at $\mu = s$, the mutational meltdown now occurs at $\mu \approx ks^2$, where $k \sim 1/D$ is a coefficient that depends on the diffusion coefficient D of the genetic sectors. This change in the phase transition line and other important consequences of spatial dynamics and noise are the subject of Chapters 3, 4, and 5.

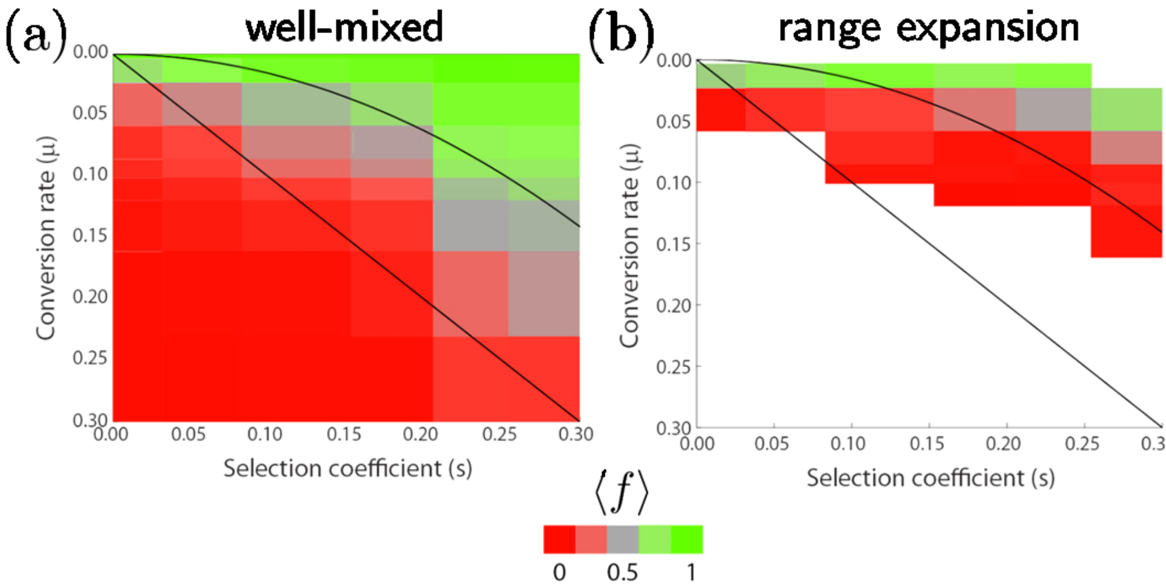


Figure 1.4.4: A plot of the fraction of a fit yeast cell strain at long times in (a) well-mixed populations and (b) in range expansions on an agar plate (Figure and data courtesy of M. Wahl). The fit yeast strain has an irreversible mutation rate μ to a less fit strain over which it has some selective advantage s. Both of the parameters are varied in the experiment and there is mutational meltdown in the transition between the green and red regions. The straight line is the expected well-mixed result $\mu = s$ for the transition line. The curved line is a fit to $\mu = ks^2$, which is the expected transition line for a two-dimensional range expansion, derived from directed percolation phase transition theory (see Chapter 3). In the experiments, $k \approx 1$. We expect k to increase with decreasing genetic drift.

We have now given a brief overview of all topics covered in this thesis. We have also shown that the subsequent chapters of this thesis can be tied together by the repeated application of the mathematical description of a random walker in the presence of an absorbing site. In addition, we have given two important biological motivations for studying non-equiliubrium phase transitions in noisy reaction-diffusion equations like Eq. (1.39). We start, however, with a detailed look at nutrient diffusion and absorption in clusters of cells in Chapter 2.

Chapter 2

Nutrient Shielding in Clusters of Cells

This chapter appears in its entirety in M.O. Lavrentovich, J. H. Koschwanez, and D.R. Nelson, *Phys. Rev. E* **87**, 062703 (2013)

2.1 Introduction

Nutrient uptake is essential for all life and has been studied in a variety of model organisms; however, the physical mechanisms involved in the uptake are not yet well understood. A cell commonly takes up nutrients from its surrounding medium via facilitated diffusion or active transport: Specialized transporters on the cell surface move the nutrient down (facilitated diffusion) or up (active transport) a concentration gradient from the ambient medium into the cell. The cell's absorption of nutrients in its immediate vicinity sets up and maintains a concentration gradient outside the cell that gradually depletes the nutrient from the surrounding medium. This is the method used by budding yeast cells for acquiring glucose, for example [25, 26, 27].

Nutrient consumption via facilitated diffusion is ubiquitous in nature and occurs in a variety of contexts, including oxygen consumption by human tissue cells, calcium uptake by intestinal cells, bacterial absorption of various sugars, etc [28, 29, 30]. In many of these examples, cells grow in clusters and shield each other from the available nutrients. In Fig. 2.1.1, we see confocal microscope images of yeast colony cross sections in which yeast cells marked green are growing due to an abundance of glucose. Nutrient (glucose) shielding in larger colonies (Colonies 2 and 3) prevents cell growth in the colony interiors (red regions in Fig. 2.1.1) and only an outer shell of thickness ℓ is able to grow (illustrated for Colony 3 in Fig. 2.1.1).

Typical bacterial and yeast cell colonies (e.g., those in Fig. 2.1.1) are dense cell clusters with large cell packing fractions $\phi \gtrsim 0.5$ [31, 32]. In this chapter, we characterize both densely and sparsely packed cell clusters. Dilute clusters with $\phi \lesssim 0.1$ can be constructed artificially, by suspending microbial cells in a gelatinous matrix. Immobilized microbial cells have a very wide range of industrial and environmental applications (see Ref. [33] for a review). For example, gelatinous beads are seeded with yeast cells and used in reactors to produce ethanol [34]. To better understand the growth dynamics and physical properties of these systems, it is important to characterize the nutrient transport properties of cell clusters as a function of both single cell nutrient uptake kinetics and the geometry of specific cell packings.

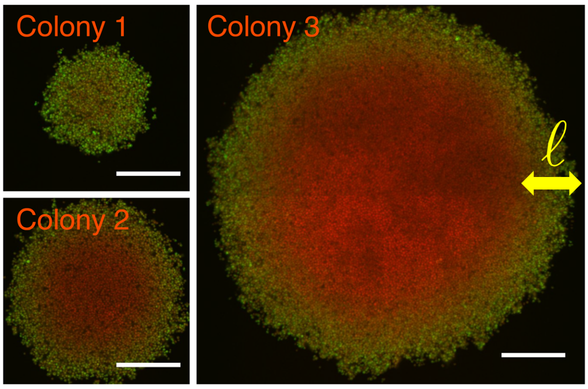


Figure 2.1.1: Confocal microscope images of cross-sections through the bottom of three budding yeast colonies (scale bars represent 100 μm). The red (darker) color is the constitutive expression of a protein, whose level is largely independent of growth rate, in all cells and the green (lighter) color is ribosomal protein expression, indicating growth. Colony 1: 0.5 mM glucose, 43 h after inoculation; Colony 2: 1.5 mM glucose, 47 h after inoculation; Colony 3: 4.5 mM glucose, 56 h after inoculation. Small colonies such as Colony 1 receive enough nutrients for all cells to grow. Colonies 2 and 3 are larger and growth occurs only in an outer shell of thickness ℓ (illustrated for Colony 3). The red cells in the interior are shielded from the nutrients. We assume the colonies are roughly spherical caps (see Sec. 2.4 for a discussion of the shape and experimental details).

A nutrient concentration ψ in some medium, such as water or gel, with a constant diffusion coefficient D_0 obeys the diffusion equation

$$\partial_t \psi = D_0 \nabla^2 \psi. \tag{2.1}$$

In the steady state, the left hand side of Eq. (2.1) vanishes, and the equation reduces to Laplace's equation. This is an important and well-studied equation in electrostatics, as it is the equation for the electrostatic potential $\Phi(\mathbf{r}) \equiv \psi(\mathbf{r})$ in regions of space without any charges.

In the case of nutrient diffusion, Eq. (2.1) must be satisfied everywhere in the medium outside of the cell. However, to completely solve Eq. (2.1) in the steady state, we must specify boundary conditions. One natural boundary condition is to set the concentration field at infinity to some constant value, i.e., $\psi(|\mathbf{r}| \to \infty) \to \psi_\infty$, corresponding to a large nutrient bath with a uniform concentration ψ_∞. We also need boundary conditions on each cell surface. For example, if the cell is a perfect nutrient absorber, then for all points $\mathbf{r}$ on the cell surface S, the nutrient concentration vanishes, i.e. $\psi(\mathbf{r}) = 0$. In the electrostatic analogy, this condition means zero electrostatic potential on every cell surface, i.e., each cell is a perfect grounded conductor. Conversely, if the cell does not absorb any nutrient, (i.e. it is a perfect reflector), then Fick's first law of diffusion tells us that the derivative of the concentration along a direction $\hat{\mathbf{n}}$ perpendicular to the cell surface must vanish. More precisely, the local nutrient flux density $J(\mathbf{r})$ into the cell at some point $\mathbf{r} \in S$ on the surface satisfies

$$J(\mathbf{r}) = D_0 \hat{\mathbf{n}} \cdot \nabla \psi(\mathbf{r}) \tag{2.2}$$

so that $J(\mathbf{r}) = 0$ for all $\mathbf{r} \in S$ implies $\hat{\mathbf{n}} \cdot \nabla \psi(\mathbf{r})|_{\mathbf{r} \in S} = 0$. In the electrostatic analogy, this would correspond to a perfect insulator with no surface charge, with a vanishing normal electric field. Of course, living cells are neither perfect absorbers nor perfect reflectors. A more realistic boundary condition interpolates between these two ideal cases.

A boundary condition on the cell can be derived from a more microscopic model of the nutrient transporters. For example, Berg and Purcell modeled transporters as small perfectly absorbing disks on the surface of an otherwise reflecting cell [35, 36]. They showed that the cell requires very few transporters to act as an effectively perfect absorber: A cell with as little as a 10^{-4} fraction of its surface covered by transporters takes in half the nutrient flux of a perfect absorber! Zwanzig and Szabo later extended this result to include the effects of transporter interactions and partially absorbing transporters [37, 38]. They showed that a homogeneous and partially absorbing cell surface model captures the average effect of all the transporters. As discussed below, in many cases of biological interest, the cell cannot be treated as a perfect absorber. The same partially absorbing boundary condition used by Zwanzig and Szabo will be derived in a different way in the next section.

Although Eq. (2.1) is easily solved in the steady state for a single, spherical cell with the appropriate boundary conditions [35, 36], the complicated arrangement of cells in a typical multi-cellular system, such as a yeast cell colony, implies a complex boundary condition that makes an exact solution intractable – one would have to constrain $\psi(\mathbf{r})$ and its normal derivative on a highly irregular object, like the surface of a cluster of grapes. In this chapter, we explore an "effective medium" approximation to the exact solution of this problem.

Effective medium theory treats a cluster of cells, or nutrient sinks, as a region with uniform effective nutrient transport properties (such as an effective diffusivity and nutrient absorption constant) that depend on the arrangement of cells in the cluster and the individual cell nutrient absorption properties. A key feature of the effective medium theory is that these effective transport properties are derived in a self-consistent way. These theories have been used to calculate many effective properties in heterogeneous systems such as conductivity, elasticity, and reaction rates (see [39, 40] and Chapter 18 in [41] for reviews).

The chapter is organized as follows: We develop our theoretical model for nutrient uptake in single cells, dilute cell clusters, and dense clusters in Sec. 2.2. In Sec. 2.3 we compare our analytic results with numerical solutions of Eq. (2.1) in the steady-state for clusters with hundreds of partially absorbing cells. We discuss experimental tests of our model in Sec. 2.4 and provide concluding remarks in Sec. 2.5.

2.2 Theoretical Model

We now discuss how to couple single cell nutrient uptake kinetics to the nutrient uptake behavior of an entire cluster or colony of cells via effective medium theory. In what follows we assume the cells are all identical and spherical. Although our experimental model is the budding yeast cell, the theoretical treatment is quite general and can be adapted to any cell cluster that passively absorbs nutrients that reach it via diffusion.

2.2.1 Single Cell Nutrient Uptake

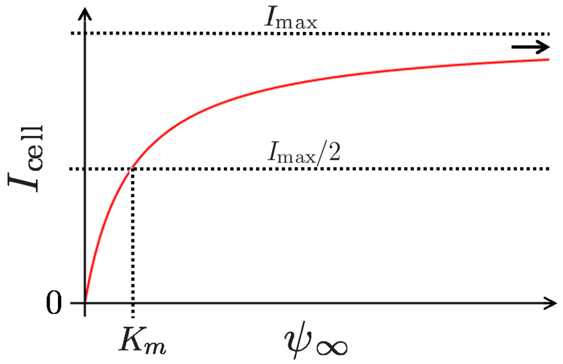


Figure 2.2.1: A plot of the nutrient current I_{cell} (solid red line) into a cell as a function of the ambient nutrient concentration ψ_∞, which follows Michaelis-Menten kinetics. The maximum current $I_{\max}$ and the kinetics parameter K_m are also shown.

We first review nutrient uptake by a single cell, with nutrient transporters operating via facilitated diffusion and following Michaelis-Menten kinetics. We smooth out the effect of a discrete set of transporters within the cell wall and study a radially symmetric model of nutrient uptake. A Michaelis-Menten model then means that the total nutrient current I_{cell} into a cell is related to the ambient nutrient concentration ψ_∞ via

$$I_{\text{cell}} = \frac{I_{\max}\psi_\infty}{K_m + \psi_\infty}, \tag{2.3}$$

where $I_{\max}$ (sometimes called $V_{\max}$ in the literature) is the saturating nutrient current into the cell as $\psi_\infty \to \infty$ and K_m is the concentration at which $I_{\text{cell}} = I_{\max}/2$, as shown in Fig. 2.2.1. Even if a cell does not obey these kinetics for all ψ_∞, one can often define a range of concentrations ψ_∞ characterized by effective kinetic parameters $I_{\max}$ and K_m, a characterization often used in experimental studies of nutrient uptake. Many studies infer an effective $I_{\max}$ and K_m from the measured amount of nutrient consumed by a dilute suspension of cells [27, 42, 43].

The parameters $I_{\max}$ and K_m are determined by the microscopic kinetics of each individual cell transporter and the density of these transporters on the cell surface. A third contribution arises from the structure of the cell wall, which influences the rate of transport of the nutrient into the cell. We can find $I_{\max}$ and K_m by appealing to a microscopic model of the transporters. Berg and Purcell, for example, modeled each transporter as a small, perfectly absorbing disk on the surface of a perfectly reflecting cell [35, 36]. Zwanzig showed that the large scale physics of this model are approximated by a partially absorbing boundary condition on the surface of the cell [37]. We will now derive this boundary condition in another way and connect $I_{\max}$ and K_m to our model.

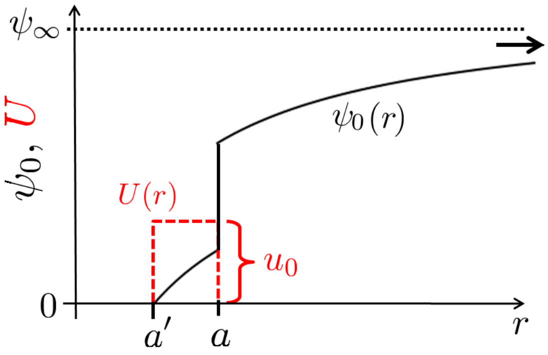


Figure 2.2.2: A radially symmetric model of nutrient uptake in a single cell. We specify a steady concentration ψ_∞ far away from the cell. The cell is centered at the origin and has radius a. The steady state concentration profile $\psi_0(r)$ (black solid line) is calculated for a rectangular potential barrier (red dashed line) of height u_0 (designed to model the complex uptake dynamics of the cell wall) and a perfectly absorbing nutrient sink at $r = a'$. The concentration profile approaches ψ_∞ at large r and exhibits a jump discontinuity at $r = a$. The nutrient currents are constrained to be continuous.

To model the cumulative effect of many microscopic details, such as transporter kinetics, the cell wall, etc., on the nutrient uptake we introduce a potential energy barrier $U(\mathbf{r})$. Once a nutrient molecule surmounts this barrier, it gets absorbed by the cell. The nutrient concentration $\psi(\mathbf{r})$ obeys the Fokker-Planck equation

$$\partial_t \psi = D_0 \nabla \cdot \left[\nabla \psi + \frac{1}{k_B T} (\nabla U) \psi \right], \tag{2.4}$$

where k_B is the Boltzmann constant and T is the temperature of the nutrient solution [14]. Simple diffusion is recovered when the potential is constant. For simplicity, let's suppose that the nutrient must overcome a radially symmetric potential barrier $U(r)$ that has a rectangular "lip" of height u_0 at $r = a$ and with width $w \equiv a - a'$ (see Fig. 2.2.2). Thus,

$$U(r) = \begin{cases} 0 & r < a' \\ u_0 & a' \le r \le a \\ 0 & r > a, \end{cases} \tag{2.5}$$

where u_0 is the height of the barrier. We assume perfect absorption at a' and a constant nutrient concentration infinitely far away,

$$\begin{cases} \psi(r = a', t) = 0 \\ \psi(r \to \infty, t) = \psi_\infty . \end{cases} \tag{2.6}$$

To determine the nutrient flux into the cell, we solve Eq. (2.4) for the steady state profile $\psi_0(r)$. In addition to the boundary conditions [Eq. (2.6)], we ensure the continuity of the nutrient flux at $r = a$ via the "jump conditions" at $r = a$, as discussed in Ref. [44]. The resulting concentration profile reads

$$\psi_0(r) = \begin{cases} \psi_\infty - \dfrac{aa'\psi_\infty}{a' + (a - a')e^{u_0/k_B T}} \dfrac{1}{r} & r \geq a \\ \dfrac{a(r - a')\psi_\infty}{(a' + (a - a')e^{u_0/k_B T})} \dfrac{1}{r} & a' < r < a, \end{cases} \tag{2.7}$$

with $\psi_0(r) = 0$ for $r < a'$. The shape of the solution $\psi_0(r)$ is shown in Fig. 2.2.2.

Let's now consider narrow potential barriers relative to the cell radius ($w = |a - a'| \ll a$). Then, from Eq. (2.7), we find the concentration gradient just outside the cell surface:

$$\partial_r \psi_0 |_{r \to a^+} \approx \frac{1}{w} \exp\left[-\frac{u_0}{k_B T} \right] \psi_0(a^+) \equiv \kappa \psi_0(a^+), \tag{2.8}$$

where the + superscript indicates that we take the limit $r \to a$ from outside the cell. Eq. (2.8) reveals that the gradient of ψ_0 normal to the cell surface is proportional to $\psi_0(r)$ just outside. Notice that $\kappa \to \infty$ when $w \to 0$ (we also let $u_0 \to 0$), so that $\psi_0(a^+) \to 0$ at the cell surface to keep the flux finite. Thus, the cell is perfectly absorbing within our model if there is no potential barrier. Similarly, for a very large barrier ($u_0 \to \infty, w$ finite), we have $\kappa \to 0$ so that there is no flux of nutrient into the cell and $\partial_r \psi_0(a^+) \to 0$, signifying a perfect reflector.

The proportionality between a field and its gradient at a boundary is called a *radiation* boundary condition in the physics literature and can be derived quite generally [45]. This boundary condition is a natural coarse-grained description of the Berg and Purcell model of transporters as absorbing disks. Zwanzig and Szabo [37, 38] have used the radiation boundary condition to successfully model the physics of both perfectly and partially absorbing disks on scales larger than the disk spacing, thus confirming our expectation that the coarse-grained nutrient uptake can be modeled by the ubiquitous radiation boundary condition with an appropriate choice of κ.

The absorptive strength of the cell can be parameterized by the dimensionless number $\nu \equiv \kappa a$, where a is the cell radius. In chemical engineering, ν is sometimes referred to as a Sherwood number [46]. If P_{abs} is the probability that a nutrient particle at the cell surface will be absorbed by the cell (instead of escaping to infinity), then first-passage techniques from probability theory [3] lead to

$$P_{\text{abs}} = \frac{\nu}{1+\nu}. \tag{2.9}$$

Thus, $\nu = \kappa a \ll 1$ indicates poor nutrient absorption while $\nu \gg 1$ indicates a good absorber. Note that at $\nu = 1$, the nutrient has equal probability of being absorbed at the cell surface or escaping to infinity.

We now connect ν with the measurable biological parameters $I_{\max}$ and K_m. Recall that the nutrient flux into a single cell is related to the ambient nutrient concentration

ψ_∞ via the Michaelis-Menten relation Eq. (2.3). Suppose for now that the cells are well-separated, so the nutrient uptake I_{cell} of any given cell is independent of the others (i.e., there is no nutrient shielding), and that the ambient nutrient concentration ψ_∞ is held constant, so that the parameters $I_{\max}$ and K_m at each cell can assume their steady-state values. We assume the nutrient solution experiences no macroscopic flows, such as convection currents, that would bias the isotropic absorption kinetics of the cell. The nutrient concentration $\psi(r \equiv |\mathbf{r}|)$ then satisfies $\nabla^2\psi(\mathbf{r}) = 0$ in the steady state, with the boundary conditions $\psi(r \to \infty) = \psi_\infty$ and $\hat{\mathbf{n}} \cdot \nabla\psi|_S = \kappa\psi|_S$ at each cell surface, as discussed above.

Upon inserting Eqs. (2.7) and (2.8) into Fick's first law [3] [see also Eq. (2.2)], we find the steady-state nutrient current into an individual cell,

$$I_{\text{cell}} = \int_S D_0 \hat{\mathbf{n}} \cdot \nabla\psi(r) a^2 \, \mathrm{d}\Omega \approx \frac{4\pi D_0 \psi_\infty \nu a}{1+\nu}, \tag{2.10}$$

where we integrate over the surface of the cell S (so that $\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi$ in spherical coordinates) and again assume $w = a - a' \ll a$. Comparison of Eq. (2.10) with Eq. (2.3) leads to ν as a function of biological parameters. In the limit of low ambient nutrient concentration ($\psi_\infty \ll K_m$), we have

$$\nu = \frac{I_{\max}}{4\pi a D_0 K_m - I_{\max}}. \tag{2.11}$$

It is also possible to define an effective ν for a reflecting spherical cell uniformly covered by identical, partially absorbing disks with radius a_{disk} and absorption parameter κ_{disk}. In this case, using the boundary condition Eq. (2.8) on each disk surface, Zwanzig and Szabo find that the effective parameter ν for the entire cell (for $a_{\text{disk}}\kappa_{\text{disk}} \ll 4/\pi$)

is $\nu = N_{\text{disk}} a_{\text{disk}}^2 \kappa_{\text{disk}}/4a$, where N_{disk} is the number of disks on the cell surface (see Ref. [38]).

A rough estimate of ν for glucose uptake by a *S. cerevisiae* cell follows from values for $I_{\max}$ (4.2×10^7 molecules/sec), K_m (7.4 mM), a (2 μm), and D_0 (670 μm^2/sec) found in the literature [27, 47, 48]. The $I_{\max}$ is particulary difficult to estimate as the nutrient uptake rate in experiment is calculated per gram of dried yeast taken out of a liquid culture. To get an uptake rate per cell, we estimate that a yeast cell has a 2 pg dry weight [49]. We find that these yeast cells are in fact very poor absorbers with $P_{\text{abs, yeast}} \approx \nu_{\text{yeast}} \approx 6 \times 10^{-4} \sim 10^{-3}$ within an order of magnitude. Gram-negative bacterial cells are quite different! Again using literature values for $I_{\max}$ (2×10^7 molecules/sec), K_m (1 μM), and a (0.5 μm) [42, 50] for a single *Escherichia coli* cell, we find that $P_{\text{abs, gram-n.}} \approx \nu_{\text{gram-n.}} \approx 0.09 \sim 10^{-1}$. Thus, this bacterium is $\sim 10^2$ times more absorbent than the yeast cell: This striking difference has profound biological implications since, as we will show in the next sections, the parameter ν greatly influences the growth of a cell colony.

The large disparity in ν values may be due to the thicker cell walls of *S. cerevisiae*, compared to gram-negative bacteria like *E. coli.* The presence of a cell wall can have two effects. First, the diffusion coefficient of the transported nutrients may be lower in the cell wall medium than it is in the bulk solution. Second, the absorbing surface lies at the plasma membrane, not the surface of the wall. The second effect implies that the absorbing surface is at a distance $a - w$ from the cell center (rather than a), where a is the cell radius and w the thickness of the wall. So, even if the diffusion coefficient inside the wall is the same as in the bulk and we have a perfectly absorbing plasma membrane, the effective value of ν is $\nu = a/w - 1$. This argument is consistent with the measured glucose uptake kinetics for the gram-positive bacterium *Luconostoc mesenteroides*, which has a cell wall thicker than *E. coli* and thinner than *S. cerevisiae*. We find $P_{\text{abs, gram p.}} \approx \nu_{\text{gram p.}} \approx 5 \times 10^{-2}$ from the literature values of the parameters [43, 51]. Of course, other factors apart from the cell wall thickness could be relevant.

The experimental results discussed here for single cells tell us that it is important to consider a large range of the parameter $\nu = \kappa a$: ν can range over at least two orders of magnitude ($10^{-3} < \nu < 10^{-1}$) for yeast, gram-positive, and gram-negative bacteria. Thus, neither perfectly reflecting nor perfectly absorbing boundary conditions are relevant for nutrient uptake in many cell populations. Instead, we develop a theory for the nutrient absorption by a cluster of cells with arbitrary ν. This is the subject of the next section.

2.2.2 Nutrient Uptake in Cell Clusters

We now consider nutrient absorption by a cluster or colony of cells. In general, this is a very complicated problem involving solving the diffusion equation in the interstitial area of the cluster while making sure the boundary condition [Eq. (2.8)] is satisfied at each cell surface. Although analytical results are possible for a single cell, we must resort to numerical solutions and approximations when dealing with a cluster. A typical 1 mm diameter yeast cell colony contains over 10^6 cells, making the exact solution for nutrient uptake by such a colony intractable even numerically.

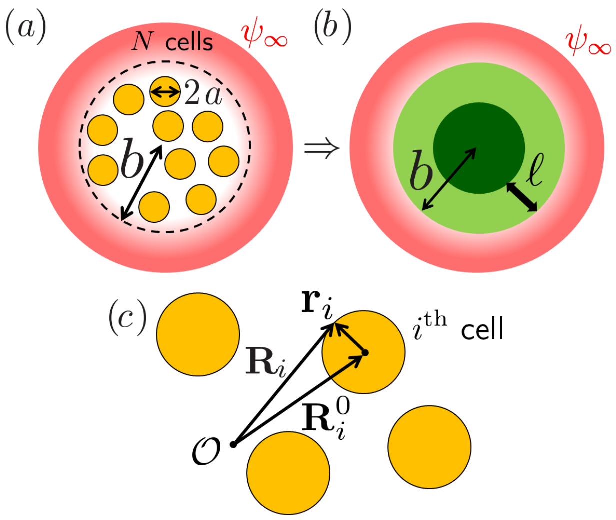


Figure 2.2.3: Schematic of the effective medium approximation, which replaces N orange cells in a spherical cluster with radius b (enclosed by the dashed lines) in (a) by a homogeneous, attenuating medium in (b) shown in dark and light green. The light green rim of width ℓ illustrates the section of the cluster receiving enough nutrients to grow. ψ_∞ is the limiting nutrient concentration in the variably shaded red region, outside the cluster. To analyze this problem, we use the coordinate system in (c), with an origin at $\mathcal{O}$, cell centers located at $\{\mathbf{R}_i^0\}_{i=1}^N$, and vectors $\mathbf{r}_i$ pointing from the cell center to the surface of the cell.

To model nutrient uptake (and nutrient shielding) in a group of cells, we again consider the steady-state diffusion equation with boundary conditions provided by a disorderly cluster of N identical spherical cells all with radius a and no overlaps (see Fig. 2.2.3). Let the cells be located inside a spherical region of radius $b \gg a$ with centers at positions $\mathbf{R}_i^0$, where $i = 1, \ldots, N$. We also employ local spherical coordinates at each cell to write vectors $\mathbf{r}_i = (a, \Omega_i)$ pointing at positions on the cell surface, where $\Omega_i \equiv (\theta_i, \phi_i)$ is a pair of polar angles that specifies the direction of $\mathbf{r}_i$ relative to the center of the ith cell (see Fig. 2.2.3 (c)). Consider a particular configuration of these N cells and denote by $\Psi(\mathbf{r})$ the *exact* solution to the steady-state concentration field in the interstitial region for this particular configuration.

Upon assuming a time-independent steady-state, we modify Eq. (2.1) to include a set of Lagrange multiplier functions $\sigma_i(\mathbf{r})$ defined on the cell walls,

$$D_0 \nabla^2 \Psi(\mathbf{r}) = \sum_{i=1}^{N} \int \sigma_i(\mathbf{r}_i)\, \delta(\mathbf{r} - \mathbf{R}_i)\, \mathrm{d}\Omega_i + s(\mathbf{r}). \tag{2.12}$$

We also set $\mathbf{R}_i \equiv \mathbf{R}_i^0 + \mathbf{r}_i$ (see Fig. 2.2.3(c)) and choose $\mathbf{r}$ to be inside the cell cluster. The Lagrange multiplier functions $\{\sigma_i(\mathbf{r}_i)\}$ will be chosen to satisfy the radiation boundary condition at each cell surface (the functions $\{\sigma_i(\mathbf{r}_i)\}$ would be charge densities in electrostatics). These boundary conditions [given Eq. (2.8) with $\nu = \kappa a$ for a single cell] read, for all i and $\mathbf{R}_i$,

$$\Psi(\mathbf{R}_i) = \nu^{-1}\, \mathbf{r}_i \cdot \nabla \Psi(\mathbf{R}_i) \equiv \mathcal{Q}_i \Psi(\mathbf{R}_i), \tag{2.13}$$

where $\mathcal{Q}_i \equiv \nu^{-1}\, \mathbf{r}_i \cdot \nabla$ is a convenient gradient operator used in the detailed effective medium calculation in Appendix A.1. The source function $s(\mathbf{r})$ allows us to incorporate additional boundary conditions on the concentration field.

We now average over all possible cell configurations (consistent with excluded volume interactions between cells) to obtain the average transport properties of the nutrient in a cell cluster. Upon averaging both sides of Eq. (2.12), we find

$$D_0 \nabla^2 \psi(\mathbf{r}) = \left\langle \sum_{i=1}^{N} \int_{S_i} \sigma_i(\mathbf{r}_i)\, \delta(\mathbf{r} - \mathbf{R}_i)\, \mathrm{d}\Omega_i \right\rangle + s(\mathbf{r}) \tag{2.14}$$

$$\approx \int \Sigma(\mathbf{r}' - \mathbf{r})\, \psi(\mathbf{r}')\, \mathrm{d}\mathbf{r}' + s(\mathbf{r}), \tag{2.15}$$

where $\psi(\mathbf{r}) \equiv \langle \Psi(\mathbf{r}) \rangle$. The bracket average is an ensemble average over cell configurations and $\Sigma(\mathbf{r})$ is a linear response function describes how the cells deform the concentration field.

The linear response approximation, justified here by comparisons with simulations, is only valid away from the cluster edges and for sufficiently small concentration field deformations. In addition, this approximation often breaks down for time-dependent diffusion [52] because the transient diffusive dynamics are dominated by slowly decaying modes due to large voids inside the cell cluster [53]. Ref. [54] uses a more microscopic description of nutrient diffusion to examine the validity of the linear response approximation in more detail. We will forgo these complications here and exploit the linear approximation above, checking our assumptions using experiments and simulations, as had been done for the physics of fluorescence quenching [55].

An exact evaluation of $\Sigma(\mathbf{r})$ involves an ensemble average denoted by brackets in Eq. (2.14). This average requires the full probability distribution $P(\{\mathbf{R}_i^0\}_{i=1}^N)$ of observing N cells with centers $\{\mathbf{R}_i^0\}_{i=1}^N$. Unfortunately, an exact solution obtained in this way would require knowledge of all of the correlations between the cell positions, which may not be experimentally accessible. We will assume for now that this distribution is known. Later, the effective medium theory developed in Sec. 2.2.5 will approximate $\Sigma(\mathbf{r})$ in a self-consistent way using just the one and two cell center distributions.

What happens to the configurationally averaged solution $\psi(\mathbf{r})$ of Eq. (2.15) over distances large compared to the size of a single cell? Specifically, how does the cell colony absorb nutrients *on average*, as if it were the homogeneous medium illustrated in Fig. 2.2.3(b)? To answer this question, we perform a gradient expansion of Eq. (2.15). As discussed in Ref. [54], such a gradient expansion neglects intrinsically non-local contributions to $\Sigma(\mathbf{r})$ due to the fluctuations in the concentration field that cannot be averaged over large distances. The mean field approximation used here then uses the resulting transport coefficients to describe the absorptive properties of the cluster. Simulations have shown that this approach correctly models the physics in related systems [56].

Equation (2.15) in Fourier space reads

$$-D_0 q^2 \psi(\mathbf{q}) = \Sigma(\mathbf{q})\psi(\mathbf{q}) + s(\mathbf{q}), \tag{2.16}$$

where we have applied the convolution theorem and all the functions are now their Fourier transformed functions of the 3d wave-vector $\mathbf{q}$. On average, the cells in the cluster should be distributed isotropically and $\Sigma(\mathbf{q})$ can only depend on $q \equiv |\mathbf{q}|$. Expanding Σ around $q = 0$ gives us the desired gradient expansion:

$$\begin{aligned} -D_0 q^2 \psi(\mathbf{q}) &= \Sigma(q=0)\psi(\mathbf{q}) + \frac{\Sigma''(q=0) q^2 \psi(\mathbf{q})}{2} + \mathcal{O}(q^4) + s(\mathbf{q}) \\ &\equiv k\psi(\mathbf{q}) + \delta D q^2 \psi(\mathbf{q}) + \mathcal{O}(q^4) + s(\mathbf{q}), \end{aligned} \tag{2.17}$$

where we have identified an absorptive term $k\psi$ and a correction to the diffusion term $\delta D q^2 \psi$. We neglect higher powers of q (i.e., higher order derivatives of the concentration field) in our coarse-grained reaction-diffusion description of nutrient transport in a large colony of cells. Upon returning to real space, we find the desired macroscopic transport equation for the configurationally averaged nutrient concentration:

$$D\nabla^2 \psi(\mathbf{r}) - k\psi(\mathbf{r}) = s(\mathbf{r}) \tag{2.18}$$

where $D = D_0 + \delta D$. The coefficients D and k characterize the macroscopic diffusion and absorption, respectively, of the nutrient in the cell colony. The crucial question is how the effective diffusion constant D and absorption k depend on parameters such as ν and the cell volume fraction ϕ.

In general, we expect that k will be a positive, increasing function of the cell volume fraction ϕ, since the cluster becomes more absorbent as we introduce more cells. It will also increase with $\nu = \kappa a$ as each cell will absorb more nutrients as ν increases. The sign of δD is more subtle, because of two competing factors: The confinement of the nutrient by the cells in the interstitial space will *decrease* the effective diffusion constant D, while the nutrient gradients in random directions induced by nutrient uptake by cells will *increase* D. Since ν controls the amount of absorption, we expect that (for small ϕ) $\delta D < 0$ for $\nu \ll 1$ and $\delta D > 0$ for $\nu \gg 1$. This is confirmed by our effective medium calculation of D in Sec. 2.2.5.

We assumed that the radii of the cells in the cluster are monodisperse. This approximation might not be realistic in cell clusters at different stages of their cell cycle. However, in both polydisperse and monodisperse cases, we can still define a cell volume fraction ϕ. The effective transport coefficients D and k will depend on this volume fraction and the polydispersivity of the radii. For a fixed ϕ, polydisperse cell radii will decrease the total cell surface area and, consequently, *decrease* the nutrient absorption k. For example, the ratio of the total cell surface area S_{poly} of a cluster of cells with a Gaussian radius distribution (with average $\langle a \rangle$ and variance $\sigma_a \ll \langle a \rangle$) to the total surface area S_{mono} for a cluster of cells with the same radius $\langle a \rangle$ is given by [41]

$$\frac{S_{\text{poly}}}{S_{\text{mono}}} = \frac{\langle a^2 \rangle \langle a \rangle}{\langle a^3 \rangle} = \frac{\sigma_a^2 + \langle a \rangle^2}{3\sigma_a^2 + \langle a \rangle^2} < 1. \tag{2.19}$$

Thus, a theory of monodisperse cells is reasonable provided $\langle a \rangle \gtrsim \sqrt{3}\,\sigma_a$. Also, we don't expect the correction to vary much with ν since the main effect of polydispersivity seems to be from the surface area decrease. Finally, computer simulations for two discrete sphere sizes reveal that the ratio $k_{\text{bidisp}}/k_{\text{mono}}$ of absorption coefficients for (perfectly absorbing) cells satisfies $k_{\text{bidisp}}/k_{\text{mono}} \approx (S_{\text{bidisp}}/S_{\text{mono}})^2$ [57, 58] over a wide range of packing fractions. Thus, it may be possible to approximate polydispersity in our theory by reducing the absorption coefficient k by such a geometric factor. However, unless stated otherwise, we henceforth ignore this complication and instead consider cell clusters with monodisperse radii.

2.2.3 The Macroscopic Screening Length ξ

An important nutrient screening length associated with Eq. (2.18) for cell clusters is

$$\xi \equiv \sqrt{\frac{D}{k}}. \tag{2.20}$$

We now use effective medium theory to calculate ξ and relate this length to cell configurations and single cell nutrient uptake kinetics.

Consider first a spherical cell cluster (or "colony") of radius b in which ψ satisfies the effective medium result Eq. (2.18). Assume as well that $\nabla^2\psi = 0$ outside of this spherical region. Then, with the boundary condition $\psi(r \to \infty) = \psi_\infty$ and continuity of the nutrient current at $r = b$, we find

$$\psi(r) = \begin{cases} \dfrac{D_0\psi_\infty\xi}{Dr\cosh(b/\xi)}\,\dfrac{\sinh(r/\xi)}{1+(D_0/D-1)T(b/\xi)} & r \le b \\ \\ \psi_\infty\left[1+\dfrac{b}{r}\left[\dfrac{T(b/\xi)-1}{1+(D_0/D-1)T(b/\xi)}\right]\right] & r > b \end{cases}, \tag{2.21}$$

where $T(x) \equiv \tanh(x)/x$. When $b \gg \xi$, the concentration field $\psi(r) \approx \frac{D_0\psi_\infty\xi}{Dr}\exp[(r-b)/\xi]$ near the colony surface, and thus decays exponentially as we move further into the interior. Thus, ξ is a nutrient penetration depth.

Equation (2.21) leads to the total flux I_{cluster} into the cell *cluster* in the same way as the single cell discussed above. We will see in the next two sections that a good approximation is $D \approx D_0$. Then, the nutrient flux into the cluster is

$$I_{\text{cluster}} = 4\pi D_0\psi_\infty\left[b - \xi\tanh\left(\frac{b}{\xi}\right)\right]. \tag{2.22}$$

Thus, when $\xi \ll b$ so the nutrient does not penetrate far into the cell colony, $I_{\text{cluster}} \approx 4\pi D_0\psi_\infty b$ and the entire colony acts as if it is a single, perfectly absorbing sphere with radius b. Conversely, if $\xi \gg b$, then nutrient penetrates deep into the colony. In this limit we have

$$I_{\text{cluster}} = \frac{4\pi D_0\psi_\infty b^3}{3\xi^2} + \mathcal{O}(\delta^4), \qquad \delta = \frac{b}{\xi} \ll 1, \tag{2.23}$$

so that the total nutrient absorption is down relative to a perfectly absorbing colony by a factor of $\delta^2/3 \ll 1$. I_{cluster} now scales with colony volume ($\propto b^3$), since each cell in the colony contributes to the nutrient uptake.

A related biologically relevant parameter is the thickness ℓ of the outer shell of actively growing cells, as illustrated in Fig. 2.1.1 and by the green band in Fig. 2.2.3(b): Suppose that cells require some minimum concentration ψ_{min} of nutrient in order to grow (or exhibit some level of growth-coupled fluorescence). We can estimate ℓ by finding the location r_{min} inside a spherical colony such that $\psi(r_{\text{min}}) = \psi_{\text{min}}$ [see Eq. (2.21)] and setting $\ell \equiv b - r_{\text{min}}$. For $b \gg \xi$ and $D \approx D_0$, this length is related to b, ξ, ψ_{min} and ψ_∞ via (see Fig. 2.2.4)

$$\ell \approx b + \xi W\left[-\frac{\psi_\infty}{\psi_{\text{min}}}\, e^{-b/\xi}\right], \tag{2.24}$$

where W is the Lambert-W function [59]. The thickness ℓ can also be determined experimentally using a fluorescence reporter, as shown in Fig. 2.1.1. Experimental estimates of ℓ in yeast cell colonies are discussed in Sec. 2.4.

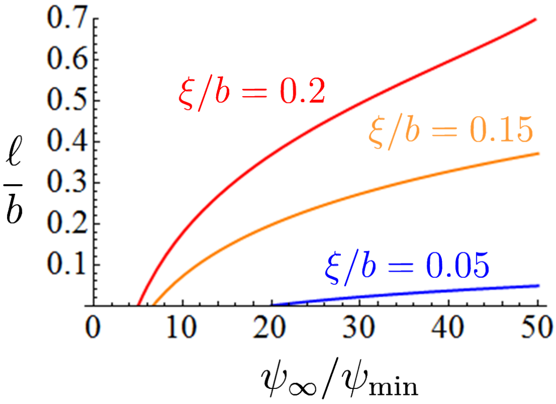


Figure 2.2.4: Plot of the thickness ℓ of actively growing cells in a spherical cluster of size b as a function of the nutrient bath concentration ψ_∞ for various values of the penetration depth ξ.

2.2.4 The Dilute Cell Limit

As mentioned in Sec. 2.1, dilute clusters of microbial cells can be realized in experiment by embedding the cells in a matrix. The initial cell density in such a matrix can be tuned over a wide range and is often quite small. Moreover, since the matrix usually has a negligible influence on the diffusion of small molecules like glucose [60], we now develop an effective medium theory for clusters with cell densities low enough to neglect interactions. This theory will serve as an important and instructive limiting case.

We approximate the nutrient flux into each cell with the single cell result Eq. (2.10), replacing ψ_∞ with the local value $\psi(\mathbf{r})$, where $\mathbf{r}$ is the location of the cell. This approximation does not take into account the finite size of our cells or the deformation of the concentration field $\psi(\mathbf{r})$ around each cell. Thus, we cannot use this dilute limit to find a correction to the diffusion constant D. However, as we will see in the next section, we expect this to be a very small correction, especially for small packing fractions ϕ.

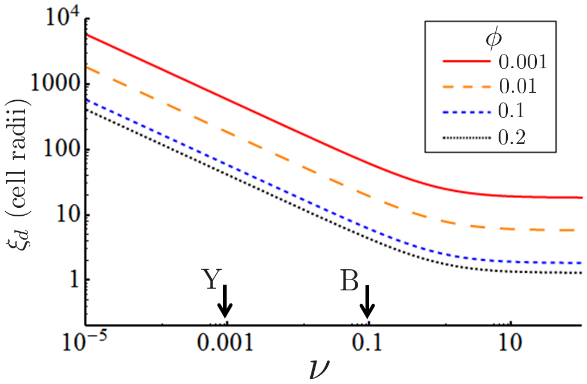


Figure 2.2.5: Plot of the screening length ξ_d in the small ϕ limit as a function of ν for various values of ϕ on a log-log scale. ξ crosses over to the perfectly absorbing limit at $\nu \sim 1$. The arrows Y and B denote, respectively, typical values of ν for baker's yeast and bacteria.

To approximate the nutrient absorption k, we assume that each cell independently contributes to a nutrient flux per unit volume

$$j(\mathbf{r}) = \frac{4\pi D_0 \psi(\mathbf{r}) \nu a n}{1+\nu} = \frac{3 D_0 \phi \nu}{(1+\nu) a^2}\, \psi(\mathbf{r}), \tag{2.25}$$

where n is the cell number density which we relate to the cell volume fraction $\phi = \frac{4}{3}\pi a^3 n$. We conclude that the macroscopic transport coefficients in this dilute approximation are

$$D = D_0 \text{ and } k = \frac{3 D_0 \phi \nu}{(1+\nu) a^2}. \tag{2.26}$$

More sophisticated theories can find a correction to D that is linear in ϕ and higher order corrections to k (see Ref. [40] for a review). Equation (2.26) implies that the screening length $\xi = \sqrt{D/k}$ in the dilute limit is given by

$$\xi \to \xi_d \equiv \sqrt{\frac{1+\nu}{3\phi\nu}}\, a. \tag{2.27}$$

Fig. 2.2.5 shows ξ_d as a function of the absorption parameter $\nu = \kappa a$: the screening length is a strong function of ν for $\nu < 1$ and crosses over to a perfectly absorbing regime when $\nu > 1$. If the cells in a dilute cluster at low volume fraction have polydisperse radii in a Gaussian distribution of cell radii with average $\langle a \rangle$ and standard deviation $\sigma_a \ll \langle a \rangle$, Eq. (2.27) becomes

$$\xi_{d,\,\mathrm{poly}} = \sqrt{\frac{1+\nu}{3\phi\nu}} \left[\langle a \rangle + \sqrt{3}\,\sigma_a\right]. \tag{2.28}$$

As expected, the reduced total cell surface area in the polydisperse cluster leads to a larger screening length. Note that the ratio $\xi_{d,\,\mathrm{poly}}/\xi_d$ is independent of ν and ϕ in this limit.

For yeast cells with $\phi \approx 10^{-1}$ and $\nu \approx 10^{-3}$, Eq. (2.27) predicts $\xi \approx 60a$. Thus, glucose penetrates far into the yeast cell cluster when $\phi \approx 0.1$ (e.g., for yeast embedded in gel), allowing a substantial fraction of the yeast population to grow. Conversely, we expect that $\xi \approx 6a$ for a bacterial colony ($\nu \approx 10^{-1}$) at $\phi \approx 0.1$ volume fraction. We clearly need to go beyond the dilute limit and determine the dependence of ξ on ϕ and ν more carefully: Yeast and other cell colonies rarely grow at low volume fractions; cells typically clump together and pack themselves in an amorphous structure with a volume fraction approaching that of the random close packing density, $\phi \approx 0.6 - 0.7$ [61].

2.2.5 Effective Medium Theory for Dense Cell Clusters

For $\phi \gtrsim 0.1$, the dilute approximation breaks down, and we must solve Eq. (2.12) for $\Psi(\mathbf{r})$ more exactly. Upon defining the diffusive Green's function $G_0(\mathbf{r}) = (4\pi D_0 r)^{-1}$, we can rewrite Eq. (2.12) as an integral equation,

$$\Psi(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\, G_0(\mathbf{r} - \mathbf{r}') \left[\sum_{i=1}^{N} \int \mathrm{d}\Omega_i\, \sigma_i(\mathbf{r}_i)\, \delta(\mathbf{r}' - \mathbf{R}_i) + s(\mathbf{r}') \right], \tag{2.29}$$

where we integrate over all positions $\mathbf{r}'$ ($\mathrm{d}\mathbf{r} \equiv \mathrm{d}^3 r$). In principle, we could choose the functions $\sigma_i(\mathbf{r}_i)$ to enforce the radiation boundary condition at each cell surface [Eq. (2.13)], and then average over the cell positions to obtain $\Sigma(\mathbf{r})$ in Eq. (2.15). However, as discussed in Section 2.2.2, we do not have access to the full probability distribution $P(\{\mathbf{R}_i^0\}_{i=1}^N)$ necessary to do the averaging. An alternative approach exploits an approximate solution which accounts for each cell independently to first order, the effects of pairs of cells at second order, etc. This expansion requires knowledge of the probability distributions of a single cell position, a pair of cells, a triplet, etc. Such scattering expansions are plagued with divergences due to the long-range nature of the diffusive interaction between the cells and must be treated via careful resummations [54, 62, 63]. Particular care is required, because the Green's function, or propagator, in this expansion is $\hat{G}_0(\mathbf{q}) = (D_0 q^2)^{-1}$ in Fourier space, which is singular as $q \to 0$.

To bypass such complications, we imagine cells imbedded in an "effective medium" with effective transport properties, to be determined self-consistently. These transport properties include screening which cuts off the long range behavior of $G_0(r)$, and renders $\hat{G}(\mathbf{q})$ finite as $\mathbf{q} \to 0$. In particular, we assume a modified "effective medium" propagator given by

$$\hat{G}_\Sigma(\mathbf{q}) \equiv \frac{1}{D_0 q^2 + \Sigma(\mathbf{q})}, \tag{2.30}$$

which describes the transport properties of the nutrient as it wanders through the homogeneous attenuating medium shown in Fig. 2.2.3(b). A self-consistency condition on $\Sigma(\mathbf{q})$ leads to a renormalized (and better behaved) scattering expansion.

In Appendix A.1, we follow Cukier and Freed's effective medium calculation of $\Sigma(\mathbf{q})$ [64]. However, our cells will be partially absorbing, which generalizes the perfectly absorbing case Cukier and Freed considered. The calculation assumes that the cells have uniformly distributed centers in the cluster. Excluded volume interactions are included in an approximate way, by assuming that the centers $\mathbf{R}^0_{i,j}$ of any pair of cells i and j are distributed according to the low density hard sphere pair distribution function

$$P(\mathbf{R}^0_i, \mathbf{R}^0_j) = \frac{1}{V^2}\,\theta(\left|\mathbf{R}^0_i - \mathbf{R}^0_j\right| - 2a), \tag{2.31}$$

where $\theta(x)$ is the step function and V is the cluster volume. This generalization was also considered by Cukier [65], but we find an important second order correction to his results from the pair distribution function Eq. (2.31). The details are contained in Appendix A.1.

We find that the nutrient screening length in units of the cell radius $\alpha \equiv \xi/a$ satisfies

$$\alpha^2 = \frac{3\phi\alpha\nu(1+\coth\alpha)}{1+\alpha+\nu} + 36\phi^2\left[\frac{\alpha\nu(1+\coth\alpha)}{1+\alpha+\nu}\right]^2 \times\left[\frac{1}{4\alpha^2} + k_1(2\alpha)i_0(\alpha)\left[\frac{\alpha i_1(\alpha)}{\nu} - i_0(\alpha)\right]\right], \tag{2.32}$$

where $i_\ell(x)$ and $k_\ell(x)$ are the modified spherical Bessel functions of the first and second kind, respectively. For a specific $\nu = \kappa a$, Eq. (2.32) is a self-consistent equation for $\alpha = \xi/a$. It is soluble numerically (we used Newton's method of successive approximation) for ξ as a function of ν and ϕ.

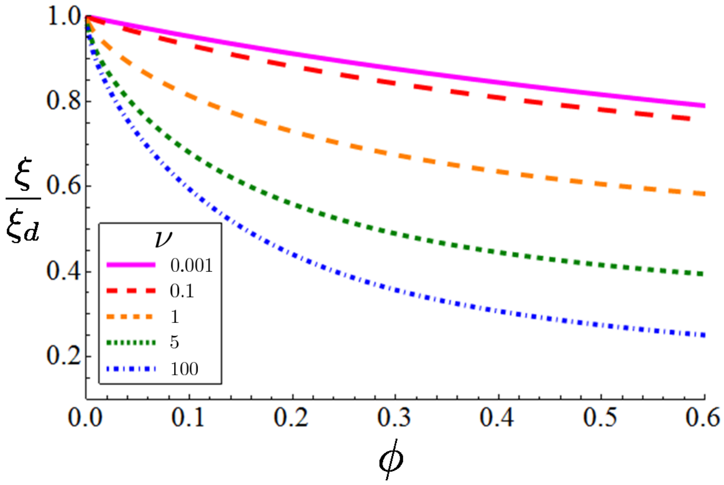


Figure 2.2.6: Plot of the screening length ξ calculated with the effective medium theory divided by the dilute limit result ξ_d as a function of the volume fraction ϕ for various values of ν. The low density result becomes increasingly inaccurate at larger packing fractions.

The ratio ξ/ξ_d, where ξ_d is the dilute limit formula Eq. (2.27), appears as a function of ϕ for various ν in Fig. 2.2.6. Note that the effective medium theory always predicts $\xi < \xi_d$, a plausible result since this more sophisticated theory allows for nutrient absorption mediated by repeated interactions with individual cells.

For the values of ν typical of yeast clusters and gram negative bacteria ($\nu \sim 10^{-3} - 10^{-1}$), ξ is no more than 20% different from ξ_d even for large $\phi \sim 0.5$. These results suggest that effective medium theory is a reasonable approximation for dense clusters of cells with $\nu = \kappa a \ll 1$, in contrast to the more problematic perfectly absorbing case $\nu \to \infty$ [1].

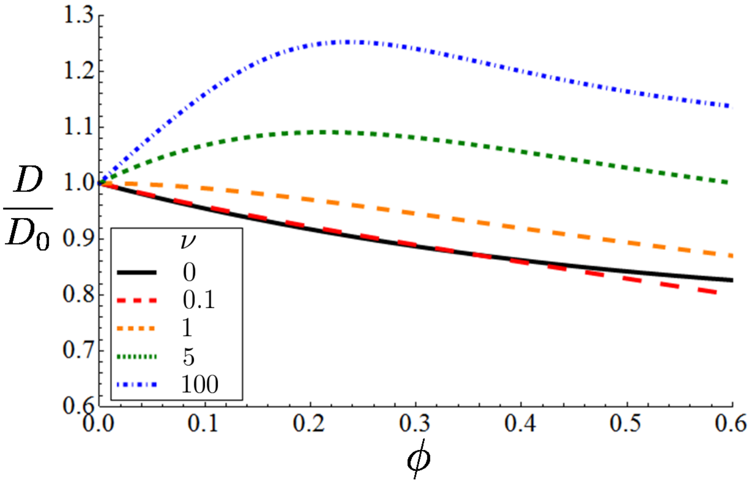


Figure 2.2.7: Plot of the diffusion coefficient D in the cell cluster calculated with the effective medium theory divided by the bare diffusion coefficient D_0 in the absence of the cluster as a function of the volume fraction ϕ for various values of ν.

The second order term in the Taylor series expansion of $\Sigma(\mathbf{q})$ around $\mathbf{q} = 0$ yields an implicit equation for the effective diffusion coefficient D of the nutrient inside the cell cluster. The somewhat cumbersome equation is Eq. (A.20) in Appendix A.1. The solutions for various ν are plotted in Fig. 2.2.7 as a function of ϕ. In all cases of biological interest, the corrections to the bare diffusion coefficient D_0 in the absence of the cluster are relatively small, i.e., approximately 10-20% and never exceeding 30%.

---

[1] The effective medium calculation in the perfectly absorbing $\nu \to \infty$ limit has serious deficiencies: even the second order correction calculated in Appendix A.1 yields unphysical results for large values of ϕ [64].

In the biologically relevant regime $\nu \lesssim 0.1$, the correction to the diffusion coefficient is close to the $\nu = 0$ (perfectly reflecting) limiting case given by

$$\frac{D(\nu = 0)}{D_0} = \frac{1}{1 + \phi/2 - \phi^2/4}, \tag{2.33}$$

which is derived from Eq. (A.20). As expected, the diffusion coefficient D is smaller than D_0 when the cells are perfectly reflecting. Also, the structure of Eq. (2.33) is consistent with the modified Maxwell-Garnett theory for diffusion of particles in a medium with perfectly reflecting inclusions derived in Ref. [66]. Equation (2.33) is also consistent with experimental results for the diffusion coefficient of water in a suspension of spherical colloidal particles [67]. At high volume fractions, we expect the effective medium theory to start to break down and higher order terms in $\Sigma(\mathbf{q})$ to contribute to transport within the colony. Thus, the curves in Fig. 2.2.7 should be treated as approximations and specific features, such as the crossover between the $\nu = 0$ and $\nu = 0.1$ lines, might not be physical.

We now check the effective medium result against simulations and experiments.

2.3 Numerical Simulations

To obtain further insight into nutrient absorption, we now go beyond effective medium theory and solve the diffusion equation numerically in a densely randomly packed cluster of spherical cells. We employ the finite element solver discussed below to numerically solve the steady-state diffusion equation with the appropriate boundary conditions. Given the complicated nature of the problem, we can only test effective medium theory for clusters of up to 400 cells in this way. However, since the effective medium approximation is derived in the limit of an infinite number of cells, if it works for simulations with 100-400 cells, it should be even more reliable when we have even more cells, as in a typical growing yeast colony.

2.3.1 Cell Cluster Simulation

To check the analytic results, we solved the steady state diffusion equation exactly for cell clusters with hundreds of cells. The numerical solution was found with the COMSOL 3.5a finite element solver [68]. A MATLAB program was written to input in the locations and radii of all the cells in the cluster. The COMSOL program included a computer-assisted design (CAD) feature that was then able to parse the MATLAB output and create a particular arrangement of spherical cells that defined our domain of interest. The coordinate list for the sphere cluster was created via a Bennett model, originally designed to quickly generate amorphous, dense random packings of identical spheres [69]. These arrangements approximate the disordered packing of cells observed in yeast colonies in the experiments.

To prevent the cells from touching and creating singularities in the finite element mesh, identical spherical cells with radius a are placed at the sphere centers of a cluster generated by the Bennett model using spheres with larger radius $\tilde{a} > a$. This guarantees a gap of at least $2(\tilde{a} - a)$ between adjacent spheres. A high volume fraction $\phi \approx 0.63$ (corresponding to random close packing) is generated with $\tilde{a} = 1$ and $a = 0.999$. Smaller values of ϕ are generated by decreasing a for a fixed $\tilde{a} = 1$. For example, $a = 0.9$ results in a volume fraction $\phi \approx 0.48$ (e.g., the cluster in Fig. 2.3.1).

A large bounding sphere concentric with the center of mass of the cell cluster allowed us to impose a constant nutrient concentration "at infinity": $\Psi(|\mathbf{r}| \to \infty) = \psi_\infty$, to approximate a suspended cell cluster in an infinite nutrient bath buffered at concentration ψ_∞. We used COMSOL's "infinite element" option to efficiently simulate by placing extremely large finite elements between the bounding sphere and the cell cluster. With bounding spheres of radii typically five times the cell cluster radius, this method yielded good results: Our calculated solution was unchanged when varying the bounding sphere radius from 5 to 10 times the cluster radius.

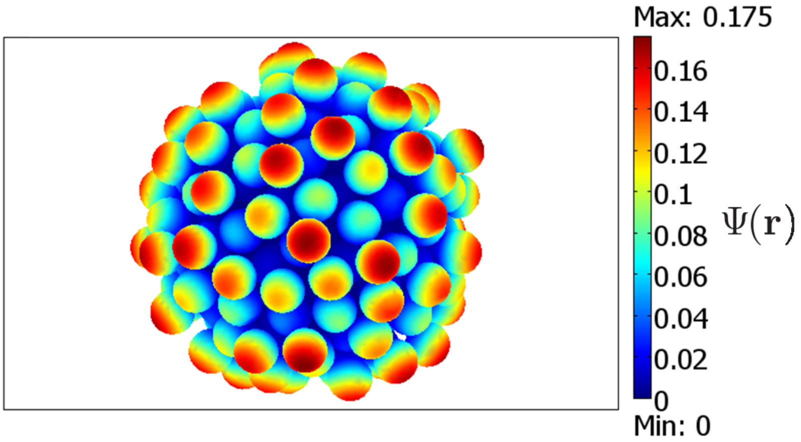


Figure 2.3.1: A simulated cluster of cells with $\nu = 1$ and $\psi_\infty = 1$. The color indicates the local nutrient concentration $\Psi(\mathbf{r})$ near the cell surfaces. In this case, the 200 cells with radius a were confined to a cluster of radius $b \approx 7.45a$, yielding a volume fraction of $\phi \approx 0.48$. The radius of the large bounding sphere on which the nutrient concentration was fixed at ψ_∞ was about $56a \approx 7.5b$. To increase computation efficiency, we used COMSOL's "infinite element" option with spherical symmetry between radii $28a$ and $56a$ (see Sec. 2.3.1).

Finally, we specified the dimensionless nutrient uptake parameter $\nu = \kappa a$ for cells with identical radius a and applied the boundary conditions given by Eq. (2.13) on each cell surface. A particular cell cluster and the corresponding steady state nutrient concentration $\Psi(\mathbf{r})$ near each cell surface is shown in Fig. 2.3.1. In this case, $\nu = 1$, the packing fraction $\phi \approx 0.48$, and the cluster radius $b \approx 7.45a$, with $\psi_\infty = 1$. The bounding sphere had a radius of $56a$ and we inserted infinite elements at distances between $28a$ and $56a$. With these parameters, the effective medium theory predicts a screening length $\xi \approx 0.72a$. This is consistent with what we observe in Fig. 2.3.1, as the concentration decays by a factor of e over a distance comparable to the cell radius a.

2.3.2 Comparison with Theoretical Results

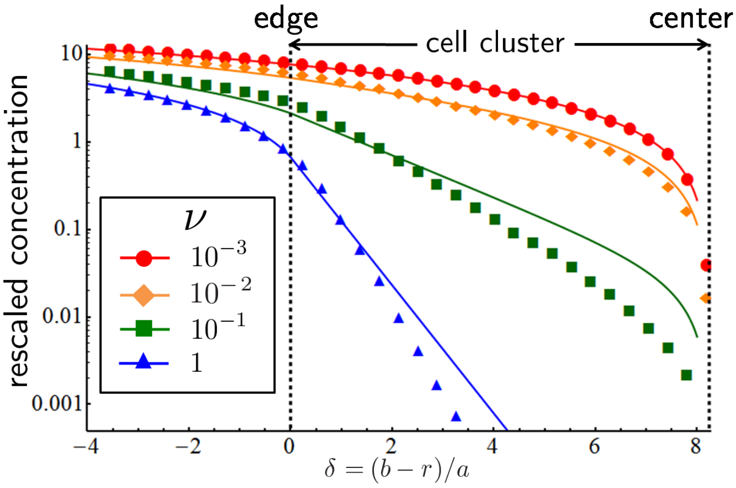


Figure 2.3.2: The radially averaged concentration $r\overline{\Psi}(r)/a\psi_\infty$ (symbols) and the effective medium result $r\psi(r)/a\psi_\infty$ (lines) for a cluster of 353 cells of radius a (with cluster radius $b \approx 8.15a$) as a function of $\delta = (b - r)/a$ on a log-linear plot. The concentrations are rescaled to highlight the exponential decay of the concentration into the bulk of the cluster. Cells occupy the region $0 < \delta < 8.15$ with the center at $\delta \approx 8.15$, as indicated by the dashed lines. The cell packing fraction is $\phi \approx 0.63$. The bounding sphere has a radius of $50a$, with variable mesh elements inserted at distances between $25a$ and $50a$.

To compare theoretical and simulation results, we tracked the decay of the radially averaged nutrient concentration $\overline{\Psi}(r) \equiv (4\pi)^{-1} \int \Psi(\mathbf{r})\, d\Omega$ into the center of the colony, and compared with the effective medium prediction $\psi(r)$ of Eq. (2.21). Fig. 2.3.2 shows a semilogarithmic plot of $r\psi(r)/a\psi_\infty$ versus $\delta = (b - r)/a$, varying $\nu = \kappa a$ over 3 orders of magnitude. Equation (2.21) predicts that this quantity decreases exponentially near the cell surface when $1 \ll b/a$. Figure 2.3.2 shows how the nutrient concentration decreases as we move into the cluster.

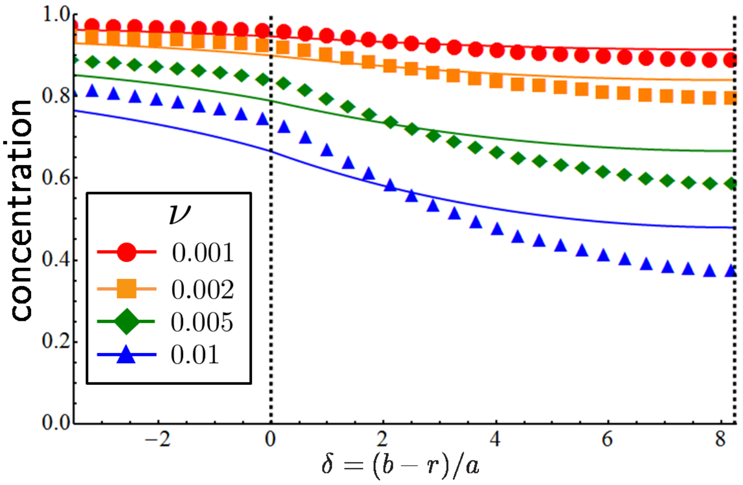


Figure 2.3.3: The radially averaged concentration $\overline{\Psi}(r)/\psi_\infty$ (symbols) and the effective medium solution $\psi(r)/\psi_\infty$ (lines) as a function of $\delta = (b - r)/a$ for the same 353 cell cluster described in Fig. 2.3.2. The dotted lines indicate the cluster edge ($\delta = 0$) and center ($\delta \approx 8.15$).

Figure 2.3.2 shows that, even for the large volume fraction $\phi \approx 0.63$, effective medium theory provides an excellent description, especially for small values of ν. Thus, this theory is appropriate for modelling yeast colonies, which have a very small $\nu \sim 10^{-3}$ value. The effective medium agreement should improve if computer resources allow more cells in the simulation, since it was designed to handle the limit where the cell number $N \to \infty$ and cluster volume $V \to \infty$, with N/V fixed. Figure 2.3.3 shows the low ν regime (poor nutrient absorbtion) in more detail. Although the differences between the simulation and effective medium theory are now more evident, the absolute difference between the simulation and the theory for the concentration remains small.

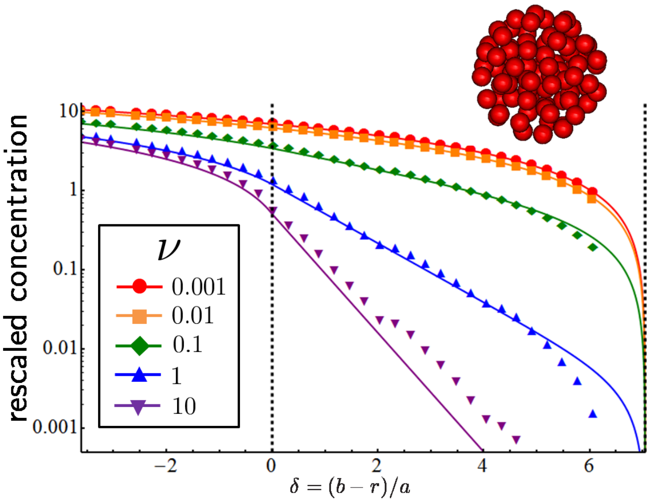


Figure 2.3.4: Exponential decay of the rescaled nutrient concentration in the effective medium theory (lines) and in a simulation (symbols) for a more dilute 85 cell cluster with $b = 7a$ and various ν. The cell packing fraction is $\phi \approx 0.24$. The cluster, shown above the graph, now occupies the region between the dotted lines, $0 < \delta < 7$.

We also studied a more dilute 85 cell cluster, to check that effective medium theory is indeed accurate for lower volume fractions and an alternative cell configuration. Instead of using the Bennett model, we placed cells with centers in three concentric shells with radii $2a$, $4a$ and $6a$. The cells within each shell were placed randomly, but their positions were adjusted to prevent the cells from touching and disrupting the finite element mesh. Figure 2.3.4 reveals even better agreement between the simulation and effective medium theory for the lower volume fraction of $\phi \approx 0.24$. Note that good agreement in this case is obtained for the highly absorbing limit $\nu \geq 1$, as well. We conclude that effective medium theory provides a good description of spherically averaged nutrient uptake, for both weakly and highly absorbing cells, for volume fractions $\phi \lesssim 0.6$.

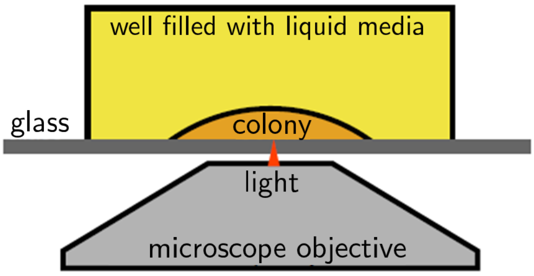
2.4 Experiments

We now compare the effective medium theory presented in Sec. 2.2 to experimental results for glucose uptake in yeast cell colonies. Yeast cell aggregates are a particularly interesting biological application because they form naturally in the wild and are a possible model for the initial emergence of multicellularity [70]. Nutrient uptake in single versus clumped cells is a crucial factor in this model system. Thus, accurately modelling nutrient transport properties in cell clusters is key to understanding their biological function and role in evolution.

Since yeast cells are poor glucose absorbers, the characteristic nutrient penetration depth in a yeast colony is much larger than a cell radius and is easily visible in colony cross sections, as shown in Fig. 2.1.1. Their poor absorption properties are also well-suited for effective medium modelling, since we expect that nutrient transport for small absorption parameter $\nu \ll 1$ is dominated by much longer length scales than the local colony geometry at the single cell scale.

2.4.1 Experimental Setup

(a) Confocal microscopy setup



(b) Fluorescence of individual cells

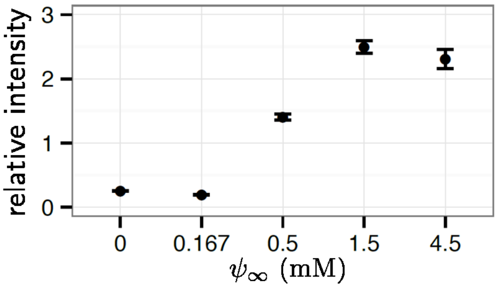


Figure 2.4.1: (Color online) (a) Confocal microscopy setup. A single yeast cell was inoculated into a well of a glass-bottomed 96-well plate containing 200 μL of yeast media and allowed to grow into a colony. The bottom of the colony was imaged using a spinning-disc confocal inverted microscope. (b) Cells were grown in shaken media, pipetted into an empty well, and imaged immediately (before they could divide and form colonies) with the setup shown in (a). At least 250 cells at each concentration were imaged, and the error bars are the 95% confidence interval using the Student's t-test. Relative fluorescence intensity in (b) is defined as ribosomal protein expression divided by constitutive expression. Cells in 0.167 mM glucose continue to divide, but at less than half the growth rate of cells in 0.5 mM glucose.

We used confocal microscopy to measure the distribution of growth in a yeast colony. To start the experiment, we filled each well of a 96-well, glass-bottomed microtiter dish (Greiner Bio-One 655892) with 200 μL of minimal, synthetic yeast media (recipe in Ref. [70]). Figure 2.4.1(a) shows the confocal setup and Fig. 2.1.1 shows representative images. We engineered a yeast strain (yJHK192) with two fluorescence protein (FP) reporters: (1) Cerulean, a stable CFP (cyan) variant, was constitutively expressed from the *ACT1* promoter; and (2) mCitrine, a YFP (yellow) variant, was expressed from the ribosomal protein *RPL7A* promoter. The expression of *RPL7A* has been shown to correlate with growth rate [71]; the correlation between growth rate and glucose concentration has been well studied in yeast [72], and we verified that the YFP fluorescence correlated with glucose concentration in cells grown in shaken media (Figure 2.4.1(*b*)). mCitrine was destabilized by adding a *CLN2* PEST-sequence to the C terminus [73]; a destabilized fluorescent protein was essential to see decreased expression of *RPL7A* in cells that had stopped growing.

When grown on agar, small yeast cell colonies (less than 1 mm across) form a spherical cap with a contact angle θ between the edge and the agar surface. The contact angle θ increases with agar concentration and $\theta \approx 40^\circ$ for 2% agar [74]. Unfortunately, in our experiment, the side view of the colonies was optically inaccessible and a precise characterization of the colony morphology was not possible. However, since our colonies are very small, we expect that surface energy minimization is most important in determining the shape so that the colonies are approximately spherical caps.

We grew colonies in three glucose concentrations: 0.5 mM, 1.5 mM, and 4.5 mM. We inoculated one cell per well using a fluorescent activated cell sorter (MoFlo FACS, Beckman Coulter, Inc.); the cells had been grown to saturation in 1 mM glucose synthetic media. Inoculation was verified under a microscope, and those wells with a cell closest to the middle of the well (at least 3 wells per glucose concentration) were selected for colony tracking. The plate was incubated without shaking at 25 °C for 3 days. Each colony was imaged a day after inoculation and an additional 2-3 times over the next 2 days. The radius of the largest imaged colony in each glucose concentration was less than half of the average colony size after two weeks of growth (610 μm in 0.5 mM glucose, 920 μm in 1.5 mM glucose, and 1170 μm in 4.5 mM glucose), indicating that the carbon source was not yet depleted. Images were taken with a 20X objective on a Nikon inverted Ti microscope with a Yokagawa spinning disc unit and an EM-CCD camera (Hamamatsu ImagEM); CFP was excited with a 447 nm laser and YFP was excited with a 515 nm laser; exposure times for all images (including single-cell images) were 200 ms (CFP) and 1000 ms (YFP). All images were focused on the bottom layer of cells in the colony, and multiple, overlapping images were taken of colonies that exceeded the field of view. Three independent experiments were performed. Figures 2.4.2 and 2.4.3 show the combined results of all experiments.

2.4.2 Data Analysis and Results

Images were processed using the Fiji distribution of ImageJ [75]. Images were converted to 8-bit, stitched together [76], and merged into a single RGB image. Contrast was not adjusted during processing. Fluorescence intensity as a function of colony radius was measured using a custom script written in Python. The basic algorithm is as follows: (1) The constitutive (CFP) image was thresholded using Li's Minimum Cross Entropy thresholding method [77]. (2) Noise and cells not attached to the colony were removed by eroding and then dilating the binary image a total of three times with a four pixel diameter circular structuring element. (3) A series of mask "rings" was made by a series of morphological erosions and subtractions from the original image. The distance between rings was approximately 8 μm, or 2 cell diameters. (4) The masks were used to generate images of concentric rings using both the constitutive (CFP) and the growth-dependent expression (YFP) image. (5) The average fluorescence of each ring was measured; the reported, relative fluorescence is the average expression (YFP) fluorescence divided by the average constitutive (CFP) fluorescence. Figure 2.4.2 shows the intensity as a function of distance from the edge of the colony; each point in the figure is a measurement of fluorescence at one of the concentric rings, and a line joins the values from a single colony.

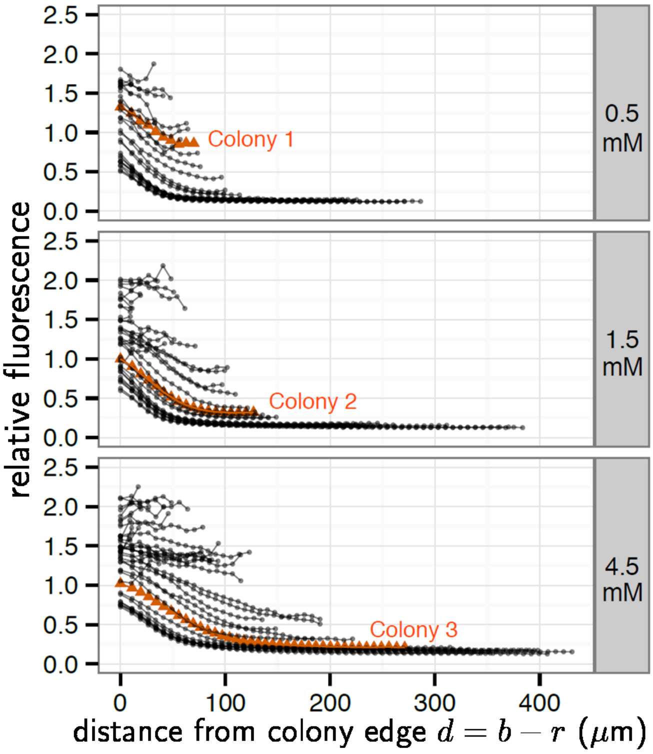


Figure 2.4.2: (Color online) Relative fluorescence intensity (ribosomal protein expression divided by constitutive expression) as a function of distance $d = b - r$ from the edge of the colony. Here b is the colony radius and r is the observation position. Each point is the average fluorescence of an 8 μm thick ring whose edge is located an equal distance from the edge of the colony. The last point on each connected set of lines represents the radius of the colony and larger radii reflect longer periods of incubation. Each line is one colony. Colonies were grown in 0.5 mM glucose (top), 1.5 mM glucose (middle), and 4.5 mM glucose (bottom). Profiles from the three colonies shown in Fig. 2.1.1 are shown as orange triangles.

In order to compare the experimental results to the effective medium theory, we assume that each colony is a hemisphere with a radius b equal to the radius of the bottom cross-section. Since the colonies are small compared to the glass well dimensions and since the glucose molecules do not stick to the glass, we also assume that the hemispheres sit on an infinite, reflecting surface. Hence, by reflection symmetry, the concentration profile $\psi(r)$ is given by the solution for spherical colonies [Eq. (2.21)], where r is the distance from the colony center (measured above the infinite plane). Of course, as discussed in the previous section, the colonies in the experiment could be spherical caps with a smaller contact angle $\theta < 90^\circ$ at the glass well bottom. A smaller contact angle corresponds to a more shallow colony with an increased nutrient penetration depth, due to nutrient diffusion from the top of the colony. However, we expect that the corrections due to smaller θ to be small when the penetration depth is small compared to the colony radius.

Some predictions of the effective medium theory are qualitatively confirmed by the experimental data. The effective medium theory described in Sec. 2.2 predicts that for these dense cell colonies, the characteristic shielding length is 50 μm, or about 12 cell diameters. This is consistent with the fluorescence curves in Fig. 2.4.2. The theory also predicts that the glucose level at the outer edge of the colony decreases like $1/b$, where b is the colony radius [see Eq. (2.21)]. We see this effect in Fig. 2.4.2, where shorter curves (corresponding to smaller colonies) have a higher fluorescence level at $d = 0$. However, the fluorescence level has a complicated, non-linear relationship to the glucose concentration in the bulk medium shown (see Fig. 2.4.1(b)) and the specific $1/b$ scaling cannot be tested.

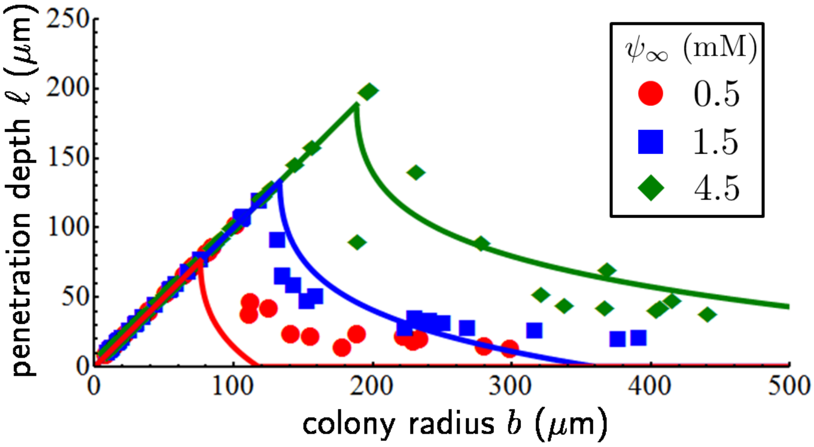


Figure 2.4.3: (Color online) Penetration depth ℓ as a function of colony radius b. Penetration depth in the experiment (symbols) is the distance from the edge of the colony at which relative fluorescence drops below 0.4. The theoretical results (solid lines) are calculated by numerically solving $\psi(r = b - \ell) = \psi_{\min}$ [Eq. (2.21)] for ℓ, where we assume that $\psi_{\min} \approx 0.25$ mM local glucose concentration corresponds to a 0.4 relative fluorescence level. As discussed in the text, this level is slightly higher than the level at which the cells stop growing and expressing the ribosomal protein. We also use a cell radius of $a = 2$ μm, a cell packing fraction $\phi = 0.56$, and $\nu = 6 \times 10^{-4}$.

Figure 2.4.3 shows the penetration distance ℓ as a function of colony size; penetration distance is defined here as the distance from the edge of the colony at which the relative fluorescence intensity falls to less than 0.4. This intensity is chosen to be well above the background fluorescence level. For the smaller colonies, the fluorescence is higher than 0.4 throughout the colony. For these colonies, we say that the penetration depth is equal to the colony radius: $\ell = b$. Equation (2.21) is used to find a theoretical result for the penetration depth ℓ inside the colony. We assume that the relative fluorescence intensity of 0.4 corresponds to a local glucose concentration of $\psi_{\min} \approx 0.25$ mM. Again, note that this level is slightly higher than the minimum glucose level required for growth. We estimate the packing fraction of cells inside a colony, $\phi \approx 0.56$, by looking at cell colonies growing on agar (data not shown). Then, assuming a cell radius $a = 2$ μm and absorption parameter $\nu = 6 \times 10^{-4}$, we find reasonable agreement between the theoretical solution for ℓ and the experimental data in Fig. 2.4.3.

We do see some disagreements between the experimental results in Fig. 2.4.3, especially at the 0.5 mM glucose concentration. In general, we expect to underestimate the penetration depth due to our approximation of colonies as hemispheres. The deviations may also indicate the presence of biological mechanisms that allow the cell clusters to adapt to particular nutrient environments. For example, yeast cells can modify their nutrient uptake parameter $\nu = \kappa a$ by expressing different nutrient transporters on their surface [25]. Our theory assumes that each cell in a colony has the same absorption strength ν, an approximation suitable provided all cells employ the same nutrient transporter system, and the ambient nutrient concentration is small enough so that we are in a linear regime of nutrient absorption in a Michaelis-Menten-type curve (see Fig. 2.2.1). So, deviations from our theory may indicate deviations from the linear regime or the presence of regulatory mechanisms in cells that modify nutrient uptake kinetics during colony growth.

2.5 Conclusions

We have developed and tested an effective medium theory of nutrient transport in clusters of cells. Simulations and experiments support the theoretical results. The key parameters in our model are the single cell nutrient absorption parameter ν and the packing fraction ϕ of cells in the cluster.

The wide variation in ν for different cell types such as yeast ($\nu \sim 10^{-3}$), gram-positive ($\nu \sim 10^{-2}$), and gram-negative ($\nu \sim 10^{-1}$) bacteria has important implications for nutrient absorption. Effective medium theory predicts very different nutrient shielding properties of cell clusters as ν varies (see Figs. 2.3.2 and 2.3.4). Thus, the fraction of actively growing cells at a cluster surface will vary significantly with cell types. For instance, we predict that the screening length ℓ for an *S. cerevisiae* colony is about 10 times longer than the length in a gram negative *E. coli* bacteria colony and 3 times longer than in a gram positive *L. mesenteroides* bacteria colony, measured in units of the respective cell diameters.

Our theory also predicts that nutrient shielding is more sensitive to the volume fraction ϕ when each cell is a good absorber (i.e., for $\nu \gtrsim 0.1$). As ϕ increases, correlations between the cells become more important as nutrient collisions with multiple cells create a stronger shielding effect. Thus, a colony of good absorbers, such as gram-negative bacteria, should be able to change its nutrient shielding properties by tuning the separation between each cell. This hypothesis could be tested by placing bacterial cells in a gelatinous matrix at various cell densities and observing their growth. As discussed in Sec. 2.2.4, these artificial colonies have many industrial and environmental applications.

It would be interesting to extend our theory to include nonspherical cell shapes and cell radius polydispersivity. Polydispersivity should increase the screening length, but a detailed understanding of its effect on cell correlations is lacking. One could also account for spatial variability in ϕ. Cell clusters should "thin out" near the cluster surface, where the cells have not had time to grow into a densely packed structure. If this density variation occurs on scales large compared to the cell radii, it should be sufficient to replace ϕ by a spatially dependent $\phi(\mathbf{r})$ in our effective medium calculations.

To extend the experimental results and test the effective medium theory more precisely, one could control for the variability of the ν parameter in the colony. One possibility is to use engineered yeast strains with fixed nutrient uptake kinetics, such as the mutants constructed by Reifenberger et al. [26]. These cells express a single type of glucose transporter in a medium with a low glucose concentration. In addition, it would be interesting to more precisely characterize the local glucose concentration in the colony by using either a more direct reporter or more precisely characterising the relationship between glucose level and ribosomal protein expression.

Chapter 3

Radial Domany-Kinzel Models with Mutation and Selection

This chapter appears in its entirety in M.O. Lavrentovich, K. S. Korolev & D.R. Nelson, *Phys. Rev. E* **87**, 012103 (2013)

3.1 Introduction

To grow and prosper, populations must often migrate into new territory. These ubiquitous range expansions occur in bacterial growth on a Petri dish, tumor growth, viral infections, human migration, species movement induced by climate change, and in many other biological systems [78, 79, 80, 81, 82, 83, 84, 85, 86, 87]. Such expansions influence the evolutionary dynamics of the population with, for example, enhanced genetic drift due to low population densities at the frontier. To understand the universal features of these diverse range expansions, it is of interest to construct simple models with the essential elements of the population's evolutionary and spatial dynamics.

Many features of a range expansion can be captured by adapting a simple "stepping stone" model [13, 88]. In this model, individuals in a collection of island subpopulations, or demes, reproduce, die, and mutate stochastically. Two commonly used stochastic processes are the Moran and Wright-Fisher models (see [10, 13, 89] and references therein). The demes track spatial configurations, which simulate the spatial distribution of an evolving population. By allowing individuals to migrate between adjacent demes, one can simulate the effect of dispersal on the evolutionary dynamics in, say, one or two dimensions. Analysis of these models highlighted many important features distinguishing the evolutionary dynamics of a well mixed versus spatially extended populations (see [13] for a review).

As discussed in Ref. [13], stepping stone models also provide a useful approximation to population genetics at the *frontier* of an expanding population. First, consider an asexual population expanding across a flat surface with an approximately linear front. Under a wide variety of conditions, the front will move with a constant speed v, set by the competition between growth and dispersal (see e.g., Ref. [90] and references therein). This growth is called a Fisher wave [91]. Second, we assume that only individuals living near the population front are able to divide and settle into the new territory. In the approximation of a *flat* population front, and in a reference frame moving with the frontier, we can simulate the front using a one-dimensional stepping stone model. This model predicts that for small deme size, number fluctuations (genetic drift) create a strong spatial genetic demixing of the population into domains of closely related individuals. Initially, demixing occurs primarily through local fixation within a deme. At long times, however, coarsening is dominated by the diffusive motion of the domain boundaries, which annihilate or coalesce on contact, creating larger domains. Selective advantages will bias the motion of the boundaries, and mutations will introduce new domains into the population [13].

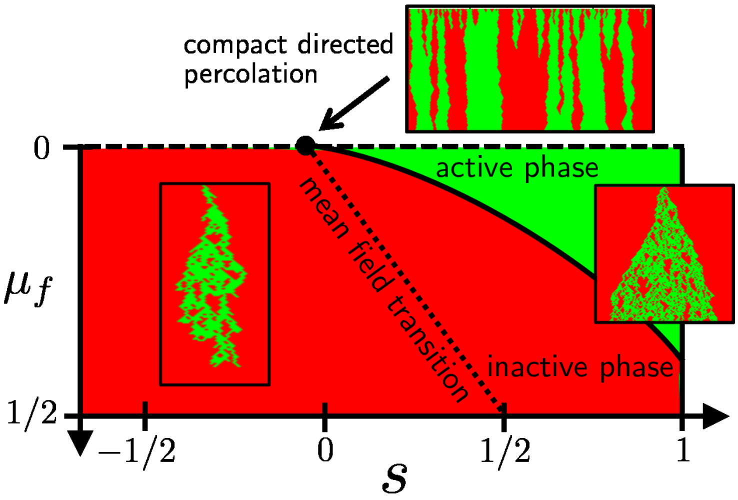


Figure 3.1.1: A sketch of the phase diagram for mutation-selection balance in linear expansions with irreversible, deleterious mutations from green to red cells occurring at rate μ_f, at a fitness cost s (see Appendix B.1 for a biological motivation). The solid line between red and green regions represents the directed percolation phase transition. We show as insets typical clusters generated from single green (light gray) cells in the active and inactive phases. The dashed line at the top represents the case of no mutations where the system falls into the compact directed percolation (CDP) universality class. The dynamics at criticality ($s = \mu_f = 0$ at the black dot) in this CDP regime is shown. The dotted line shows the position of the mean field transition (i.e., for a well mixed population). An analogous plot for inflating radial range expansions is shown in Fig. 3.3.7.

Stepping stone models (and the simplified voter-type models discussed below) of population genetics at frontiers make use of the notion of "dimensional reduction." A range expansion of organisms across a surface takes place in two spatial and one temporal dimensions. At the cost of considerable computational complexity, a full $2+1$-dimensional model of the population genetics could be constructed, as was done recently within mean field theory to model range expansions of baker's yeast, *Saccharomyces cerevisiae* [92]. However, the population density and relative proportions of many organisms stop changing (and some even stop migrating) in the wake of a front colonizing virgin territory. The crucial dynamics of genetic drift, competition, and spatial diffusion then takes place at a linear frontier, where one can approximate the $2 + 1$-dimensional dynamics by a $1 + 1$-dimensional stepping stone in a reference frame which moves with the population front.

This mapping is only exact for perfectly *flat* population fronts. For neutral mutations, the flat front approximation leads to boundaries of genetic domains behaving as simple random walks which wander from their initial positions with a characteristic displacement $\Delta x \propto \sqrt{t}$ that increases with the square root of time. Interesting complications arise when the front undulates – for neutral genetic variants, the boundaries between genetic domains couple to the undulations and wander more vigorously, which changes some of the detailed predictions of the theory [2, 9, 13, 93]. We neglect such complications here, but note that Kuhr et al. [94] have recently studied dynamical phase transitions in a two-species Eden model with an approximately flat but undulating front, and found *different* critical exponents than those of the directed percolation universality class [95] considered here. It's worth noting that not only are flat one-dimensional habitats easier to analyze, a quasi-one-dimensional environment could describe the evolutionary dynamics along the bank of a river, along a seacoast, or along a constant altitude slope of a linear mountain range. To obtain inflation (or deflation) in this context, one could invoke, for example, a receding or advancing waterline around an island. Flat front models might also describe the gradual, climate-driven shift of some populations toward the Earth's poles [84, 85].

An interesting special case of the stepping stone model is one in which a fit wild-type strain undergoes deleterious mutations that confer a fixed selective disadvantage. If the deleterious mutations are irreversible, there will be a resultant selection-mutation balance, as shown in Fig. 3.1.1. The irreversible, deleterious mutations describe the genetic dynamics of a well adapted population evolving in a fitness landscape with a single sharp peak (see Appendix B.1 for more details). The population sector dynamics exhibits a phase transition that falls into the directed percolation universality class [95]. The order parameter is the fixation probability of the advantageous strain. In the "active" phase, the advantageous strain has a finite probability of surviving after an infinite amount of time (in a spatially infinite population). In the "inactive" phase, the strain always dies out eventually. The transition into the inactive phase is sometimes called "mutational meltdown" in population genetics literature [9, 96]. We consider the simplest possible model leading to directed percolation: a single active strain with deleterious mutations. However, directed percolation behavior is observed in a wide variety of biological processes [80, 81, 97].

In this chapter we study *curved* population fronts in two dimensions that expand with some constant velocity v, motivated by populations that settle new territory from some localized, approximately circular, initial homeland with radius R_0. (A semicircular variant of this theory would describe organisms spreading into the interior after a coastal colonization event.) A particularly simple example is the growth of a circular bacterial inoculation on a Petri dish [98]. We will again assume that the dynamics are confined to the population front. A dimensional reduction to a $1+1$-dimensional model will also be possible in this case, but the spatial dimension must now *inflate* to account for the growth of the curved front. To simplify the analysis, we assume the extreme limit of a stepping stone model, with a single cell per deme. As a result, the genetic drift, or number fluctuations, at the frontier will be very strong. We will argue that the long-time, large-size dynamics of this extreme limit of the stepping stone model is equivalent to the dynamics of a generalized voter model under an appropriate coarse-graining.

In general, the curved front suppresses the effects of genetic drift by systematically increasing the perimeter of the frontier and inflating sector boundaries away from each other. The suppression becomes significant after a crossover time $t^* = R_0/v$. A scaling argument for this crossover time follows from a comparison of the diffusive and inflationary length scales in the problem. At short times, sector boundaries move diffusively and sectors grow to a characteristic diffusive size $\ell_d \sim R_0\phi \sim \sqrt{Dt}$, where ϕ is the angular sector size, R_0 is the homeland radius, and D is a diffusion constant. The diffusion constant scales as $D \sim a^2/\tau_g$ where a is the cell diameter, and τ_g is the time between generations (see Sec. 3.3 for more details). The inflationary growth due to the increasing population radius $R(t) = R_0 + vt$ has a characteristic size $\ell_i \sim \phi vt$. So, there will be a crossover time t^* at which the two length scales ℓ_i and ℓ_d are comparable. This sets $\ell_d \sim R_0\phi \sim \sqrt{Dt^*} \sim \phi vt^* \sim \ell_i$ and we find the desired result $t^* \sim R_0/v$. After time t^*, sector boundaries with a significant angular separation no longer diffusively interact, and the population separates into independent segments. Consequently, quantities such as the survival probability and the steady state fraction of an advantageous mutant become enhanced relative to linear expansions, because the advantageous mutant just has to survive until time t^* after which inflation prevents its extinction via genetic drift.

The inflationary dynamics also strongly influences the mutation-selection balance (directed percolation phase transition) by introducing a cutoff to the critical population dynamics after time t^*. Just as in a finite size effect at a conventional phase transition, this cutoff is characterized by scaling functions. We find that the dynamics scales with t^* in a similar way as for a finite population size L in a linear range expansion at criticality, but with a completely different scaling function. Moreover, inflation is qualitatively different from typical finite size effects due to transient dynamics or small system sizes [95, 99]: Inflation occurs at long times and is due to a gradually *increasing* system size.

We will also consider the effects of inflation on the dynamics of single sectors of cells with a selective advantage. A key dimensionless parameter that characterizes the strength of the selective advantage for voter-type models (in the presence of inflation) is $\kappa \approx s\sqrt{R_0/2v\Delta t}$, where $s \leq 1$ is a selection coefficient, Δt is the time between generations, and v is the circular front propagation speed. For $\kappa \ll 1$, selection is weak, genetic drift dominates the evolution, and the sector boundaries exhibit strong fluctuations. Conversely, when $\kappa \gg 1$, the selective advantage is strong, the sector boundaries move out deterministically at long times, and the fate of the sector is decided very early in the evolution. Examples of both cases are shown in Fig. 3.3.4.

In general, if a subpopulation has a strong selective advantage, it will create "bulges" in an advancing population front. The population front will also roughen due to the stochastic nature of the growth at the interface [13, 92]. These bulges and undulations will couple the population dynamics to the front geometry, changing the nature of the dynamics. As discussed above, these effects are not included in our study. However, we note that a smooth front approximation has been successfully employed in studies of expansions of bacteria, such as *Pseudomonas aeruginosa* [78]. In addition, this simplification will allow us to make connections between the radial expansions studied in this chapter and the large body of work on stepping stone models for linear expansions with uniform fronts [13].

The organization of the chapter is as follows: In Sec. 3.2 we develop a generalized

Domany-Kinzel model [100] to simulate both linear and radial range expansions. An unfortunate artifact in lattice simulations of circular range expansions is genetic boundaries and colony shapes that lock into the discrete four- or six-fold symmetry of the underlying lattice. We describe novel simulation methods (which create an amorphous packing of organisms) that resolve this problem. In Sec. 3.3 we interpret the simulation results by appealing to a generalized voter model [3, 101] description that includes both mutations and natural selection. We also discuss how our generalized voter model is connected to the stepping stone model and the directed percolation universality class. The effect of inflation on the directed percolation phase transition is also described in Sec. 3.3. In addition, we compare linear and radial expansion results for quantities such as fixation probabilities and the heterozygosity correlation function. Sec. 3.4 presents a brief discussion of an alternative "deflating" inward expansion in which the population front gets smaller, as might happen in colonization from the perimeter of an island. We conclude with final observations and a summary in Sec. 3.5.

3.2 Simulations

Many properties of the directed percolation phase transition line shown in Fig. 3.1.1 have been studied using the Domany-Kinzel (DK) lattice model [95, 100, 102]. The DK model is implemented on a hexagonal (i.e. equilateral triangle) lattice in which horizontal lines of sites are updated one at a time, starting from the top. The sites can be in one of two states (e.g., red or green) and the states of newly evolved sites are determined by the states of neighboring sites in the horizontal line above, as shown in Fig. 3.2.1(a). Since all of the updating occurs at the bottommost actively evolving horizontal line, we have the dimensional reduction discussed in the introduction. The downward direction on the lattice can be treated as a time axis.

The DK model has a natural biological interpretation. The sites represent cells, and the evolving lines of sites represent cell generations evolving as flat population fronts. Note that the DK model updates the entire front or generation of cells before moving the front. This scheme is called parallel updating [95, 103] and has non-overlapping generations. The two states of the cells in this model correspond to two genetic variants or "alleles." During each update step, new "daughter" cells are evolved according to a probabilistic rule that depends on the states of neighboring "parent" cells in the previous generation. Although these parents do not sexually reproduce, two neighboring asexual cells must compete for space for daughter cells during reproduction. Thus, not all of the parent cells will be able to propagate an associated daughter cell because of limited space at the frontier. As part of this reproductive process, successful daughters are allowed a small chance of switching colors via mutation, as discussed below. After a daughter cell is pushed forward into a frontier populated by a new generation, cell division behind the advancing frontier ceases.

3.2.1 A Generalized Domany-Kinzel Model

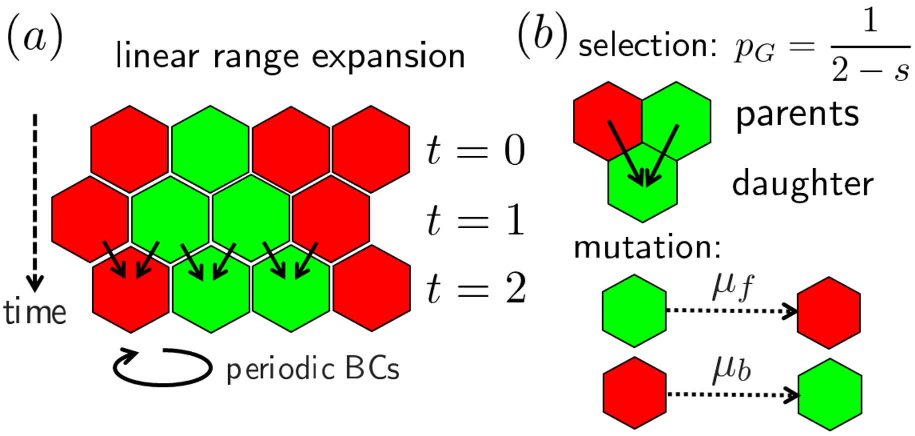


Figure 3.2.1: (a) A sketch of the Domany-Kinzel (DK) model implemented on a hexagonal lattice so that each daughter cell has two potential parents. Periodic boundary conditions are employed, and each cell is evolved using input from the states of the two cells above it in the previous generation. (b) Illustration of the two update steps in our generalized DK model. The daughter shares the color of two parents of the same color. Otherwise, the daughter is green (light gray) with probability $p_G = (2-s)^{-1} \approx 1/2 + s/4$ if its parents are different colors, where $s \leq 1$ is the selective advantage of the green strain. We set $p_R = 1 - p_G = \frac{1-s}{2-s} \approx 1/2 - s/4$. A newly evolved cell is then allowed to mutate from green (light gray) to red (dark gray) with a forward mutation probability $\mu_f \ll 1$ or vice-versa with a backward mutation probability $\mu_b \ll 1$.

In the original version of the DK model, there are two parameters p_1 and p_2 which set the update rules [100]. If a daughter cell has a green and red parent in the previous generation, the daughter is green with probability p_1. If both parents are green, the daughter is green with probability p_2. When the parents are both red, the daughter cell is always red so that no mutations from red to green cells are allowed. This parameterization is not well suited for population genetics, so a biologically motivated modification that includes two-way mutations is necessary. To adapt and extend the DK model, daughter cells are now created in two steps shown in Fig 3.2.1(b). In the first step, a green or red daughter cell is created with a probability p_G or $1 - p_G$, respectively.

We construct p_G by drawing inspiration from the Moran model, a well studied update scheme for well mixed populations [10, 90]. The probability p_G will be proportional to both the number of potential green parents n_G, and the green cell growth rate, which is normalized to unity. The red cells suffer a selective disadvantage and grow at the smaller rate $1-s$. Then, to ensure that p_G is properly normalized, we have

$$p_G = \frac{n_G}{n_G + n_R(1-s)} = \frac{n_G}{s n_G + n_T(1-s)}, \tag{3.1}$$

where n_R is the number of red parents and $n_T \equiv n_R + n_G$. After this first selection step, the surviving daughter cell mutates forward with probability $\mu_f \ll 1$ if it is green and backward with probability $\mu_b \ll 1$ if it is red. The update steps are shown for a simple linear expansion in Fig. 3.2.1(b), for which $n_T = 2$.

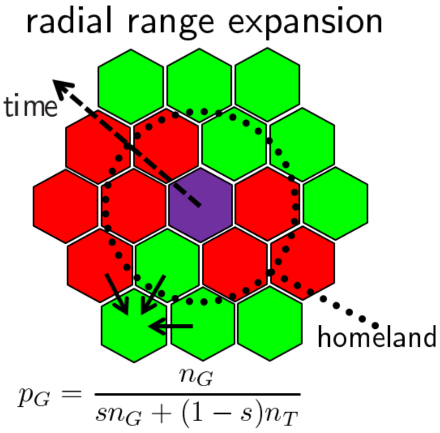


Figure 3.2.2: (Color online) An illustration of a radial expansion simulation on a hexagonal lattice with an initial state given by a “homeland” of 7 cells. We add daughter cells one at a time, each time picking an unoccupied lattice site that is closest to the center of the homeland (the central purple cell). When two or more cells are at the same distance, we chose one of them at random. We implement the generalized Domany-Kinzel update rule that allows for more than two potential parent cells. Each daughter cell will be green (light gray) with a probability p_G that will depend on the number of green (light gray) parents n_G, given by Eq. (3.1). Mutations can be implemented in a second update step as described in the text (also see Fig. 3.2.1).

A lattice model for radial range expansions can now be constructed by taking advantage of the variability of n_T in the update rule in Eq. (3.1). For example, as illustrated in Fig. 3.2.2, we can use a hexagonal lattice and evolve the cells starting from an initial homeland centered around a reference cell. We always choose to evolve the daughter cell that is closest to the central reference cell in the initial seed. This choice ensures a uniform, circular population front. Certain daughters will now have three ($n_T = 3$) instead of two ($n_T = 2$) parents. Our revised DK update scheme easily adapts to this variation. As discussed below, this algorithm leads to circular colonies, and might arise biologically if previously deposited cells generate a chemical that stimulates cell division at the frontier.

3.2.2 The Bennett Model

Unfortunately, the *radial* DK model implemented on lattices with discrete rotational symmetries (e.g. hexagonal, square, or triangular lattices) exhibits strong lattice artifacts. For example, domains of cells prefer to grow along the crystallographic directions, which can introduce artificial oscillations in angular correlation functions, as discussed in Sec. 3.3.2. To eliminate this problem, we have implemented the DK model on an amorphous lattice that does not have any special directions. An efficient way to generate such a lattice is provided by an extension of the Bennet algorithm [69].

In the original Bennet model [69], identical hard sphere particles (cells, in our case) were deposited sequentially in locations closest to a chosen reference point that becomes the center of the population. The possible locations for daughter cells are generated by enumerating all the possible ways of placing a new cell so that it touches at least two cells without overlapping any other cell. The first few close-packed cells (a triangle of three, in the planar implementation discussed below) are placed manually in an initial “seed”. It turns out that, in two dimensions, the generated cluster of identical cells has a very strong hexagonal ordering, regardless of the shape of the initial seed [104]. Following Ref. [104], we create an amorphous configuration by depositing cells with two different sizes with different probabilities. For a carefully chosen range of ratios ρ between the smaller and larger cell radius and the probability Q of depositing the smaller cell rather than the larger one, one finds packings with no preferred directions.

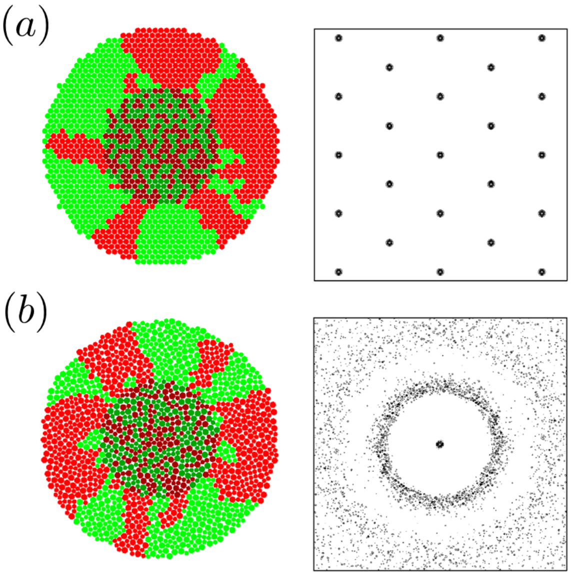


Figure 3.2.3: (*a*) A radial DK model simulation on a hexagonal lattice. The structure factor of the lattice is shown in the right panel. The cell colors were assigned with neutral selective advantage ($s = 0$) and no mutations ($\mu_f = 0$). The first 1500 cells generated by the algorithm are shown. The central shaded region is the homeland of radius $R_0 \approx 10$ cell diameters, where green (light gray) and red (dark gray) cells are placed randomly with equal probability. (*b*) The same simulation performed on a Bennett model lattice with approximately equal numbers of small and large cells ($Q = 0.5$) and a radius ratio $\rho = 0.769$. The structure factor in the right panel shows that the Bennett call packing is isotropic.

One test of rotational symmetry for a generated cluster is to compute the structure factor

$$S(\mathbf{k}) \equiv \frac{1}{N}\left|\sum_{i=1}^{N} e^{i\mathbf{k}\cdot\mathbf{r}_i}\right|^2, \tag{3.2}$$

where N is the number of cells in the cluster and $\mathbf{r}_i$ is the center of the i-th cell. A hexagonal cluster of cells with identical radii will have a striking structure factor with a clear six-fold symmetry, as shown in Fig. 3.2.3(a). By contrast, the amorphous cluster in Fig. 3.2.3(b) has a structure factor with no discrete symmetries. The latter cluster was generated using equal numbers of large and small cells ($Q = 0.5$) and size ratio $\rho = 0.769$. We use this choice of parameters in the radial range expansion simulations in the next two sections (Sec. 3.3.1 and Sec. 3.3.2). In the later sections, we use slightly different parameters ($Q = 0.6$ and $\rho = 8/11 \approx 0.73$) which yield similar cell configurations.

Once an amorphous configuration of cells is generated, a homeland of cells is established by assigning genotypes of red or green to all cells within a radius R_0 of the central reference point. Fig. 3.2.3(b) illustrates a well mixed initial homeland with R_0 equal to about ten cell diameters. This initial homeland is then evolved using the rules discussed in Sec. 3.2.1. Similar to the evolution on the hexagonal lattice in Fig. 3.2.2, a new daughter is always the one that is closest to the central reference point used to generate the lattice. This can be regarded as an approximation for cell colonies where previous generations continuously secrete a chemical promoting cell division into the open space at the frontier. (We assume steric constraints prevent cell division in the interior of the colony.) The resulting concentration of stimulant will monotonically from the center, leading to circular colonies. Alternatively, a circular population front can be imposed externally by, for example, a receding waterline around a circular island. The time t is measured in generations. Each generation represents the time needed to evolve a new rim of cells at the circular population front. We assume that this rim has a finite thickness corresponding to the average cell width, so that the population radius is given by $R(t) = R_0 + vt$. The speed v is the rim thickness divided by the generation time.

To facilitate the measurement of time in units of the radial colony size, we rescale all of our amorphous lattice coordinates such that we get the same cell number density as in a hexagonal lattice of identical cells with a diameter equal to one. We know that for the hexagonal lattice, a radial expansion will grow out by about a cell diameter per generation. Thus, with this rescaling, we can treat each generation in our amorphous lattice as all the cells in a rim with a thickness equal to one. This means that our population front speed is set to $v = 1$ and the radius of the radial range expansion at time t is now $R(t) = R_0 + t$.

In the Bennett model lattice, there are a variety of ways to choose the competing parents in the local neighborhood of each daughter. In our simulations, we choose the parents to be the previously evolved cells that touch the daughter cell. Since we update the cells one at a time, starting from the closest cell to the reference point, the adjacent parent cells must all be closer to the reference point than the daughter. This usually gives us two or three potential parents. However, sometimes the daughter cell is adjacent to cells that are all further from the reference point and have not yet been updated. In these exceptional cases, we choose the competing parents to be the two closest nonadjacent cells that are closer to the reference point.

Our amorphous cell packings will have slight inhomogeneities. For example, there will be a variation in the coordination number z at each lattice site. The lattice spacing a will also vary across the lattice. Hence, we define an effective coordination number z_r for our radial expansions as well as an effective lattice constant a_r. Since the amorphous lattice is not the densest possible packing of the cells, there will be gaps between the cells, and a_r is expected to be slightly larger than the average cell diameter. In particular, we expect the amorphous lattice to be more loosely packed than a hexagonal lattice of identical cells with diameters equal to the average cell diameter in the amorphous lattice.

The inhomogeneities, or “quenched disorder,” in our amorphous cell packings require an average over spatial variations to get the effective parameters a_r and z_r. We have found it most straightforward to estimate a_r and z_r by fitting simulation results to known analytic solutions. For example, the heterozygosity function derived in Sec. 3.3.1 is used to fit the combination a_r^2/z_r (see Fig. 3.3.2). An alternative approach might be to use the coarse-graining procedure used to produce measured correlation functions from actual cell colonies in Ref. [78]. Some quantities, such as the fraction of green cells $f(t)$ at time t, can be measured without knowing either a_r or z_r. Note that the disorder is present in both the time-like (radial) and the space-like (circumferential) direction on the lattice. Fortunately, it can be shown [95] that spatio-temporally quenched disorder is an irrelevant perturbation in systems in the directed percolation universality class and it will not change the important features of the dynamics.

3.3 Voter Models with Inflation

To understand the radial simulation results analytically, we shift to a generalized voter model description in which cells in the population are randomly updated one at a time, instead of generation by generation. The former and latter update schemes are often referred to as random-sequential and parallel updates, respectively [103]. The original Domany-Kinzel model uses parallel updates as discussed previously. Our simulations (using the DK model generalization discussed in Sec. 3.2) represent an intermediate case since we evolve the system one cell at a time, but in a deterministic way to ensure that each generation is updated sequentially. The random-sequential scheme is used in canonical models such as the voter model or more general contact processes, which have features conducive to an analytical analysis [103]. Parallel updates are easier to implement in simulations. Despite the differences between these two update schemes, we expect that they share universal features when we look at their coarse-grained dynamics. There are exceptional cases where this is not so, such as in the roughening transition in models of polynuclear growth [95]. However, we do not expect our simulations to be exceptional, because the directed percolation transition can be modelled with either update scheme [95]. We start be considering flat frontier range expansions, and then introduce the inflation due to radial range expansions in the coarse-grained description at the end.

To start the analysis, we exploit dimensional reduction and consider the cells at the frontier of a $d+1$-dimensional population as "voters" in a d-dimensional space. It is convenient to describe the states of the cells with a set of spin variables $\{\sigma_i\}_{i\in\mathbb{Z}^d}$ where $\sigma_i = \pm 1$ and i is a d-dimensional vector of integers describing the location of the ith cell on a hypercubic lattice that approximates cell positions in our amorphous simulations. The spin variable encodes the cell state: $\sigma_i = 1$ for a green cell and $\sigma_i = -1$ for a red one. No empty sites are allowed, which is equivalent to assuming a uniform front that advances in lock-step in $(d+1)$-dimensions with each generation. We now will use the notation $\{\sigma\} \equiv \{\sigma_i\}_{i\in\mathbb{Z}^d}$ for convenience. The probability distribution of a particular spin configuration $\{\sigma\}$, $P(\{\sigma\},t)$ obeys the master equation [3]

$$\partial_t P(\{\sigma\},t) = \sum_{j\in\mathbb{Z}^d} \left[\omega_{\{\sigma\}_j\to\{\sigma\}} P(\{\sigma\}_j,t) - \omega_{\{\sigma\}\to\{\sigma\}_j} P(\{\sigma\},t)\right], \tag{3.3}$$

where $\{\sigma\}_j$ is the same configuration of spins as $\{\sigma\}$, *except* that the jth spin is flipped (has an opposite state). Spin flipping in the biological context represents the replacement of a cell at site j by a cell of the opposite type after cell division. The variable j in Eq. (3.3) runs over all spin locations and $\omega_{\{\sigma\}\to\{\sigma\}_j}$ are the probability rates of transitions between the configurations $\{\sigma\}$ and $\{\sigma\}_j$. The flipping rates $\omega_{\{\sigma\}\to\{\sigma\}_j}$ are given by the sum of three terms,

$$\omega_{\{\sigma\}\to\{\sigma\}_j} = \omega^{\text{drift}}_{\{\sigma\}\to\{\sigma\}_j} + \omega^{\text{mut}}_{\{\sigma\}\to\{\sigma\}_j} + \omega^{\text{sel}}_{\{\sigma\}\to\{\sigma\}_j}, \tag{3.4}$$

where $\omega^{\text{drift}}_{\{\sigma\}\to\{\sigma\}_j}$ is the rate due to random genetic drift, $\omega^{\text{mut}}_{\{\sigma\}\to\{\sigma\}_j}$ is the contribution from mutations, and $\omega^{\text{sel}}_{\{\sigma\}\to\{\sigma\}_j}$ is the contribution from a selection bias.

For neutral strains without mutations, only genetic drift contributes to the rates $\omega_{\{\sigma\}\to\{\sigma\}_j}$. This genetic drift arises because a cell is more likely to change its state if it is surrounded by neighbors of the opposite type [15], just as in the original voter model [3]. Thus, $\omega^{\text{drift}}_{\{\sigma\}\to\{\sigma\}_j}$ is the standard Glauber rate [105] used in the voter model:

$$\omega^{\text{drift}}_{\{\sigma\}\to\{\sigma\}_j} = \frac{1}{2\tau_g}\left[1 - \frac{\sigma_j}{z}\sum_{i\ \text{n.n.}\ j}\sigma_i\right], \tag{3.5}$$

where the summation is over the $z = 2d$ nearest neighbors of voter σ_j. The generation time τ_g corresponds to updating every cell once per generation. Mutations allow some cells to stochastically change their states independently of the states of their nearest neighbors. We model these processes by introducing the rates

$$\omega^{\mathrm{mut}}_{\{\sigma\}\to\{\sigma\}_j} = \frac{\mu_f(1+\sigma_j)}{2\tau_g} + \frac{\mu_b(1-\sigma_j)}{2\tau_g}, \tag{3.6}$$

where μ_f and μ_b are the mutation probabilities for the cell during a generation time τ_g. These probabilities correspond to the mutation probabilities in the DK model simulations described in Sec. 3.2. To make further contact with the DK model, green cells are replaced by red cells with a rate that is smaller than the reverse rate by a factor of $1-s$, where $s \leq 1$ represents a selective advantage. The resultant contribution to the flipping rate is

$$\omega^{\mathrm{sel}}_{\{\sigma\}\to\{\sigma\}_j} = \frac{s}{4z\tau_g} \sum_{i\ \mathrm{n.\,n.}\ j} (\sigma_i - 1)\left(1+\sigma_j\right), \tag{3.7}$$

where we sum over the nearest neighbors of spin σ_j.

We may now apply a coarse-graining procedure on the master equation (Eq. (3.3)) with the rates in Eq. (3.4) to find a stochastic differential equation for the local, coarse-grained fraction $f(\mathbf{x},t)$ of green cells:

$$f(\mathbf{x},t) \equiv \frac{1}{\Omega} \sum_{i\in\mathcal{B}_\Omega(\mathbf{x})} \frac{1+\sigma_i(t)}{2}, \tag{3.8}$$

where $\mathcal{B}_\Omega$ is a set of spins in a neighborhood around frontier point $\mathbf{x}$ with volume Ω. In Appendix B.2 we derive that for $\Omega = 1$, $f(\mathbf{x},t)$ obeys the equation

$$\begin{aligned}\partial_t f(\mathbf{x},t) = D\,\nabla^2 f(\mathbf{x},t) + \bar{s}\, f(\mathbf{x},t)[1-f(\mathbf{x},t)] \\ - \bar{\mu}_f f(\mathbf{x},t) + \bar{\mu}_b\,(1-f(\mathbf{x},t)) + \eta(\mathbf{x},t),\end{aligned} \tag{3.9}$$

where $D = a^2/(z\tau_g)$ is the diffusion coefficient, $\bar{s} = s/\tau_g$, $\bar{\mu}_f = \mu_f/\tau_g$, $\bar{\mu}_b = \mu_b/\tau_g$, and η is a noise with correlations given by

$$\langle \eta(\mathbf{x}, t)\eta(\mathbf{x}', t)\rangle \approx \frac{2f(\mathbf{x}, t)[1 - f(\mathbf{x}, t)]}{\tau_g} a^d \delta(t - t')\delta(\mathbf{x} - \mathbf{x}'). \tag{3.10}$$

This Langevin equation can also be derived starting from a stepping stone model where the lattice sites are subpopulations undergoing a Moran process (see Appendix B.1) and which exchange individuals between neighboring sites. By setting the number of individuals in each subpopulation to one, we find the same Langevin equation (Eq. (3.9)), as shown in Ref. [13]. Because of the nonlinearity in the noise correlations Eq. (3.10), the stochastic differential equation must be interpreted according to the Îto calculus [13, 15]. The time and space scaling used in the simulations sets $\tau_g = 1$.

The analysis so far is valid for both radial and linear range expansions as no reference has been made to the shape of the population front. For a linear front of length L in a two-dimensional population, the coordinate $\mathbf{x}$ is replaced by a single coordinate $x \in [-L/2, L/2]$. For a circular front, we have to be more careful. In an “inflating” front, the number N of active “voters” or cells around the perimeter increases in time. To cover a variety of interesting cases, we let the radius $R(t)$ grow in time according to a general power law

$$R(t) = R_0 \left[1 + \left(\frac{t}{t^*}\right)^{\Theta}\right], \tag{3.11}$$

where Θ is the growth exponent $0 < \Theta < \infty$, t^* is a characteristic time, and R_0 is the initial homeland radius. In our simulations, we will be restricted to $\Theta = 1$ where there is a constant front propagation velocity v such that $t^* = R_0/v$. The choice of $\Theta = 1$ is also appropriate for biological systems [90, 98, 106, 107] expanding on a flat substrate.

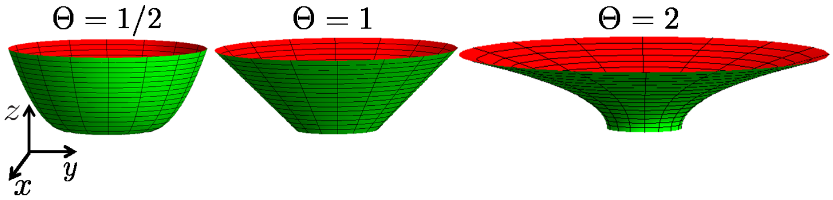


Figure 3.3.1: Three examples of surfaces generated by the set of points $\mathbf{r}(t,\phi)$ given by Eq. (3.12) with $0 \leq t \leq 2$ and $-\pi \leq \phi < \pi$. We set the initial population circumference and the crossover time to one: $R_0 = t^* = 1$. The $\Theta = 1/2$ surface is shaped like a bowl and has positive Gaussian curvature $K > 0$. Conversely, the $\Theta = 2$ surface is shaped like a trumpet and has $K < 0$. Confining a population to grow along the $\Theta = 1$ surface (with $K = 0$) is another way of implementing the regular radial range expansion with linear inflation.

The $\Theta \neq 1$ case might be used to model frontier expansions on curved surfaces, such as a parabolic surface with positive Gaussian curvature, or a negatively curved surface like the horn of a trumpet as shown in Fig. 3.3.1. Each position $\mathbf{r} = (x, y, z)^T$ on such a surface can be parameterized by the time t and an angle $\phi \in [-\pi, \pi)$. The angle ϕ parameterizes a location on an expanding circumference. In Cartesian coordinates,

$$\mathbf{r}(t,\phi) = \begin{pmatrix} R_0 \left[1 + \left(\frac{t}{t^*}\right)^{\Theta}\right] \cos\phi \\ R_0 \left[1 + \left(\frac{t}{t^*}\right)^{\Theta}\right] \sin\phi \\ t \end{pmatrix}. \tag{3.12}$$

The population grows upward along the positive z-axis, so that we can identify this direction with the time t. The Gaussian curvature $K(t,\phi)$ at each point (t,ϕ) of such a surface can be calculated using standard methods [108]:

$$K(t,\phi) = \frac{R_0^2 \left(1 + \left(\frac{t}{t^*}\right)^{\Theta}\right) \left(\frac{t}{t^*}\right)^{\Theta} \Theta(1-\Theta)}{t^2 + R_0^2 \left(\frac{t}{t^*}\right)^{2\Theta} \Theta^2}. \tag{3.13}$$

Note that $K(t,\phi)$ is positive when $\Theta < 1$ and negative when $\Theta > 1$. Surfaces at three different values of Θ are shown in Fig. 3.3.1.

The analysis will be general for now, and we will consider circular population fronts growing with arbitrary power $\Theta > 0$ of t, with radius $R(t)$ given by Eq. (3.11). A location along the population front is given by the x and y components of $\mathbf{r}$ in Eq. (3.12): $R(t)\{\cos\phi, \sin\phi\}$, where $\phi \in (-\pi, \pi)$ is the angle parameterizing a location on an expanding circumference. Setting $x = R(t)\phi$ in Eqs. (3.9) and (3.10), we obtain (upon setting $\mu_b = 0$) the Langevin equation for a radial expansion

$$\partial_t f(\phi, t) = \frac{D}{R(t)^2}\,\partial_\phi^2 f(\phi, t) + s f(\phi, t)[1 - f(\phi, t)] - f(\phi, t)\mu_f + \eta(\phi, t), \tag{3.14}$$

where the noise $\eta(\phi, t)$ obeys

$$\langle \eta(\phi, t)\eta(\phi', t)\rangle = \frac{2a f(\phi, t)[1 - f(\phi, t)]}{R(t)}\,\delta(t - t')\delta(\phi - \phi'), \tag{3.15}$$

and we have set the generation time to one: $\tau_g = 1$. When we compare our simulation results to these voter models for radial range expansions, we will use an effective lattice constant a_r and coordination number z_r, as discussed in Sec. 3.2. The corresponding diffusion constant is then $D = D_r \approx a_r^2/(z_r \tau_g)$, with $\tau_g = 1$. Similar equations to Eq. (3.14) and Eq. (3.15) were derived via the stepping stone model in Ref. [13, 78] for the linearly inflating case, where $R(t) = R_0 + vt$.

3.3.1 Neutral Evolution

An important exactly soluble case is a *neutral* radial range expansion, considered previously for linear inflation ($\Theta = 1$ in Eq. (3.11)) and inflation with a rough front

[9, 13, 78, 109]. We consider first the dynamics without mutations for simplicity. An illustration of a range expansion undergoing neutral dynamics for a well mixed initial homeland with a fraction $f_0 = 1/2$ of green (light gray in print) cells is given in the left panels of Fig. 3.2.3 (a) and (b). We see in the figure that the domain walls between green and red (light and dark gray, respectively, in print) sectors will initially annihilate in pairs due to genetic drift leading to domain coarsening. A useful measurement of the coarsening is the angular heterozygosity (for a spatially homogeneous initial condition)

$$H(\phi,t) \equiv \Big\langle f(\phi+\phi',t)[1-f(\phi',t)] + f(\phi',t)[1-f(\phi+\phi',t)] \Big\rangle_{\phi'}, \tag{3.16}$$

where $\langle\ \ \rangle_{\phi'}$ includes an angular average over ϕ' and over realizations of the noise in the Langevin equation (Eq. (3.14)) for $f(\phi,t)$. The heterozygosity $H(\phi,t)$ represents the probability of observing two cells of *different* types an angular distance ϕ apart at time t. Since our voter-type model has a single cell at each lattice point, $H(\phi,t)$ satisfies the boundary condition $H(\phi=0,t)=0$.

We set $s = \mu_f = 0$ in Eq. (3.14) and use the Îto calculus to get the equation obeyed by the angular heterozygosity $H(\phi,t)$ for the initial condition $f(\mathbf{x},t=0) \equiv f_0$. We also assume that we can extend the range of ϕ from the interval $[-\pi,\pi]$ to the entire real line $(-\infty,\infty)$, an approximation will be valid as long as the average angular sector size is small compared to 2π, the angular size of the population [9]. The heterozygosity $H(\phi,t)$ for an arbitrary power law inflation $R(t) = R_0[1+(t/t^*)^{\Theta}]$ obeys a diffusion equation with a time-dependent diffusivity $D(t) = 2D_r/[R(t)]^2$. The time dependence can be removed by introducing a rescaled, “conformal” time variable

$$t_c(\Theta, t) \equiv \frac{t}{\Theta}\left\{\frac{1}{1+(t/t^*)^\Theta} + (\Theta-1)\,{}_2F_1\left[1, \frac{1}{\Theta}, 1+\frac{1}{\Theta}; -\left(\frac{t}{t^*}\right)^\Theta\right]\right\}, \tag{3.17}$$

where ${}_2F_1(\alpha, \beta, \gamma; z)$ is the Gaussian hypergeometric function [4]. Changing variables to the conformal time yields a simple diffusion equation for $H(\phi, t_c)$, namely

$$\partial_{t_c} H(\phi, t_c) = 2D_r\, \partial_\phi^2 H(\phi, t_c). \tag{3.18}$$

Eq. (3.18) can now be solved for the uniform initial condition $H(\phi, t_c = 0) = H_0 \equiv 2f_0(1-f_0)$ and the boundary condition $H(\phi = 0, t_c) = 0$. The solution, after changing back to the original time variable, is

$$H(\phi, t) = H_0 \,\mathrm{erf}\left[\frac{R_0|\phi|}{\sqrt{8D_r t_c(\Theta, t)}}\right], \tag{3.19}$$

where $t_c(\Theta, t)$ is given by Eq. (3.17).

We recover the known solution for a $\Theta = 1$ linear expansion by replacing $t_c(\Theta, t)$ with t in Eq. (3.19) [13]. The limiting value $t_c(\Theta, t \to \infty)$ is particularly interesting as it tells us for which values of Θ we can expect a “cutoff” in the dynamics, i.e., when the time variable $t_c(\Theta, t)$ approaches a finite value and $H(\phi, t)$ approaches a limiting shape as $t \to \infty$ in the radial expansion. In particular, we find

$$t_c(\Theta, t \to \infty) = \begin{cases} \infty & \Theta \leq \frac{1}{2} \\ \dfrac{\pi(\Theta-1)t^*}{\Theta^2 \sin(\pi/\Theta)} & \Theta > \frac{1}{2} \end{cases}. \tag{3.20}$$

Thus, when the exponent Θ is larger than the characteristic value 1/2 (which matches the diffusive wandering behavior of domain walls), the inflation will eventually dominate the diffusive dynamics.

The inflation-induced crossover also modifies the average angular sector width $\Delta\phi(t)$. The width follows from Eq. (3.19), upon noting that (see Eq. (C2) in Ref. [13])

$$\Delta\phi(t) = \left(\frac{\partial H}{\partial \phi}\right)^{-1}\Bigg|_{\phi\to 0^+} = \frac{\sqrt{2\pi D_r t_c(t,\Theta)}}{R_0 H_0}. \tag{3.21}$$

Eq. (3.21) will be valid only when $\Delta\phi(t) \ll 2\pi$ since we do not take into account sectors that take over the entire population. When $\Theta \leq 1/2$, the angular width $\Delta\phi(t)$ diverges at long times so that only a *single* genetic variant is present. Note the striking difference between this behavior and the $\Theta > 1/2$ case where the number of surviving sectors approaches a finite value given by

$$N_{\text{surv}} = \frac{2\pi}{\lim_{t\to\infty}\Delta\phi(t)} = R_0 H_0 \sqrt{\frac{2\Theta^2 \sin(\pi/\Theta)}{D_r(\Theta-1)t^*}}. \tag{3.22}$$

This result is in agreement with the results of Refs. [9, 13] for the case $\Theta = 1$ and $t^* = R_0/v$.

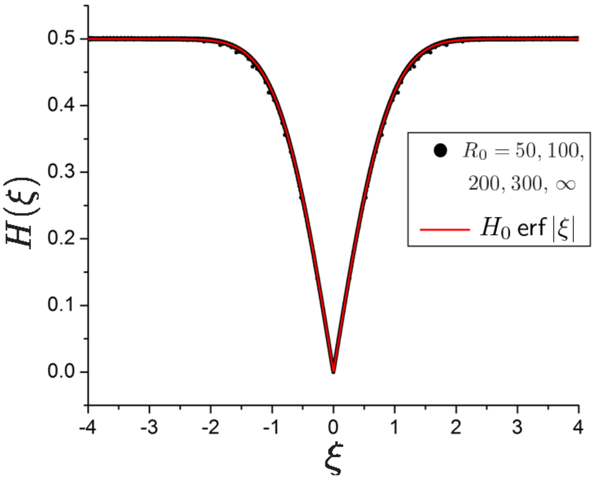


Figure 3.3.2: Nearly perfect collapse of the simulation data using the scaled variable $\xi = R_0\phi/\sqrt{8D_r t_c(t)}$ for the heterozygosity with $H_0 = 1/2$, corresponding to random initial conditions, with equal probabilities for the two different colors. The initial colony radius is given by $R_0 = 50, 100, 200, 300$. Linear range expansions correspond to the limit $R_0 \to \infty$. All of the radial expansion simulations were fit with a single parameter D_r. The scaling variable is $\xi = x/\sqrt{8Dt}$ for linear expansions.

We now specialize to $\Theta = 1$ and $t^* = R_0/v$ in order to check our results with simulations and known results. In this case, we have the simple relation

$$t_c(t) \equiv t_c(1,t) = \frac{t}{1+t/t^*} = \frac{t}{1+vt/R_0}. \tag{3.23}$$

The transformation from t to $t_c(t)$ in Eq. (3.23) is a special case of the Mobiüs transformation and maps the interval $t \in (0,\infty)$ to $t_c \in (0,t^*)$ [110]. Both the linear and radial expansion heterozygosities collapse to a single universal function $H(\xi) = H_0\,\mathrm{erf}\,|\xi|$, with $\xi = x/\sqrt{8Dt}$ for linear expansions, and $\xi = R_0\phi/\sqrt{8D_r t_c(t)}$ for radial ones. The comparison requires fitting the radial diffusion constant $D_r = a_r^2/(z_r\tau_g)$, which depends on the effective radial lattice constant a_r and the coordination number z_r for the Bennett model cellular packing ($\tau_g = 1$ in the simulations). We find excellent data collapse and agreement between simulation and the voter model theory. These theoretical predictions for the heterozygosity are described in more detail in Ref. [13] and are compared to simulation data in Fig. 3.3.2. The connection between the radial and linear range expansions is also discussed in Ref. [109] for expansions with rough fronts.

3.3.2 Heterozygosity in Radial Range Expansions

It is also possible to derive analytic solutions for the heterozygosity in the biologically relevant limit of neutral mutations. Neutral mutations are common and can have an important effect on a population's genetic composition [11]. We now compare strictly neutral mutational dynamics with genetic drift in radial versus linear range expansions. Many previous results for the linear case are reviewed in Ref. [13].

Mutations spoil the clean connection between linear and radial range expansions via a conformal time coordinate (Eq. (3.23)). Instead, mutations (we now assume both μ_f and μ_b are nonzero) dominate the dynamics at long times in both types of range expansion. The equation of motion for $H(\phi, t)$ can be derived from Eq. (3.14) via the Îto calculus, just as in the neutral case. An additional term $\mu_b f(\mathbf{x}, t)$ is added to the left-hand side of Eq. (3.14) to take into account backward mutations from red (dark gray) to green (light gray) cells at rate μ_b. The equation of motion for a radial range expansion for a homogeneous initial cell density $f(\mathbf{x}, t=0) = f_0$ is

$$\partial_t H(\phi, t) = \frac{2D}{[R(t)]^2}\partial_\phi^2 H(\phi, t) - 2(\mu_f + \mu_b)H(\phi, t) + 2(\mu_f - \mu_b)f(t) + 2\mu_b, \tag{3.24}$$

where the third and fourth terms on the right-hand side of Eq. 3.24 come from the coupling between the equations for the first and second moments of $f(\mathbf{x}, t)$. These terms include a contribution from the average fraction of green cells $f(t) \equiv \langle f(\mathbf{x},t)\rangle_{\mathbf{x}}$, given by

$$f(t) = \frac{\mu_b + [f_0(\mu_f+\mu_b) - \mu_b]e^{-t(\mu_f+\mu_b)}}{\mu_f+\mu_b}, \tag{3.25}$$

(see Ref. [13]) which decays to a steady state value $f(t \to \infty) = \mu_b/(\mu_f+\mu_b)$. To get the analogous linear equation, one can replace the angular coordinate ϕ with a spatial coordinate x along the linear front and set $R(t) \to 1$.

The solution to Eq. (3.24) for an initially homogeneous random initial condition $H(\phi, t=0) = H_0 = 2f_0(1-f_0)$ is

$$H(\phi,t) = H_\infty[(R_0+vt)\phi] + e^{-(\mu_f+\mu_b)t} H_{\mathrm{tr}}(\phi,t), \tag{3.26}$$

which at long times approaches the steady state $H_\infty[(R_0+vt)\phi]$, where

$$H_\infty(x) = \frac{2\mu_f\mu_b}{(\mu_f+\mu_b)^2}\left[1 - e^{-|x|\sqrt{\frac{\mu_f+\mu_b}{D}}}\right]. \tag{3.27}$$

The transient term $H_{\mathrm{tr}}(\phi,t)$ is given by

$$\begin{aligned} H_{\mathrm{tr}}(\phi,t) &= H_0 e^{-t'}\,\mathrm{erf}\left[\frac{\xi_r}{\sqrt{1+t/t^*}}\right] \\ &+ \frac{2t(\mu_f-\mu_b)[f_0\mu_f - (1-f_0)\mu_b]e^{-t'}}{\mu_f+\mu_b}\int_0^1 \mathrm{erf}\left[\frac{\sqrt{t^*/t+y}\,\xi_r}{\sqrt{(1+t^*/t)(1-y)}}\right] e^{t'y}\mathrm{d}y \\ &+ \frac{4t\mu_f\mu_b e^{-t'}}{\mu_f+\mu_b}\int_0^1 \mathrm{erf}\left[\frac{\sqrt{t^*/t+y}\,\xi_r}{\sqrt{(1+t^*/t)(1-y)}}\right] e^{2t'y}\,\mathrm{d}y + \frac{2e^{t'}\mu_f\mu_b}{(\mu_f+\mu_b)^2}\left[e^{-2\sqrt{2t'}\,\xi_r} - 1\right], \end{aligned} \tag{3.28}$$

where we have defined a dimensionless time $t' \equiv (\mu_f+\mu_b)t$, and a dimensionless length scale

$$\xi_r = \frac{(R_0 + vt)\phi}{\sqrt{8D_r t}}. \tag{3.29}$$

The steady-state shape $H_\infty[(R_0 + vt)\phi]$ is exactly the same for linear range expansions with mutations (see Ref. [13]) if we identify $(R_0 + vt)\phi = x$ as the separation between two cells in the linear expansion. In both expansions, mutations set the average domain size at long times. Also, a careful asymptotic analysis of the integrals in Eq. (3.28) reveals that the transient term in Eq. (3.26) decays as $e^{-(\mu_f+\mu_b)t}$. A similar analysis shows that Eq. (3.28) reduces to the correct transient term for a linear range expansion in the limit $t^* \to \infty$. More details about the linear range expansion heterozygosity are discussed in Ref. [13].

Equation (3.28) reveals two competing time scales: the mutational relaxation time $(\mu_f + \mu_b)^{-1}$ and the inflation time $t^* = R_0/v$. Mutations equilibrate on the first time scale, while the second determines when inflation sets in. The effect of inflation is observable for intermediate time scales such that $t^* \ll t \ll (\mu_f + \mu_b)^{-1}$. The average angular domain size $\Delta\phi(t) = [\partial_\phi H(\phi, t)|_{\phi=0}]^{-1}$ in this intermediate time regime is given by

$$\begin{aligned}\Delta\phi(t) = \frac{\sqrt{2\pi D t_c(t)}}{H_0 R_0}\Bigg\{ 1 - 2t' \\ + \left[\frac{\pi}{H_0}\left[f_0 + \frac{\mu_b(2f_0 - 1)}{\mu_b + \mu_f}\right]\sqrt{\frac{t}{t^*}} + \mathcal{O}\left(\sqrt{\frac{t^*}{t}}\right)\right] t' + \mathcal{O}(t'^2)\Bigg\}^{-1},\end{aligned} \tag{3.30}$$

where $t' = (\mu_f + \mu_b)t \ll 1$ and $t_c(t)$ is the conformal time coordinate (Eq. (3.23)). The factor in front of the curly braces in Eq. (3.30) is the neutral result of Eq. (3.21) with $\Theta = 1$.

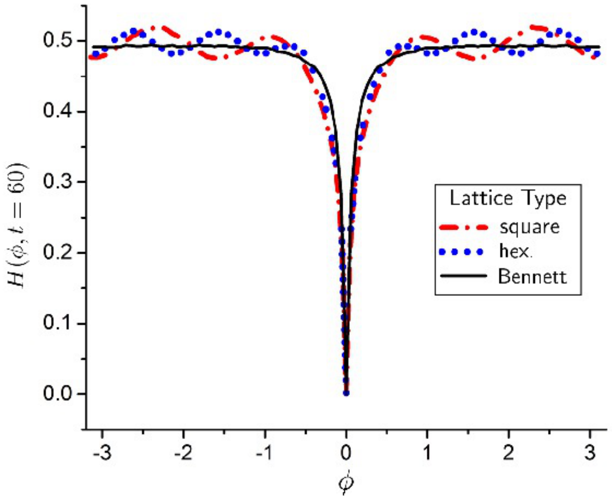


Figure 3.3.3: A calculation of the heterozygosity $H(\phi, t)$ at time $t = 60 \gg t^*$ on different lattices. $R_0 = 15$, $v = 1$, $\mu_f = 0.1$, and $s = 0.57$ in the simulations. The length and time units are measured using the rescaling discussed in Sec. 3.2. Oscillations in the heterozygosity on regular lattices reflect the $4-$ and $6-$fold rotational symmetries. These oscillations are absent for the amorphous Bennett model lattice with $\rho = 0.769$ and $Q = 0.59$.

The heterozygosity correlations are sensitive to dynamics of our model on different lattices. When we are in the active phase of the dynamics with selection and one-way mutations ($\mu_b = 0$) on a regular lattice, the heterozygosity function exhibits oscillations as shown in Fig. 3.3.3. The number of oscillations reflects the rotational symmetry of the lattice, as shown for the hexagonal (6 oscillations) and square (4 oscillations) lattices in Fig. 3.3.3. The oscillations arise because the green cell domains grow preferentially along the crystallographic directions of the lattice. These lattice artifacts obscure the evolutionary dynamics. As discussed in Sec. 3.2, the problem is corrected by using an amorphous lattice which smooths out the oscillations.

3.3.3 Single Seed Dynamics and Selective Advantage

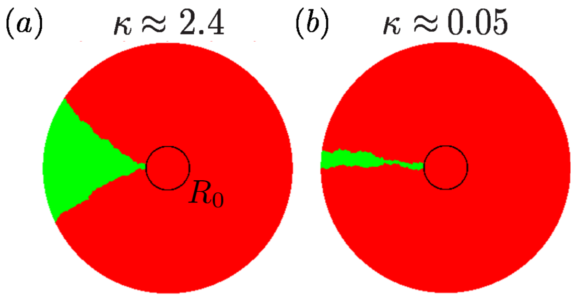


Figure 3.3.4: Two examples of single sector dynamics with $R_0 = 50$ (in units of the average cell diameter) evolved over about 250 generations for two very different values of the key inflationary parameter $\kappa \approx s\sqrt{R_0/2v\tau_g}$ discussed in the text. In (a) a combination of selection bias and inflation dominates genetic drift after a short initial transient, with $\kappa \approx 2.4$. The initial sector width in radians at the edge of the homeland is $\phi_0 \approx 0.1$. In (b) $\kappa \approx 0.05 \ll 1$, so genetic drift can influence the result expected from selection bias and inflation alone for an extended period. We set $\phi_0 \approx 0.2$ in this case. We see that the fate of the sector in (a) is decided much faster than the fate of the sector in (b), which still might eventually die out.

Finding exact solutions is difficult when we include selection. Unlike the cases considered in the two previous sections, it is not possible to apply the Îto calculus to Eq. (3.14) to exactly calculate the angular heterozygosity $H(\phi, t)$. The nonlinear term $sf(\mathbf{x}, t)[1 - f(\mathbf{x}, t)]$ in Eq. (3.14) leads to an infinite set of equations involving ever higher order moments of the field $f(\mathbf{x}, t)$. Although various approximations can be tried, accurate analytic results are limited [13]. Hence, to study the effects of selection on radial range expansions we instead consider the fate of a "seed," i.e. a single initial green sector of size $\phi_0 \ll 1$ with selective advantage s. We assume that mutations are rare, so that $\mu_f = \mu_b \approx 0$. During each generation (corresponding to a time step τ_g), the size ϕ only changes due to the competition between red and green cells at the two boundaries of the sector. We make the further simplifying assumption, easily imposed in our simulations, of neglecting the bulge expected for a seed representing a favorable mutation at the frontier [9, 92]. This approximation should be adequate for $0 < s \ll 1$, and it allows analytic progress for nearly circular colonies.

Each generation in the radial expansion simulations consists of a rim of cells with an approximately single cell thickness (see Sec. 3.2). Hence, the green sector size changes each generation due to a single green cell competing with a neighboring red cell at each of two boundaries. Specifically, each boundary can move by some average amount $\pm\Delta\phi \approx \pm\tilde{a}_r/R(t)$ per generation, where $\tilde{a}_r$ is an effective spacing between adjacent cells along the circular population front approximately equal to the effective lattice spacing a_r (see Appendix B.3). The probability distribution for the angular size ϕ of a sector satisfies the Fokker-Planck-like equation (also known as a Kolmogorov forward equation – see Ref. [13])

$$\partial_\tau p_s(\phi,\tau) = -\frac{w}{1-\tau}\,\frac{\partial p_s}{\partial\phi} + \Delta\,\frac{\partial^2 p_s}{\partial\phi^2}, \tag{3.31}$$

where $\tau = t_c(t)/t^* = vt/(R_0+vt) = vt/R(t)$ is a dimensionless conformal time coordinate. The term in Eq. (3.31) proportional to w represents a bias due to selection, and the term proportional to Δ represents genetic drift. As shown in Appendix B.3, w is proportional to the selective advantages enjoyed by the green cells,

$$w \approx \gamma\,\frac{\tilde{a}_r s}{v\tau_g} \qquad (\text{for } s \ll 1), \tag{3.32}$$

and

$$\Delta \approx \frac{\tilde{a}_r^2}{R_0 v\tau_g} \qquad (s \ll 1), \tag{3.33}$$

where γ is a lattice-dependent factor of order unity. For our disordered Bennett model lattice, we find $\tilde{a}_r \approx 0.90$ (in units of the average cell diameter) and $\gamma \approx 1.2$. We impose the absorbing boundary condition $p_s(\phi = 0, \tau) = 0$, just as in the linear case [9]. Note from Eq. (3.31) that inflation suppresses both drift and diffusion by folding the entire time evolution of the range expansion into the conformal time τ. Inflation suppresses the diffusion more than the drift term since they scale with $R(t)^{-2}$ and $R(t)^{-1}$, respectively.

The Langevin equation corresponding to Eq. (3.31) is

$$\frac{d\phi}{d\tau} = \frac{w}{1-\tau} + \sqrt{2\Delta}\,\eta(t), \tag{3.34}$$

where $\eta(t)$ is a Gaussian white noise (with $\langle \eta(t)\eta(t') \rangle = \delta(t - t')$). Note that unlike the multiplicative noise in the Langevin equation for the coarse-grained cell density $f(\mathbf{x}, t)$ (Eq. (3.9)), Eq. (3.34) has a simple noise term. Hence, single sector dynamics is more easily analyzed using the "dual" description of sector boundary motion, instead of the full density $f(\mathbf{x}, t)$ dynamics in Eq. (3.9) (also see Appendix B.3). When $R_0 \gg \tilde{a}_r$, we might hope to neglect genetic drift relative to the bias term. Upon setting $\Delta = 0$, Eq. (3.34) is solved by

$$\begin{aligned} \phi(\tau) &= \phi_0 - w \ln(1-\tau) \\ &= \phi_0 + w \ln\left[\frac{R(\tau)}{R_0}\right], \end{aligned} \tag{3.35}$$

$$S_\infty \approx \operatorname{erf}\left[\frac{\phi_0}{\sqrt{4\Delta}}\right] + \left[1 - \frac{\ln\left(\frac{w\phi_0}{2\Delta}\right)}{\sqrt{4\pi\Delta}}\,\phi_0\right]\frac{\phi_0 w}{2\Delta} \tag{3.36}$$

similar to the logarithmic spiral sector boundaries found in Ref. [92].

$$K(t,\phi) = \frac{t^2\left(\frac{t}{t^*}\right)^{\Theta}\Theta(1-\Theta)}{\left[1+\left(\frac{t}{t^*}\right)^{\Theta}\right]\left[t^2+R_0^2\left(\frac{t}{t^*}\right)^{2\Theta}\Theta^2\right]^2} \tag{3.37}$$

To obtain the conformal time to fixation τ_f, we set $\phi(\tau_f) = 2\pi$ to find

$$\tau_f = 1 - \exp\left[-\frac{2\pi - \phi_0}{w}\right] \tag{3.38}$$

for the purely deterministic model. When noise is included in our radial voter model, we find from our simulations that $\tau_f \approx 1$ for all $s \lesssim 0.5$ and $\phi_0 \lesssim \pi$. Thus, a sector with small initial angular width $\phi_0 \ll 1$ can only take over a population in a reasonable time provided $w \propto s$ is large. When ϕ_0 is very small, the takeover predicted by the deterministic solution in Eq. (3.35) is preempted by number fluctuations, i.e., genetic drift. If a small initial sector survives the genetic drift, it can eventually take over the entire population in a deterministic fashion.

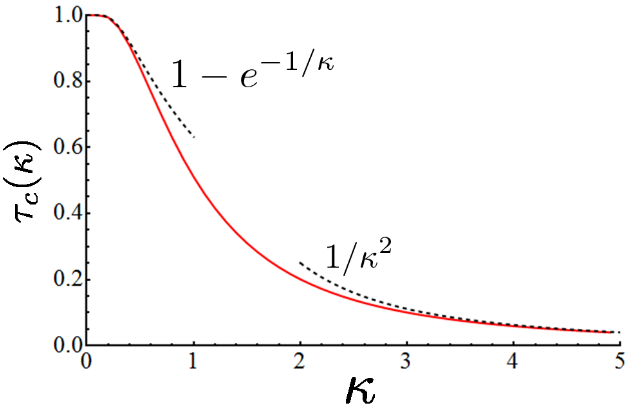


Figure 3.3.5: A plot of the crossover time τ_c that marks when genetic drift is no longer able to keep up with inflation. The time τ_c is given as a dimensionless time coordinate, related to the time t via $\tau = vt/(R_0+vt) = vt/R(t)$. The dashed lines show the asymptotic approximations to τ_c given in Eq. (3.43).

To go beyond the deterministic approximation, we must account for the diffusive dynamics of the sector boundaries caused by genetic drift. To treat the case $\Delta > 0$, we define a new angle variable

$$\psi(\tau) = \phi(\tau) + w\ln(1-\tau), \tag{3.39}$$

such that the sector probability distribution $p_s(\psi, t)$ now obeys

$$\partial_\tau p_s(\psi,\tau) = \Delta\,\frac{\partial^2 p_s(\psi,\tau)}{\partial\psi^2}. \tag{3.40}$$

Our boundary condition $p_s(\phi = 0, \tau) = 0$ moves in time, $p_s(\psi = w\ln(1-\tau), \tau) = 0$. The diffusive motion of the variable ψ, with uncertainty $\delta\psi = \pm\sqrt{2\Delta\tau}$, will be unable to keep up with the absorbing boundary condition at $\psi = w\ln(1-\tau)$ after a crossover time τ_c, such that

$$\sqrt{\tau_c} = -\frac{w}{\sqrt{2\Delta}}\,\ln(1-\tau_c) \equiv -\kappa\ln(1-\tau_c), \tag{3.41}$$

with the key parameter

$$\kappa \equiv \frac{w}{\sqrt{2\Delta}} \approx s\sqrt{\frac{t^*}{2\tau_g}} \approx s\sqrt{\frac{R_0}{2v\tau_g}}. \tag{3.42}$$

A plot of $\tau_c(\kappa)$ is shown in Fig. 3.3.5. The behavior in the two limits $\kappa \gg 1$ and $\kappa \ll 1$ is given by

$$\tau_c(\kappa) \approx \begin{cases} 1 - e^{-1/\kappa} & \kappa \ll 1 \\ \kappa^{-2} & \kappa \gg 1 \end{cases}. \tag{3.43}$$

When $\kappa \gg 1$ (strong selection), the transition from trajectories dominated by genetic drift to those dominated by deterministic natural selection is very fast. The inflationary effects on the survival probability are then negligible, and the relevant time scale is the diffusion time $\tau_l \approx \phi_0^2/(2\Delta)$. For times $\tau \gg \tau_l$, the survival probability of a sector for radial voter models is described by the long-time linear range expansion result [9]

$$S_r(\tau \gg \tau_l) \approx \frac{1 - e^{-w\phi_0/\Delta}}{1 - e^{-2\pi w/\Delta}} \qquad (\kappa \gg 1). \tag{3.44}$$

If $\tau \ll \tau_l$, however, we must either use the full linear survival probability solution of Eq. (3.31) (Eq. 3.2.15 of Ref. [3]), or else exploit the small time solution

$$S_r(\tau \ll \tau_l) \approx 1 - \sqrt{\frac{4\Delta\tau}{\pi}} \exp\left[-\frac{\phi_0^2 + 2w\phi_0}{4\Delta\tau}\right] \quad (\kappa \gg 1). \tag{3.45}$$

In our simulations, it is difficult to access the regime $\kappa \gg 1$ for small s since this requires large values of the initial radius R_0. In practice, we are limited to the range $0 \leq \kappa \lesssim 10$. Fig. 3.3.4($a$) shows a simulation for $\kappa \approx 2.4$, with only minor diffusion motion in the sector boundaries visible at the onset of the range expansion. In this case, if the sector survives beyond the relatively small time τ_c, it is quite likely to survive indefinitely.

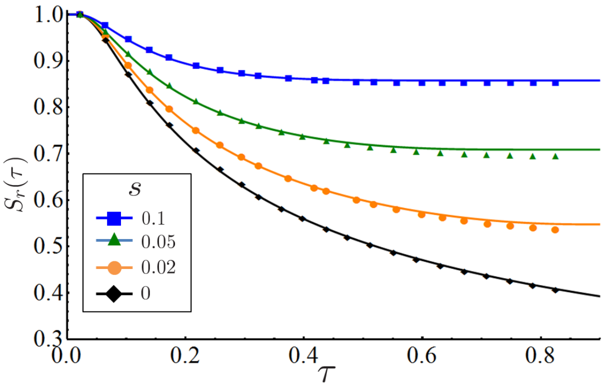


Figure 3.3.6: Radial survival probabilities $S_r(\tau)$ for a green sector (like that in Fig. 3.3.4) of size $\phi_0 \approx 0.036$ initialized on the rim of a homeland with $R_0 = 300$, evolved over 1750 generations, for a variety of selective advantages s. The solid lines represent the adiabatic approximation to $S_r(\tau)$ (Eq. (3.47)). The survival probabilities were averaged over 8×10^4 independent runs. We fit the black line (the $s = 0$ case) with the lattice parameter $\tilde{a}_r$, finding $\tilde{a}_r \approx 0.9$. The other lines use $\gamma \approx 1.2$ (see Appendix B.3).

When $\kappa \ll 1$, however, genetic drift dominates the deterministic sector size prediction in Eq. (3.35) over the entire interval $\tau \in (0, \tau_c)$. As discussed above, a single sector is quite unlikely to take over the whole population for values of ϕ_0 not too close to 2π. We can then approximate the diffusive motion of $\psi = \phi + w \ln(1-\tau)$ by confining it to a semi-infinite interval between $b(\tau) \equiv w \ln(1-\tau)$ and ∞, instead of imposing an upper limit of 2π.

When $\tau < \tau_c$, the boundary condition $p_s(\psi = w\ln(1-\tau), \tau) = 0$ moves slowly compared to the diffusion, allowing an adiabatic approximation [3]: The Fokker-Planck equation (Eq. (3.31)) for the sector size probability distribution $p_s(\phi, \tau)$ is treated as a drift-diffusion equation with diffusion coefficient Δ and a time-dependent drift velocity $|db(\tau)/d\tau| = w/(1-\tau)$, evaluated instantaneously at τ. With the initial probability distribution $p_s(\phi, 0) = \delta(\phi - \phi_0)$, the solution to Eq. (3.31) in this approximation follows from the method of images [3]:

$$p_s(\phi,\tau) \approx \frac{1}{\sqrt{4\pi\Delta\tau}}\left[e^{-\frac{(\phi-\phi_0-w\tau/(1-\tau))^2}{4\Delta\tau}} \right.$$
$$\left. - e^{-\frac{w\phi_0}{\Delta(1-\tau)}} e^{-\frac{(\phi+\phi_0-w\tau/(1-\tau))^2}{4\Delta\tau}}\right]. \tag{3.46}$$

The change $dS_r(\tau)/d\tau$ in the radial survival probability must be given by the probability flux $[\Delta\partial_\phi p_s(\phi,\tau)]|_{\phi=0}$ evaluated at the absorbing boundary. Integrating this probability flux for a radial expansion yields

$$S_r(\tau) = 1 - \Delta \int_0^\tau \mathrm{d}\tau' \; [\partial_\phi p_s(\phi,\tau')]|_{\phi=0} \approx 1 - \int_0^\tau \mathrm{d}\tau' \, \frac{\phi_0 e^{-\frac{(\phi_0(1-\tau')+w\tau')^2}{4\Delta(1-\tau')^2\tau'}}}{\sqrt{4\pi\Delta(\tau')^3}} \tag{3.47}$$

$$\approx 1 - e^{-\frac{\phi_0 w}{2\Delta}}\left[1 + \frac{w\phi_0^2}{4\Delta^2}\left(\phi_0 + \frac{w}{2}\right)\right] \operatorname{erfc}\left[\frac{\phi_0}{2\sqrt{\Delta\tau}}\right]$$
$$+ \frac{\sqrt{\tau}w\phi_0}{\sqrt{4\pi\Delta^3}}\left(\phi_0 + \frac{w}{2}\right)e^{-\frac{2\phi_0 w\tau+\phi_0^2}{4\Delta\tau}} \quad (\tau \ll 1), \tag{3.48}$$

where an asymptotic (small τ) approximation is used to evaluate the integral in Eq. (3.47). Fig. 3.3.6 shows that Eq. (3.48) is a remarkably good approximation to $S_r(\tau)$ even in the selection dominated regime $\tau > \tau_c$.

In the weak selection limit $\kappa \to 0$, the movement of the boundary condition can be ignored entirely. Equation (3.48) then reduces to the survival probability of a diffusing particle on a semi-infinite interval with an absorbing boundary at the origin [3]:

$$S_r(\tau) \xrightarrow{\kappa\to 0} \operatorname{erf}\left[\frac{\phi_0}{\sqrt{4\Delta\tau}}\right]. \tag{3.49}$$

Fig. 3.3.6 shows that the theory matches the simulation results for radial expansions with $\phi_0 \approx 0.036$ and $R_0 = 300$ well, especially at small values $\tau = vt/(R_0 + vt)$. As

$\tau \to 1$ (i.e. in the limit of long times), both inflation and selection prevent the extinction of the sector and $S_r(\tau)$ approaches a nonzero limit for all values of s, including $s = 0$. An asymptotic analysis of the integral in Eq. (3.47) for small $w \propto s$ shows that the limiting value $S_r(\tau \to 1) \equiv S_\infty$ for small initial angular sector size ϕ_0 is approximately

$$S_\infty \approx \operatorname{erf}\left[\frac{\phi_0}{\sqrt{4\Delta}}\right] + \left[1 - \frac{\ln\left(\frac{w\phi_0}{2\Delta}\right)}{\sqrt{4\pi\Delta}}\,\phi_0\right]\frac{\phi_0 w}{2\Delta} + \mathcal{O}(w\phi_0^2). \tag{3.50}$$

For completeness, we also consider the regime $\tau_d \equiv 2\pi^2/\Delta \ll 1$, where genetic drift is strong enough to allow an initially small green sector in Fig. 3.3.4 to diffusively take over the entire population. In this case, the finite range of the sector size ϕ becomes important. In the absence of selection ($\kappa \to 0$), the survival probability can be computed using Laplace transform techniques (see e.g., Ref. [3, 9]) and is given by:

$$S_r(\tau) \xrightarrow{\kappa\to 0} \frac{\phi_0}{2\pi} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\,e^{-\Delta n^2\tau/4}\sin\left[\frac{n\phi_0}{2}\right]. \tag{3.51}$$

Note that the results for the survival probability in Eq. (3.51) and Eq. (3.49) coincide when Δ is small, and the finite range of the sector size becomes irrelevant.

3.3.4 Directed Percolation with Inflation

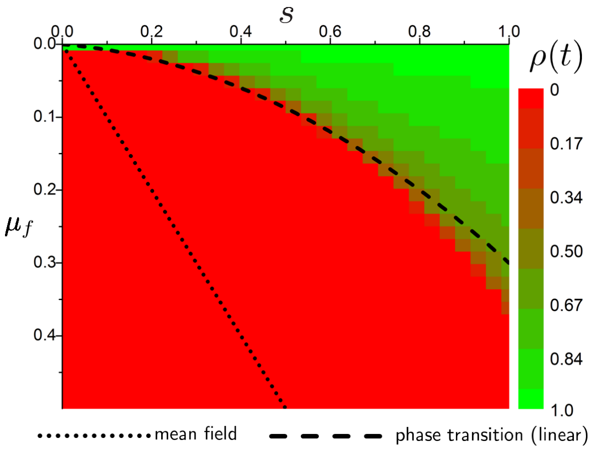


Figure 3.3.7: The color map of the average fraction of green cells at long times ($t = 1800$ generations) for a radial expansion with a homeland radius of $R_0 = 200$, averaged over 400 runs, for positive values of μ_f and s. The average fraction $\rho(t)$ approaches a steady state at long times so that $\rho_\infty = \lim_{t\to\infty}\langle f(t)\rangle = \lim_{t\to\infty}\langle f(\mathbf{x},t)\rangle_{\mathbf{x}} \approx \rho(t = 1800)$. This approximation only breaks down extremely close to the phase boundary between the red (dark gray) and green (light gray) regions and the difference is not visible at the resolution of the plot. The dotted line shows the mean field transition. The dashed line shows the approximate position of the directed percolation phase transition for a linear expansion, as shown in Fig. 3.1.1.

As illustrated in Fig. 3.1.1, linear range expansions within the Domany-Kinzel (DK) model on a hexagonal lattice for $\mu_b = 0$ lead to a phase transition driven by competition between the selective advantage s of the green cells and deleterious forward mutations μ_f, from green to red. In population genetics, a two allele model with irreversible mutations is an important limiting case in the theory of quasi-species [111, 112, 113]. In this theory, the phase transition corresponds to an "error threshold" at which a well adapted population's fitness distribution near a single fitness maximum becomes delocalized due to the large phase space available for deleterious mutations (see Appendix B.1 for more details). In this section, we will compare the well studied DK model results on a hexagonal lattice [95, 100, 102, 103] with our radial range expansion model results on a disordered Bennett lattice.

When selection is strong, the green strain is able to survive even when $\mu_f > 0$. However, for large enough μ_f, we expect a “mutational meltdown” such that the advantageous green strain eventually goes extinct [9, 95]. It is known that in the biologically relevant region of reasonably small, nonzero deleterious mutation rates (e.g. $0 < \mu_f < 0.5$), the model exhibits a phase transition in the directed percolation universality class. This transition is in the one-dimensional compact directed percolation (CDP) or “voter model” universality class when $\mu_f = \mu_b = 0$. The phase diagram as a function of parameters analogous to our parameters s and μ_f for $\mu_b = 0$ has been studied extensively (see [95, 103] and references therein) and is illustrated in Fig. 3.1.1. Note that the CDP transition is connected to an entire line of directed percolation phase transitions.

Near the phase transition line, various dynamical quantities as power laws in time, similar to the $1/t$ behavior of the density of surviving green cells with random initial conditions found in mean field theory (Eq. (B.1)) [95, 103]. In addition, for a fixed mutation rate $\mu_f > 0$, one can show that there are four critical exponents that characterize the scaling of the system. Two exponents, $\nu_\perp$ and $\nu_\|$, describe the diverging width and length of the growing green clusters parallel and perpendicular to the time-like direction. If the directed percolation transition occurs at a critical selective advantage $s_c(\mu_f)$, and $\delta = s - s_c$ ($0 < \delta \ll 1$) measures the distance from the critical point in the active phase, then a typical width and length of the green sectors diverge as $\xi_\perp \sim \delta^{-\nu_\perp}$ and $\xi_\| \sim \delta^{-\nu_\|}$, respectively.

Two additional exponents, β and β', describe order parameters relevant for two distinct initial conditions: (1) the average density of active sites (green cells) at long times, $\rho_\infty \equiv \langle f(t \to \infty) \rangle$ starting from $f(0) = 1$; and (2) the long-time sector survival probability $S_\infty \equiv S_{\mathrm{ss}}(t \to \infty)$ of a single green cell (more generally, any small green sector) seeded in an all red homeland. Here $S_{\mathrm{ss}}(t)$ is the probability that a single sector has survived out to time t. We then have $\rho_\infty \sim \delta^\beta$ and $S_\infty \sim \delta^{\beta'}$ for small distances $\delta > 0$ into the active phase. In the biological context, the exponent β describes the mutational meltdown in a population of green cells during a range expansion, and β' characterizes the survival probability of a rare advantageous mutant that spreads through the population while experiencing deleterious back mutations to a less fit strain. See Fig. 3.3.7 for $\rho_\infty(\mu_f, s)$ for radial expansions. Although we don't expect a sharp phase transition (see below), the general characteristics mimic the linear expansion diagrammed in Fig. 3.1.1.

An important feature of the directed percolation (DP) dynamics described by Eqs. (3.9, 3.10) is the so-called *rapidity reversal* symmetry that implies $\beta = \beta'$. Rapidity reversal is a special kind of time reversal that can be seen explicitly in the field theoretic formulation of the Langevin equation, which is known to be equivalent to a Reggeon field theory [95, 103]. Rapidity reversal symmetry is only valid asymptotically as $\delta \to 0^+$, where it implies that $\rho(t) \equiv \langle f(t) \rangle \propto S_{\mathrm{ss}}(t)$ near the directed percolation phase transition for large values of t. Indeed, both quantities scale for long times like $\rho(t) \sim S_{\mathrm{ss}}(t) \sim t^{-\alpha}$, where $\alpha = \beta/\nu_\parallel = \beta'/\nu_\parallel$. The asymptotic relation $\rho(t) \propto S_{\mathrm{ss}}(t)$ is also valid for systems with time-independent, small system sizes since the size is invariant under time reversal (see Ref. [99] for an example). However, unlike a typical finite size effect, inflation is inherently asymmetric in time, and our simulations will show that inflation breaks the rapidity reversal symmetry, leading to new physics beyond conventional finite size scaling.

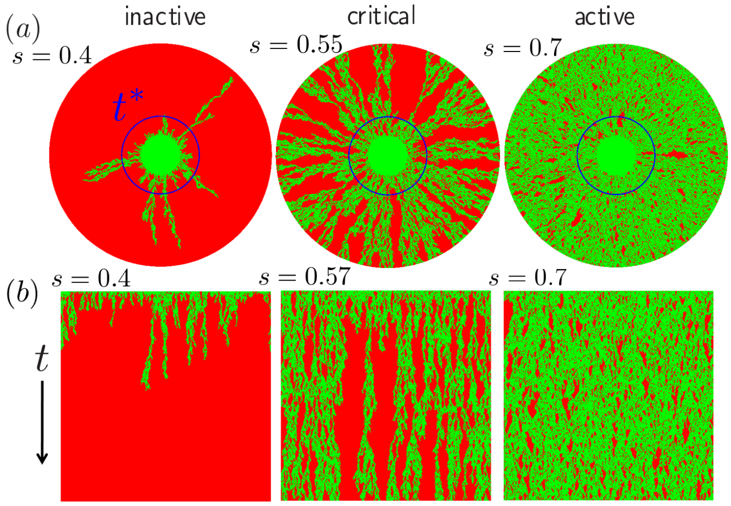


Figure 3.3.8: Linear and radial range expansions with $\mu_b = 0$ and $\mu_f = 0.1$ for various s near the directed percolation phase transition. Green (light gray) organisms have a selective advantage s, but are subject to forward mutations to a less fit red (dark gray) strain. In the radial expansions (a) $R_0 = 50$ and the system evolved for ~ 250 generations. The dark blue solid circles indicate the crossover time $t^* = 100$. In (b) a linear range expansion of 300 cells evolves for 400 generations.

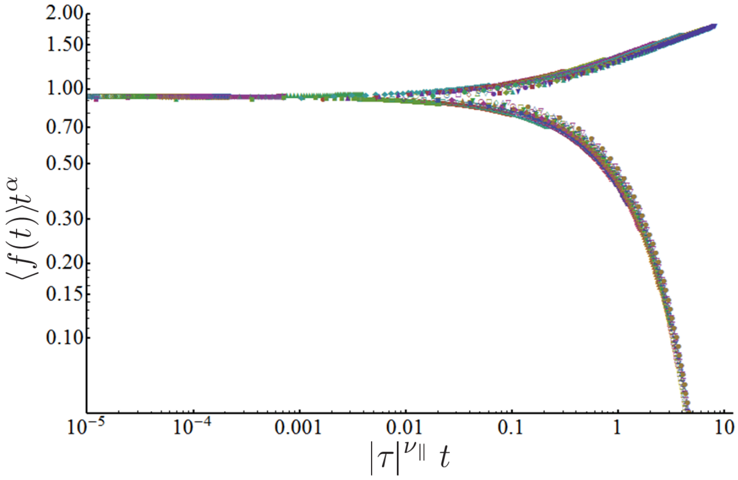


Figure 3.3.9: Data collapse for a radial range expansion with $R_0 = 2048$ for the average number of green cells over about 600 generations. We test many values for $\tau = s - s_c = \pm 0.01, 0.02, \ldots$ for $\mu = 0.1$. The collapses are consistent with the directed percolation values $\alpha \approx 0.159464(6)$ and $\nu_\parallel \approx 1.733847(6)$. The critical selection parameter was approximately $s_c \approx 0.55(1)$.

Consider the DP transition in the radial setting. The plot in Fig. 3.3.7 of the averaged green cell density after 1800 generations ($\rho(t = 1800) \approx \rho_\infty$) for an all green homeland with $R_0 = 200$ shows that we have a crossover between an active and an inactive phase. Despite the existence of these two phases, interpreting the critical point for radial expansions as a DP phase transition requires care! Inflation inhibits the boundaries of green sectors that have formed during time t^* from interacting diffusively at longer times $t \gg t^*$; hence a large enough green sector could, with small probability, survive to the inflationary regime even in the inactive phase, after which it will persist for a very long time. Of course, increasing the forward mutation rate μ_f still creates a crossover in the radial expansion dynamics, since any green sector will eventually become contaminated with red mutations.

The crossover in a radial range expansion is slightly shifted relative to the linear expansion transition line (dashed line in Fig. 3.3.7). The precise value of the transition is not universal and varies with the lattice model details. The direction of the shift in our case might be explained by the larger effective coordination number $z_r > 2$ in the disordered Bennett lattice relative to the hexagonal lattice: A larger z should move the transition closer to the mean field line (dotted line in Fig. 3.3.7) as each daughter in a new generation can evolve from more potential parents.

For early times $t \ll t^*$, inflation should only have a minor effect on the coarsening dynamics at the transition. As a check, Fig. 3.3.9 shows a collapse of the data for the average fraction $\rho(t) = \langle f(t) \rangle$ of green cells for a variety of selective advantages starting with an all green initial population with $R_0 = 2048$. Our simulations run for about 600 generations for different values of s at a fixed $\mu_f = 0.1$. We successively collapsed the data using the remarkably precise estimates for the directed percolation exponents

$\alpha \approx 0.159464(6)$ and $\nu_{\|} \approx 1.733847(6)$ quoted in Ref. [95]. As expected from Fig. 3.3.7, the position of the apparent critical point for times $t \ll t^*$ is shifted slightly relative to the linear case ($s_c \approx 0.55$ versus $s_c \approx 0.57$ for $\mu_f = 0.1$). Fig. 3.3.9 confirms that the early time dynamics of our model is indeed accurately governed by directed percolation.

At late times $t \gg t^*$, the population fronts become locally flat due to the increasing population radius. Thus, one might naively expect that the $t \gg t^*$ dynamics at the transition is also governed by critical DP dynamics. However, inflation causally disconnects portions of the population after time t^*. Thus, the $t \gg t^*$ dynamics consists of a set of noninteracting DP processes. The fluctuations around time t^* set the initial conditions for each process. These initial conditions can be quite different (ranging from all green cells to a single green cell) and lead to different DP dynamics with different power laws [95, 103]. Since we average over the behaviors of all the sectors in a radial range expansion, the $t \gg t^*$ regime does not correspond to the usual DP dynamics with a single initial condition.

One case in which the long-time behaviors of the linear and radial range expansions *are* similar is in the limit $\mu_b \ll s$ for an all green initial condition, in which the expansions both achieve an active steady state. The dynamics is deep in the "active" phase – mutations generate small red sectors which die out with high probability. If these sectors are sufficiently dilute, we can neglect their interactions. This phase is illustrated on the right panel in Fig. 3.3.8(a). Selective advantage bias tends to squeeze out red sectors, which we model by first replacing w by $-w$ in Eq. (3.31) to study the probability distribution of *red* sector widths. However, unlike the analysis of green sector widths in Sec. 3.3.3, the "homeland" radius R_0 must now be replaced by $R(t_r) = R_0 + vt_r > R_0$, where t_r is the time at which a red sector is generated by mutations.

To obtain the long-time behavior of the density $\rho_\infty \equiv \rho(t \to \infty) = \langle f(t \to \infty)\rangle$ of green sites, we examine the survival rate of red sectors formed at very large times $\bar{t} \gg t^*$. In this case, the fate of the red sector is only weakly influenced by inflation. The red sector size probability distribution then obeys

$$\partial_t p_s(\phi, t) = \overline{w}\,\frac{\partial p}{\partial \phi} + \overline{\Delta}\,\frac{\partial^2 p}{\partial \phi^2}, \tag{3.52}$$

with the *constant* coefficients, for a fixed time $t' > t^*$,

$$\overline{w} \approx \frac{\gamma s \tilde{a}_r}{\tau_g R(\bar{t})} \quad \text{and} \quad \overline{\Delta} \approx \frac{\tilde{a}_r^2}{\tau_g (R(\bar{t}))^2}. \tag{3.53}$$

As usual, we impose an absorbing boundary condition $p(\phi = 0, t) = 0$ on Eq. (3.52). If we change variables to make Eq. (3.52) a conventional diffusion equation, then the boundary condition now *advances* on the domain where diffusion and selection bias take place. In such a case, the absorbing state (extinction of a red sector) is always reached [3]. Finally, note that we are no longer strictly in the small s regime and hence have no simple expression for γ in Eq. (3.53). However, we expect it to be close to unity (see Appx. B.3). Since we are only interested in an approximation of the scaling of $\langle f(t \to \infty)\rangle$ with μ_f and s, we set $\tilde{a}_r = \gamma = 1$ in the following for simplicity.

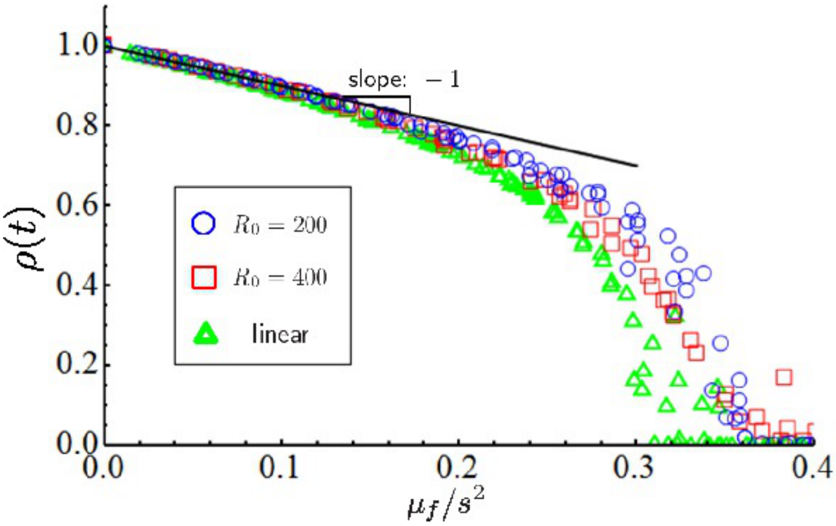


Figure 3.3.10: We plot the average green cell density $\rho(t) \equiv \langle f(t) \rangle$ at long times t versus a control parameter μ_f/s^2 for an initially all green population for radial and linear expansions. For radial expansions, $f(t)$ is averaged over 10^3 runs at $t = 2\times10^3$ generations. In the linear case, we averaged $f(t)$ over 2×10^3 runs, at time $t = 2\times10^4$, for a system size of 10^3. Away from the transition region ($\mu_f/s^2 \lesssim 0.3$), the time t is long enough so that the density $\rho(t)$ for both expansions is approximately equal to the steady-state density $\rho_\infty \equiv \langle f(t \to \infty) \rangle$, with an error smaller than the point size. The black line shows the predicted scaling of the steady-state density for small $\mu_f/s^2 \lesssim$. We varied s between 0 and 1 and μ_f between 0 and 0.5.

The average excursion area (red sector size) $\langle A_{\text{red}} \rangle$ of the random walk described by Eq. (3.52) before it reaches the absorbing state, to first order in the initial sector size of one cell, $\phi_0 \approx \tilde{a}_r/R(\bar{t})$, is given by a known result from diffusion theory (see e.g., Ref. [114]):

$$\langle A_{\text{red}} \rangle = \frac{\overline{\Delta} R(\bar{t})\phi_0}{(\overline{w})^2} \approx \frac{1}{s^2}. \tag{3.54}$$

Next, consider a population of green cells (in the active state) evolving from time $t = \bar{t}$ to time $\bar{t} + T$ (with $\bar{t} \gg T, t^*$). The number of red sectors N_{red} that are seeded in the population during this time will be approximately $N_{\text{red}} \approx 2\pi\mu_f vTR(\bar{t})/\tilde{a}_r^2$. The total area covered by the population from time $\bar{t}$ to time $\bar{t} + T$ is $A_{\text{tot}} \approx 2\pi R(\bar{t})T$. Upon dividing the average area covered by the green cells by A_{tot}, we find the steady-state density of the green cells:

$$\begin{aligned}\rho_\infty &\equiv \langle f(t\to\infty)\rangle \\ &= \frac{A_{\text{tot}} - N_{\text{red}}\langle A_{\text{red}}\rangle}{A_{\text{tot}}} \approx 1 - \frac{\mu_f}{s^2} \quad (\mu_f \ll s). \end{aligned} \tag{3.55}$$

Note the crucial difference between Eq. (3.55) and the steady state value $1-\mu_f/s$ in the mean field (well mixed) solution in Eq. (B.1): For fixed values of μ_f, we require larger values of $s \approx \sqrt{\mu_f} > \mu_f$ to get a comparable reduction in ρ for the circular range expansion.

Eq. (3.55) is derived in the absence of inflation, so the same expression for ρ holds for linear range expansions, similar to the result obtained for a related model in Ref. [9]. This is consistent with the similarity between the active phase pictures for radial and linear range expansions in Fig. 3.3.8(a) and (b), respectively. We now check Eq. (3.55) against our data in the phase diagram of Fig. 3.3.7 for both linear and radial expansions. Fig. 3.3.10 shows collapse of the data for $\mu_f/s^2 \ll 1$, i.e., deep in the active phase. As argued in Ref. [9], ρ is driven below $1-\mu_f/s^2$ by collisions between red sectors, which merge to shield green sectors from invading the red regions, increasing the total fraction of red cells. Red sector collisions are suppressed by inflation, so ρ stays closer to the $1-\mu_f/s^2$ line for radial expansions as opposed to linear ones, as is evident in Fig. 3.3.10. Note that the slope of the line at small μ_f/s^2 in Fig. 3.3.10 is consistent with our approximation $\tilde{a}_r = \gamma = 1$.

As we keep increasing μ_f/s^2, the steady-state density ρ_∞ eventually goes to zero. In the linear case, there is a sharp transition at around $\mu_f/s^2 \approx 0.3$, consistent with the results of Ref. [9]. This extinction of the green organisms is triggered, of course, by crossing the directed percolation phase transition line [95]. A sharp transition is less evident in Fig. 3.3.10 for *radial* expansions: the density ρ seems to decay to zero more slowly as we approach the crossover. We shall now study the dynamics of the radial range expansion near this crossover.

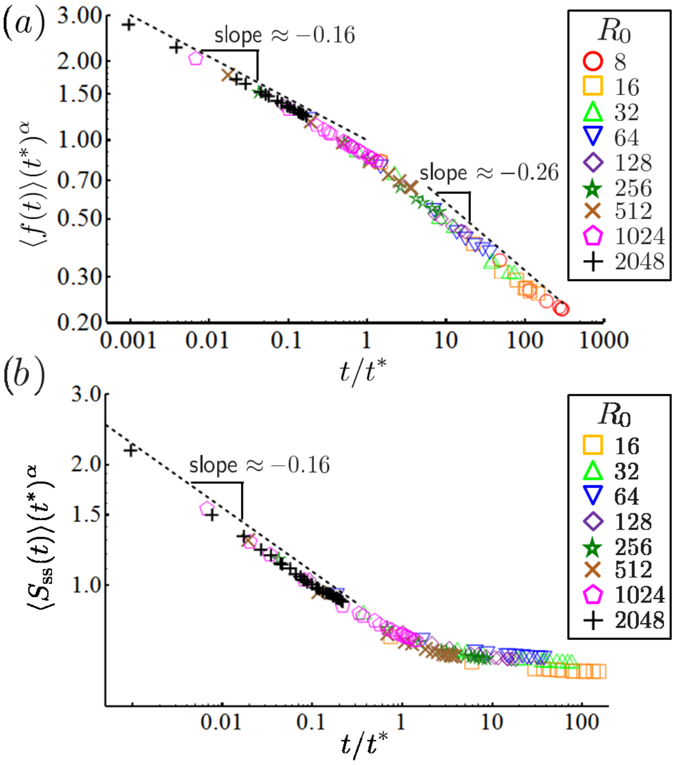


Figure 3.3.11: Data collapses for radial range expansion simulations with many values of $t^* = R_0/v$ (with $\mu_f = 0.1$ and $s = 0.547$). In (a) we plot the scaled green cell density $\rho(t) = \langle f(t)\rangle$ with an all green initial condition for $R_0 = 8, 16, 32, \ldots, 2048$. The regimes $t \ll t^*$ and $t \gg t^*$ can be described by power laws. The $t^{-\alpha}$ decay for $t < t^*$ with $\alpha \approx 0.16$ is expected from directed percolation theory [95], but the exponent $\alpha \approx 0.26$ that arises for long times appears to be new. In (b) we plot the scaled sector survival probability $S_{\text{ss}}(t)$ of a single green cell inoculated in an all red homeland $R_0 = 16, 32, \ldots, 2048$. The dashed line indicates the expected directed percolation result: a decay with slope $-\alpha \approx -0.16$. Note, however, that the survival probability levels off to a finite value for $t \gg t^*$ due to inflation.

To explore directed percolation dynamics when inflation has set in, i.e. for $t > t^* = R_0/v$, we use a dynamic scaling hypothesis for the average green fraction

$\rho(t) = \langle f(t) \rangle$ and for $S_{\text{ss}}(t)$, where $S_{\text{ss}}(t)$ is the survival probability of a single green cell inoculated at the edge of a green homeland. We let $\delta = s - s_c(\mu_f)$ be the distance from the linear inoculation ($R_0 \to \infty$) critical point. We approximate the critical point $s_c(\mu_f)$ by looking at the short time ($t \ll t^*$) dynamics of a radial range expansion with a large homeland radius R_0. For example, we found that $s_c(\mu_f = 0.1) \approx 0.55$ (see Fig. 3.3.9). Note that we cannot use the hexagonal lattice for linear expansions to find the critical point because we expect the point to shift when we use the disordered Bennett lattice.

Inspired by finite size scaling ideas near critical points, we introduce a scaling transformation $\delta \to b\delta$ and expect the scaling relations

$$\rho(t) = b^{-\beta} \tilde{\rho} \left[b^{-\nu_\parallel} t, b\delta, b^{-\nu_\parallel} t^* \right], \tag{3.56}$$

$$S_{\text{ss}}(t) = b^{-\beta'} \tilde{S}_{\text{ss}} \left[b^{-\nu_\parallel} t, b\delta, b^{-\nu_\parallel} t^* \right], \tag{3.57}$$

where we have suppressed metric factors [95], and $\tilde{\rho}$ and $\tilde{S}_{\text{ss}}$ are scaling functions. We now set $b^{-\nu_\parallel} t^* = 1$ and find that our data for various t^* (corresponding, say, to different homeland radii $R_0 = vt^*$) at the critical point ($\delta = 0$) should collapse upon plotting $\rho(t)(t^*)^\alpha$ and $S_{\text{ss}}(t)(t^*)^\alpha$ versus t/t^*. These scaling forms are tested in Fig. 3.3.11. The data collapse reasonably well, and we indeed get a crossover at $t/t^* \approx 1$ in both cases. The inflation crossover time t^* enters just as a finite size effect would in conventional critical phenomena.

Both of the collapses are consistent with a single value α given by directed percolation theory [95]. This suggests that we retain the rapidity reversal symmetry of directed percolation in a linear geometry at early times. However, the rapidity reversal prediction [95] $\rho(t) \propto S_{\text{ss}}(t)$ is violated in the inflationary regime $t > t^*$. Inflation prevents green sectors from dying out for a single seed initial condition (Fig. 3.3.11b), and the survival probability $S_{\rm ss}(t)$ approaches a nonzero constant at long times. Conversely, inflation evidently enhances the loss of the total fraction of green cells for the green homeland initial condition, and the density $\rho(t)$ decays to zero as $t \to \infty$ (Fig. 3.3.11a). In a conventional finite size effect, $S_{\rm ss}(t)$ and $\rho(t)$ are consistent with rapidity reversal symmetry, and both decay exponentially to zero at long times [99].

The levelling off of the survival probability $S_{\rm ss}(t)$ in Fig. 3.3.11(b) for $t \gg t^*$ is consistent with the analysis of single sectors without mutations in Sec. 3.3.3: The survival probability of a single sector approaches a constant at long times (see Eq. 3.50) even at $s = 0$. Unlike a sector in a linear expansion, the diffusive motion of the sector boundaries in a radial range expansion is suppressed in the inflationary regime. The boundaries are unlikely to collide in this regime, and a sector that makes it to the inflationary regime can survive even as $t \to \infty$. In a linear expansion, $S_{\rm ss}(t)$ decays to zero at the directed percolation transition as $t^{-\alpha}$ for large times t.

The more rapid decay of the average fraction of green cells $\rho(t)$ in Fig. 3.3.11(a) for $t > t^*$ is more subtle: $\rho(t)$ seems to decay approximately as $t^{-0.26}$. We can understand this trend qualitatively by appealing to the mean field solution for $\rho(t)$ (see Eq. (B.1)) which decays at the much faster rate t^{-1} at criticality. Since inflation suppresses both diffusion and noise in our Langevin equation for $f(\mathbf{x}, t)$ (Eq. (3.14)), it is plausible that $\rho(t) = \langle f(t) \rangle$ decays faster than for the linear expansion where diffusion and noise are more important. However, it is not clear that the larger decay exponent in the inflationary regime is universal, and we do not have an explanation for its particular value of -0.26.

Another important dynamical quantity is the pair connectedness correlation function [103], i.e., the probability $\Upsilon(x,t;x_0,t_0)$ of finding a directed, causal chain of entirely green cells between a green cell at point (x_0,t_0) and any point (x,t) for $t>t_0$. For radial Domany-Kinzel expansions, $\Upsilon(x,t;x_0,t_0)$ is the probability that a green cell at position $x_0=R(t_0)\phi_0$ along the population circumference at time t_0 has a green cell descendant at some later time t at position $x=R(t)\phi$. Thus, $\Upsilon(x,t;x_0,t_0)$ characterizes the spatial correlations of genetic lineages in the green cells. The pair connectedness function does not keep track of the red cell lineages, including those red cells that evolved from green cell ancestors. A knowledge of $\Upsilon(x,t;x_0,t_0)$ might be useful for genetic inference, where spatial distributions of genetic variants at time t are used to infer ancestral genetic distributions (see Ref. [13] for more information).

We can probe $\Upsilon(x,t;x_0,t_0)$ by employing a single green cell initial condition at the origin and monitoring the resultant green sector formed by descendants of the single green ancestor. We compute the total number of green (i.e. active) cells $N_a(t)$ at time t and the mean square spread of green clusters that survive to time t: $X_a^2(t)\equiv\langle|\Delta x|^2\rangle$, where

$$\langle|\Delta x|^2\rangle=\frac{\int(\Delta x)^2\Upsilon\,(\Delta x,t;0,0)\;\mathrm{d}(\Delta x)}{\int\Upsilon(\Delta x,t;0,0)\,\mathrm{d}(\Delta x)},\tag{3.58}$$

and we average over all green cells located at displacements $\Delta x=R(t)\Delta\phi$ away from the initial green cell. The quantity $N_a(t)$ is related to $\Upsilon(x,t;0,0)$ via [103]

$$N_a(t)=\text{const.}\times\int\mathrm{d}x\,\Upsilon(x,t;0,0)\tag{3.59}$$

and obeys a scaling relation (see Ref. [95, 103])

$$N_a(t) = b^{\nu_\perp - \beta - \beta'} \tilde{N}_a(b^{-\nu_\perp} t, b\delta, b^{-\nu_\parallel} t^*). \tag{3.60}$$

Similarly, we have

$$X_a^2(t) = \frac{\text{const.}}{N_a(t)} \int \mathrm{d}x \, |x|^2 \, \Upsilon(x, t; 0, 0) \tag{3.61}$$

and the scaling relation

$$X_a^2(t) = b^{2\nu_\perp} \tilde{R}_a^2(b^{-\nu_\perp} t, b\delta, b^{-\nu_\parallel} t^*), \tag{3.62}$$

where we have again suppressed metric factors [103].

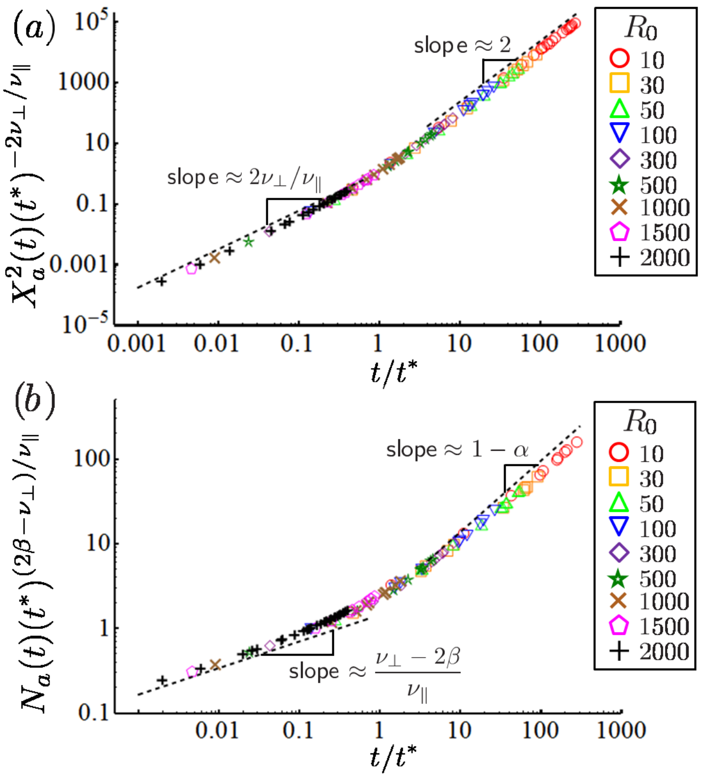


Figure 3.3.12: Data collapse for radial expansion simulations with a single green cell in an otherwise red initial homeland perimeter (with $\mu_f = 0.1$, $\mu_b = 0$, and $s = 0.547$) on log-log plots. (a) The rescaled mean squared cluster spread $X_a^2(t)$. The dashed line for $t \ll t^*$ indicates the directed percolation power law scaling with exponent: $2\nu_\perp/\nu_\parallel \approx 1.3$. The dashed line in the inflation-dominated $t \gg t^*$ regime agrees with a $R_a^2(t) \sim t^2$ behavior discussed in the text. (b) Plot of the rescaled average number of green cells $N_a(t)$. The expected directed percolation scaling for $t \ll t^*$ is shown with a dashed line, with $(\nu_\perp - 2\beta)/\nu_\parallel \approx 0.31$. For $t \gg t^*$, we find $N_a(t) \sim t^{1-\alpha} \approx t^{0.84}$, consistent with arguments incorporating inflation given in the text. The simulation consisted of 9.6×10^4 independent runs for each homeland with radii $R_0 = 10, 30, 50, 100, \ldots, 2000$, evolved to a final population of size $R(t) = 2800$.

Upon setting $b^{-\nu_\parallel} t^* = 1$, we now find data collapse when we plot $X_a^2(t)(t^*)^{-2\nu_\perp/\nu_\parallel}$ and $N_a(t)(t^*)^{(\beta+\beta'-\nu_\perp)/\nu_\parallel}$ versus t/t^* at criticality ($\delta = 0$) (see Fig. 3.3.12). In our simulations, $X_a^2(t)$ can be computed by finding the average squared angular spread $\langle\phi^2\rangle$ and multiplying by $R^2(t) = (R_0 + vt)^2$. We find excellent data collapses for both quantities. Notice that for times $t \ll t^*$, we get the expected power-law behavior given by the directed percolation universality class [103].

For $t \gg t^*$, we have argued that a green sector evolved from a single green cell will expand deterministically. The angular width of the sector will increase somewhat due to the selective advantage, but inflation will be the dominant effect at long times. Hence, the main contribution to X_a^2 for $t \gg t^*$ arises from large green sectors evolving with fixed angular widths. We then expect $X_a^2(t) \sim t^2$ at long times, consistent with the long-time behavior in Fig. 3.3.12(a). The main contribution to $N_a(t)$ also arises from this deterministic expansion epoch. However, we expect the density of green cells *inside* each sector to decay according to $t^{-\alpha}$, since the interior of a green sector should be describable as a critical directed percolation process. Thus, we expect that $N_a(t) \sim t^{1-\alpha}$ for $t \gg t^*$. This expectation also matches the simulation results (see Fig. 3.3.12(b)). By contrast, directed percolation in a finite geometry is unable to sustain large cluster growth. Thus, at long times, $N_a(t)$ and $X_a^2(t)$ have different scaling functions that rapidly decay to zero [99].

The scaling result for the mean square spread X_a^2 of surviving green clusters suggests a simple prediction for directed percolation with arbitrary power law inflation $R(t) = R_0[1 + (t/t^*)^{\Theta}]$. Namely, for $\Theta > \nu_\perp/\nu_\parallel \approx 0.632$, the inflation should be able to overtake the DP critical cluster growth, so that $X_a^2 \sim t^{2\Theta}$ and $N_a \sim t^{\Theta-\alpha}$ at long times $t \gg t^*$. X_a^2 and N_a should still have their regular DP power laws for $t \ll t^*$. However, the shape and position of the crossover region at $t \sim t^*$ will likely change with Θ. Conversely, when $\Theta < \nu_\perp/\nu_\parallel$, the cluster growth is always faster than the inflation, and we expect the inflation to be *irrelevant* so that $X_a^2 \sim t^{2\nu_\perp/\nu_\parallel}$ and $N_a \sim t^{(\nu_\perp - 2\beta)/\nu_\parallel}$ at long times for all $\Theta < \nu_\perp/\nu_\parallel$. Of course, a rigorous check of this simple prediction is necessary. It would also be interesting to study the borderline case $\Theta = \nu_\perp/\nu_\parallel$.

3.4 Radial Range Expansions with Deflation

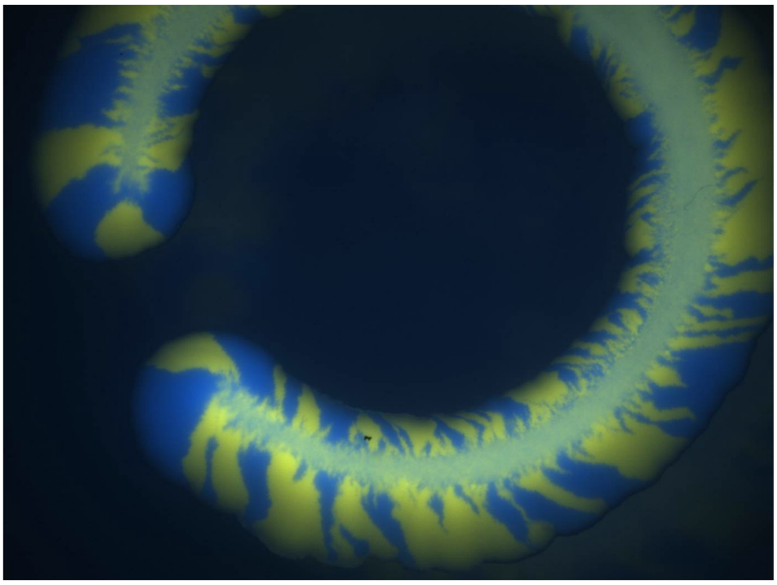


Figure 3.4.1: An inoculation of two bacterial (*Escherichia coli*) strains on a Petri dish, labelled with blue and yellow constitutively expressed fluorescent proteins. The strains are otherwise identical and, consequently, represent the $s = \mu_f = \mu_b = 0$ case in our analysis. The initial inoculation has the shape of an partial ring to allow for nutrients to penetrate the ring center and allow for the bacteria to grow inward, creating a “deflating” radial range expansion. A regular “inflating” radial range expansion moves away from the ring. This single experiment illustrates both kinds of radial expansion discussed in this chapter!

Another type of radial expansion starts with a population around the circumference of a circle of radius R_0. The population then grows *inward* toward the center of the circle, colonizing the interior. An example of such a “deflating” range expansion is the settling of an island by a new species that arrives simultaneously around the shoreline and then grows inland. The population front at time t in this case is a circle of radius $R(t) = R_0 - vt$. Note that $t^* = R_0/v$ is now a sharp temporal *cutoff* (instead of a crossover), at which the population front size vanishes and the invading species completely takes over the circular island.

A deflating range expansion can be realized in experiment by inoculating a ring of organisms on a Petri dish, as shown in Fig. 3.4.1. As long as there are enough nutrients on the interior of the ring, the organisms will grow inward toward the ring center. Fig. 3.4.1 shows the genetic demixing occurring between two neutral bacterial strains labelled with constitutively expressed blue and yellow fluorescent proteins. The bacterial sectors exhibit a counterclockwise "twisting" that is analyzed in detail in Ref. [78]. This uniform tangential motion of the bacterial sector boundaries does not affect the genetic demixing or quantities such as the heterozygosity that only depend on the separation between two individuals along a population front. We will not model the twisting in simulations, but note that it can be easily implemented by shifting all of the cells at a population front counterclockwise at each time step.

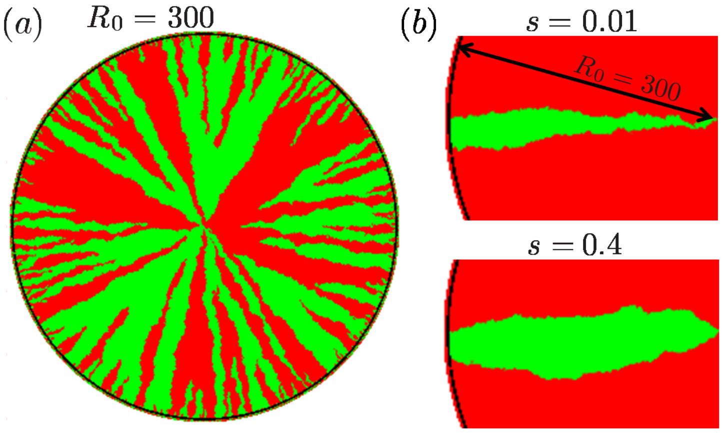


Figure 3.4.2: (a) A radial range expansion with a deflating population front, for $s = \mu_f = \mu_b = 0$, starting from an initial population around a circular rim of radius $R_0 = 300$. We use the same Bennett model lattice described in Sec. 3.2, while running the update scheme in reverse. (b) Deflationary range expansions for a single green sector with an initial size $\phi_0 = 0.1$ radians for a large ($s = 0.4$) and very small ($s = 0.01$) selective advantage. In both cases, deflation forces the green sector boundaries to collide at the center (when $t = t^* = R_0/v$).

The simulated evolution of two neutral strains in a deflating range expansion is shown in Fig. 3.4.2(a). A well mixed initial condition is used with equal proportions of green (light gray) and red (dark gray) cells. The initial population is on a circular boundary with radius $R_0 = 300$. The deflationary dynamics, similar to the inflationary dynamics considered in Sec. 3.3.1, is related to a linear range expansion via a conformal time coordinate $t_c(t) = t/(1 - t/t^*)$. The heterozygosity is again given by $H(\phi, t) = H_0 \operatorname{erf}|\xi|$, with $\xi = R_0\phi/\sqrt{8D_r t_c(t)}$. However, unlike the inflationary case, $H(\phi, t)$ now decays to zero and becomes flat as $t \to t^*$! Thus, the deflating range expansion evolves to a population with just one genetic sector as $t \to t^*$, as shown in Fig. 3.4.2(a).

The dynamics of a single green sector in a deflating radial range expansion (shown in Fig. 3.4.2(b)) also differs from the inflating case. Using the diffusion equation techniques of Sec. 3.3.3 and App. B.3, we can calculate the deflationary survival probability $S_d(t)$ of a green sector with a small selective advantage w and a small initial size $\phi_0 \ll 2\pi$. The result as $t \to t^*$ is

$$S_d(t \to t^*) \approx 1 - \int_0^\infty \frac{e^{-x}}{\sqrt{\pi x}} \, e^{-\frac{2wx\phi_0}{4\Delta x + \phi_0^2} - \frac{w^2 x \phi_0^2}{(4\Delta x + \phi_0^2)^2}} \, \mathrm{d}x, \text{ or} \tag{3.63}$$

$$S_d(t \to t^*) \approx \left[\frac{\phi_0}{\Delta} - \frac{\sqrt{\pi}\phi_0^2}{4\Delta^{3/2}}\right]\frac{w}{2} + \left[\frac{\sqrt{\pi}\phi_0}{4\Delta^{3/2}} - \frac{\phi_0^2}{\Delta^2}\right]\frac{w^2}{4} + \mathcal{O}(\phi_0 w^3) + \mathcal{O}(\phi_0^3 w). \tag{3.64}$$

Note that the convergence of sector boundaries insures that $S_d(t \to t^*)$ vanishes in the absence of selection ($w = 0$). This behavior differs markedly from the inflationary case in which $S_r(t \to \infty)$ has a limiting value given by Eq. (3.49) with $\tau = 1$. However, when $t \to t^*$ in the deflationary case, the population size at the frontier goes to zero, and the finite size of the cells becomes important. Hence, we do not expect the continuum diffusion model to hold exactly as $t \to t^*$; contrary to the continuum result (Eq. (3.63)), a small residual survival probability is possible in this limit, especially for larger initial sector sizes ϕ_0. Thus, Eq. (3.63) and Eq. (3.64) are good approximations to the limiting survival probability only when the fate of the green sector is determined sufficiently far from the sector convergence time $t = t^*$.

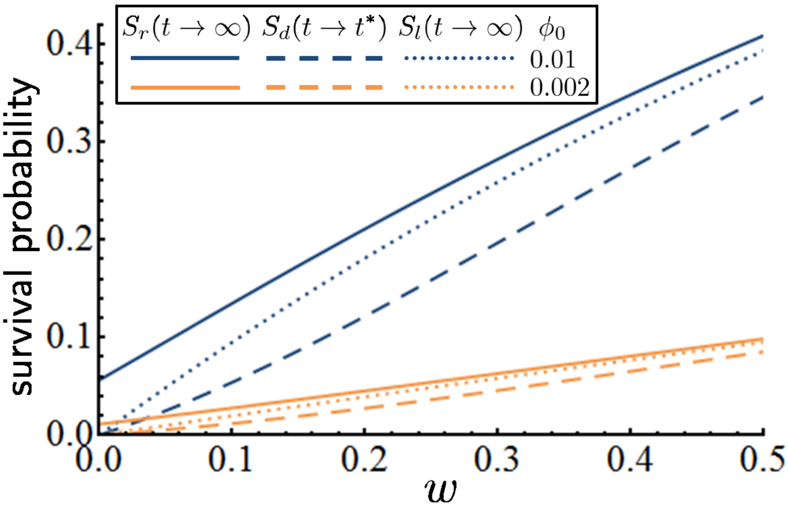


Figure 3.4.3: A comparison of the long-time survival probabilities of green sectors for small selective advantages $w \propto s$ in range expansions with radial inflation ($S_r(t \to \infty)$, solid lines), radial deflation ($S_d(t \to t^*)$, dashed lines), and for linear range expansions ($S_l(t \to \infty)$, dotted lines). The probabilities are calculated using Eq. (3.47), Eq. (3.63), and a similar result for a linear range expansion. In the forward and reverse radial expansions, the initial population radius is $R_0 = 100\, a_r$, where the effective lattice constant $a_r = 1$ for convenience. The linear expansion assumes a total population size $L = 2\pi R_0 = 200\pi a_r$. We used two different initial green sector sizes, $\phi_0 = 0.01$ (upper set of curves) and $\phi_0 = 0.002$ (lower set of curves).

Comparing Eq. (3.64) and Eq. (3.50) reveals that deflation suppresses the effects of selection relative to inflation, an effect which arises because deflation forces the green sector boundaries closer together, as illustrated in Fig. 3.4.2(b). The survival probabilities for inflationary and deflationary radial range expansions can also be compared to the limiting survival probability $S_l(t \to \infty)$ of a green sector in a linear range expansion with selective advantage w and initial size $R_0\phi_0$ in a population of linear size $L = 2\pi R_0$. The results for two different sector sizes ϕ_0 are shown in Fig. 3.4.3. As expected, the survival probability of a green sector in a linear range expansion is bracketed by the probabilities for inflating and deflating radial expansions.

3.5 Conclusions

This chapter has explored differences between the population genetics of radial and linear range expansions in the presence of genetic drift, mutation, and selection. We find that the inflation embodied in a radial expansion decreases the effect of genetic drift and biases the stochastic dynamics in a deterministic direction. The proliferation of extra sites at the outer edge of the radial expansion greatly reduces sector interactions after time $t^* = R_0/v$. For times $t \gg t^*$, the population evolves in isolated angular segments, with boundaries perturbed by genetic drift. Near the directed percolation (DP) phase transition, we found that t^* acts as a finite crossover time, after which the radial expansion no longer experiences the critical dynamics. Finite size scaling near the DP transition due to inflation is described by a different scaling function than that governing conventional finite size effects.

We presented and tested analytical approximations to the dynamics of single sectors in various limits, including selective advantages weak compared to the genetic drift and inflation. It would be interesting to develop a more precise asymptotic analysis of the radial survival probability $S_r(\phi_0, \tau)$ of a sector with an arbitrary initial angular width ϕ_0 and $\tau = t/(t + t^*)$. We also calculated the survival probability in a deflationary radial range expansion in which the sector boundaries are pushed together by a collapsing population front. We found that inflation enhances and deflation diminishes the survival probability relative to a linear range expansion with the same initial population size and selective advantage.

It would also be interesting to study the inflationary effects in more detail via a field theoretic derivation of the scaling hypothesis [95, 103]. In the linear expansion, the relevant field theory is Reggeon field theory. In the radial case, we would have to introduce the inflationary metric $\mathrm{d}s^2 = v^2\mathrm{d}t^2 + R(t)^2\,\mathrm{d}\phi^2$ (with front propagation speed $v = dR/dt$) for measuring distances between points in the range expansion rather than the usual Euclidean one $\mathrm{d}s^2 = v^2\mathrm{d}t^2 + \mathrm{d}x^2$ to take into account the expanding system size. So, we expect that the field theoretic description of radial expansions would involve studying the Reggeon field theory in this "curved" inflationary space.

It would also be interesting to extend our results by allowing for undulations in the radial population front. Such undulations produce small changes in the power laws predicted by simple diffusive models of sector growth and wandering [9]. One way to allow undulations in simulations is to generalize the Eden model proposed in [94] to an amorphous lattice and evolve initially circular homelands of individuals which now divide freely at the population periphery. Without the constraint that daughter cells always spawn as close as possible to the center of the expansion, the population front should roughen away from a nearly circular shape. In the future, we hope to study an extension of our model to three dimensions, thus allowing analysis of a greater variety of range expansions, such as tumor growth in cancer. Spherical cell colonies, for example, can be modeled using the three-dimensional Bennett model [69]. We expect that the effects of inflation are more pronounced in three dimensions since sector surface areas now increase quadratically in time due to inflation, while sector coarsening due to genetic drift is much slower at a two dimensional frontier [13].

Chapter 4

Survival Probabilities at Spherical Frontiers

This chapter will be submitted for publication.

4.1 Introduction

Early tumor evolution is driven by rare driver mutations that sweep the prevascular tumor population at the frontier and push the growing cell mass further down the path toward metastasis. Hence, an understanding of the evolutionary dynamics governing the survival of such mutations is crucial in cancer prevention [79, 115]. One important, largely unexplored aspect of this evolution is the effect of tumor geometry. An important *in vitro* model of cancer is the multicellular tumor with an approximately spherical shape, or "spheroid". The spheroid captures many of the essential features of solid tumors *in vivo* and is an important model for anti-cancer therapies [116, 117, 118].

Spheroids are especially useful for understanding small, avascular tumors. In the later stages of growth, in order for the tumor to survive, it requires a vascular system and undergoes angiogenesis [119]. The growth then becomes more complicated, and requires more sophisticated modelling efforts [120, 121]. We focus here on the earlier evolutionary dynamics of spheroidal range expansions in two and three dimensions. We assume that attractive cell-cell interactions keep such aggregates approximately spherical, i.e. that there is an effective surface tension, similar to that observed for dome-like assemblies of yeast cells on the bottom of, say, a well in a 96-well plate (Chapter 2, [74]). Although we are motivated by tumor evolution, our models are intended to be quite general. Two-dimensional and possibly three-dimensional expansions may be realized in experiment, for example, using microbial or yeast populations in hard and soft agar, respectively (Chapter 3, [78, 92]).

We will be particularly interested in computing the survival probability of a mutation that occurs among the dividing cells at the surface of a spherical or circular population of initial radius $R(t=0) = R_0$, which may or may not increase in time. In general, the radius $R(t)$ has a complicated time dependence, especially in tumor growth. At the early stages, cells divide everywhere inside the tumor and the cluster radius grows exponentially in time. After the tumor reaches a size larger than a nutrient shielding length (see Chapter 2), nutrients will no longer be able to diffuse into the tumor interior. This effect, combined with inward pressure from the surrounding non-cancerous tissue [122, 123, 124], decreases the growth rate toward the center of the tumor. The radius $R(t)$ then grows more slowly. We will model this more inhibited growth generally as

$$R(t) = R_0 \left[1 + \frac{t}{t^*}\right]^{\Theta}, \tag{4.1}$$

where R_0 is the initial tumor radius, Θ is a (possibly time-dependent) growth exponent, and t^* is a cross-over time. For a substantial portion of the growth in tumors, the radius grows linearly in time ($\Theta = 1$) and $t^* = R_0/v$, where v is the front speed [125]. Linear growth and nutrient shielding are also present in microbial populations grown in Petri dishes [92].

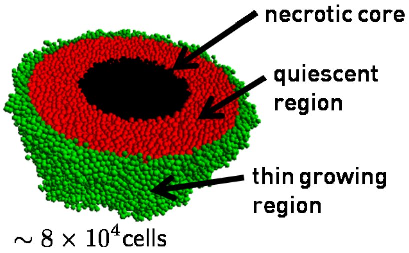


Figure 4.1.1: (Color online) A schematic of a treadmilling tumor. Due to nutrient shielding, cells divide in a thin green region at the frontier. In the red region, cells are in an arrested state and do not grow. In the necrotic core, cells undergo apoptosis and their contents are flushed out of the cluster, resulting in an overall volume loss. This volume loss can balance the gain of volume at the cluster periphery, resulting in a "treadmilling" effect and a cell mass with a constant radius [122, 123, 124, 126].

Eventually, apoptosis may be induced at the tumor center, creating a necrotic region [122, 123, 124], illustrated in Fig. 4.1.1. The cells at the tumor periphery continue to divide relatively rapidly. Thus, a "treadmilling" effect is created and the tumor experiences a rapid turnover of cells at its surface while remaining the same size, a situation we represent by a growth exponent $\Theta = 0$ in Eq. (4.1). We will show that the different growth regimes captured by varying Θ have dramatically different consequences for the fate of mutations at the tumor frontier. We will focus on $\Theta = 0$, $\Theta = 1$, and $\Theta = 1/2$, capturing, respectively, treadmilling, linearly inflating, and an intriguing borderline growth regime. Note that when the population front is inflating, a mutation occurring at the surface, even if it survives at infinite times $t \to \infty$, may not sweep the entire population (i.e., fix). Indeed, the inflating front may accommodate many surviving neutral alleles at long times [13]. We are guaranteed that a mutation either fixes or is extinguished only for non-inflating fronts. Hence, in this chapter, we calculate *survival probabilities* (instead of *fixation* probabilities) in order to compare inflating and non-inflating frontiers directly.

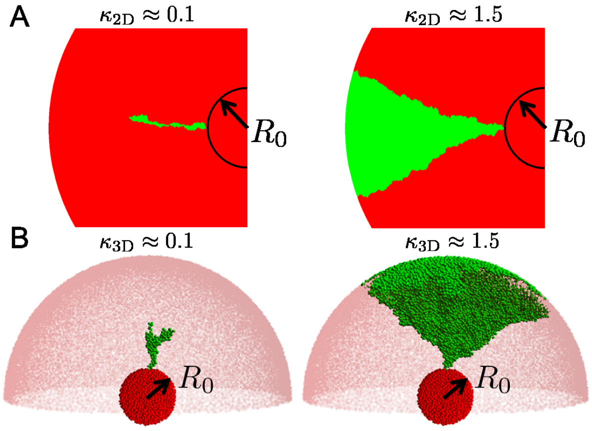

A
$\kappa_{2D} \approx 0.1$
$\kappa_{2D} \approx 1.5$
R_0
R_0
B
$\kappa_{3D} \approx 0.1$
$\kappa_{3D} \approx 1.5$
R_0
R_0

Figure 4.1.2: Examples of simulated mutant clusters (green cells) in two- and three- dimensional range expansions (see Section 4.2). We use two different values of the key dimensionless selection parameters $\kappa_{2D,3D}$ defined in Eq. (4.2). (A) Circular range expansions with a uniform front with an initial radius $R_0 = 50$ average cell diameters and a single initial mutant green cell at the population frontier. (B) Spherical range expansions with uniform fronts and a single green cell at the initial population frontier with radius $R_0 = 10$ cell diameters.

We consider the fate of mutations that occur among the dividing cells at the surface of these populations. Previous studies have focussed on the deterministic growth of the mutant population after it establishes itself [127]. However, before a mutation is established, genetic drift can drive the mutation to extinction, even if it has a selective advantage, as illustrated on the left panels of Fig. 4.1.2A and 4.1.2B in two and three dimensions, respectively. The geometry of the front strongly modifies the probability of these extinction events. For linearly inflating fronts ($\Theta = 1$) in two and three dimensions, we will show that the key dimensionless parameters for mutant survival are, respectively,

$$\begin{cases} \kappa_{\text{2D}} = s\sqrt{\dfrac{R_0}{v\tau_g}} \\ \kappa_{\text{3D}} = \dfrac{sR_0}{v\tau_g} \end{cases}, \tag{4.2}$$

where τ_g is a generation time. For $\kappa_{\text{2D,3D}} \gtrsim 1$, the selective advantage is large enough to sustain a surviving mutant cluster after a short time, allowing it to sweep the population deterministically. When $\kappa_{\text{2D,3D}} \ll 1$, however, the effects of genetic drift are important for a larger portion of the mutant cluster evolution, and can lead to an extinction of the cluster. These two cases are illustrated in Figs. 4.1.2A and 4.1.2B for two- and three-dimensional expansions, respectively.

The remainder of the chapter is organized as follows: We introduce the simulation model used to create these expansions in Section 4.2. We impose a global constraint that insures compact, approximately spherical clusters, as a way of emulating the effect of surface tension (simulations of rougher fronts are discussed in Sec. 4.5.4). In Section 4.3 we review the deterministic dynamics of mutant sectors, exemplified by the average sector shapes corresponding to the right panels of Figs. 4.1.2A and 4.1.2B. We also introduce the stochastic equation governing both the deterministic and stochastic dynamics. In Section 4.4 we review and extend results for two-dimensional expansions, as these provide insights into the three-dimensional case relevant to tumors. In Section 4.5 we calculate the survival probability of mutations in three-dimensions, focussing on linearly inflating tumors with $\Theta = 1$ and relegating details to Appendices C.1 and C.2. The smooth circular and spherical fronts generated in most of our simulations facilitate comparison with analytical results. However, Section 4.5 concludes with a discussion of rough fronts. We make concluding remarks in Section 4.6.

4.2 Simulations

The actively growing region in a tumor mass or a spherical microbial population (see Chapter 2) can be quite thin, with a width of just a few cell diameters. Genetic drift is especially strong at such thin fronts as just a few cells compete to divide into new territory at every point on the surface. We will focus on the limiting case of a single layer of growing cells, although a thicker actively growing layer of cells could be accounted for in a coarse-grained model by increasing the local effective population size [13]. We also assume, to simplify the analysis, *compact* fronts where the population front closely approximates a uniform circle or sphere at all times. The latter approximation will be valid as long as front undulations and the selective advantage s are small. When s is large, the selectively advantageous strain will create prominent bulges at the frontier, spoiling the front uniformity [92, 127]. Front undulations can also lead to important changes in the expansion evolution [94], but these (as well as prominent bulges) can be suppressed, for example, by surface tension at the frontier. As discussed in more detail in Chapter 3, these approximations have the advantage of allowing us to perform a "dimensional-reduction" by focussing on just the dynamics at the population frontier, which will be effectively one and two-dimensional inflating geometries for two- and three-dimensional expansions, respectively.

All along these frontiers, cells will compete locally to divide into new territory and form the next generation of cells. We consider two types of cells: a wild-type red cell and a green cell with a selectively advantageous "driver" mutation. To set up the green and red cell competition, we use the generalized Domany-Kinzel stochastic cellular automaton rules developed in Chapter 3. The mutant, "driver" green cells have a base growth rate $\Gamma_g = 1$ in units of inverse generation times. The red wild-type cells will grow more slowly with rate $\Gamma_r = 1 - s$, where s is a selective advantage. These rules specify the probability $p_g(n_g, z)$ of a green cell division out of z total locally competing "parent" cells, n_g of which are green:

$$p_g(n_g, z) = \frac{\Gamma_g n_g}{\Gamma_r(z - n_g) + \Gamma_g n_g} = \frac{n_g}{(1-s)z + s n_g}. \tag{4.3}$$

New cells are placed one at a time in predefined locations to create expansions with uniform fronts as described below. The probability of a particular cell color dividing into the new spot is determined by Eq. (4.3), with the probability of a red offspring given by $p_r(n_g, z) = 1 - p_g(n_g, z)$.

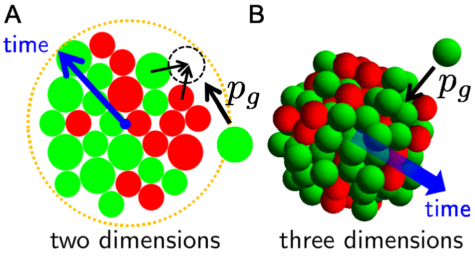


Figure 4.2.1: (Color online) Schematics of simulated radial range expansions in (A) two and (B) three dimensions. In (A), cells of two different sizes compete to divide into new territory. The two different sizes and their relative abundance are chosen to create an amorphous, isotropic cell packing, thus avoiding spurious effects for radial expansions of identically-sized cells in two dimensions, which tend to form a regular triangular lattice (see Chapter 3). In (B), cells divide on a lattice generated by the Bennett model [69]. All cells have identical diameters a in this case. In both (A) and (B), the virgin territory that is closest to the cell cluster center is settled first to create an effect similar to a surface tension. In both models, the green cells out-compete the red with probability p_g given by Eq. (4.3).

In order to achieve smooth, approximately circular or spherical fronts (thus approximating the effect of a surface tension), newly born cells appear sequentially with no overlaps and at locations chosen to be closest to some central reference point, as shown in Fig. 4.2.1. Such a compactification of the cell clusters could also arise for biological reasons, if all cells secrete a chemical that promotes cell divisions, which then are only feasible at the cluster surface for steric reasons (see Chapter 3). This methodology is reasonable for more conventional growing cell clusters, provided surface tension forces only move daughter cells a short distance after they are born. This approximation also increases our computational efficiency.

In $d = 1+1$ dimensions (outwardly growing circular range expansions), strong lattice effects are present in the evolutionary dynamics (see Chapter 3). We mitigate these effects by working with an amorphous cell packing constructed using a modified Bennett model [104], a model originally developed to understand structure in metallic glasses. In this model, cells of two different sizes are placed closest to the center of the initial cluster such that the cell touches two previously placed cells. This procedure creates, uniform, circular population fronts as one might find in expansions with an effective surface tension. In Chapter 3, it was found that isotropic, amorphous cell packings occurred if we chose cells with a small-to-large radius ratio of $\rho \approx 0.727$, with the smaller one placed with 60% probability, choices we also make here. In $d = 2 + 1$ dimensions, lattice artifacts are greatly reduced, and our range expansions are constructed using the original Bennett model [69] with identical spherical cells with diameter a. These cells are now placed closest to a central point such that each touches at least three previously placed spheres. The two kinds of expansion with relatively small population sizes are shown below in Fig. 4.2.1. Note that a cell diameter is more complicated to define in the two-dimensional expansion because of the amorphous packing; here a denotes an effective lattice spacing (approximately equal to the average cell diameter), as described in more detail in Chapter 3.

To further mitigate lattice artifacts, we initialize the mutant green cells during each simulation run in one of at least 100 different, random locations, sampled randomly and uniformly along the population periphery. In two dimensions, all cells along the initial, circular population front have equal probability of being chosen as the mutant green cell. Thus, the mutant may be either a small cell or a large cell, depending on the location chosen. In three dimensions, we first perform a random rotation of the initial spherical population as described in Ref. [128]. We then seed the mutated, green cells at the north pole of the spherical population. This ensures a uniform, random sampling of the initial population front.

This algorithm generates linearly inflating fronts with $\Theta = 1$. It is possible to simulate *treadmilling* tumors with $\Theta = 0$ by adding a step that evolves the population backwards, toward the interior of the cluster: First, we evolve the range expansion just a few generations, generating a cluster of radius $R_0 + nv\tau_g \approx R_0 + na$, where $n \approx 2.5$. Second, we go back to the cells at approximately a distance $R_0 + 1.5a$ from the cluster center and evolve the population a single cell position at a time, but *backwards*, toward the central reference point. To do this, we let the outer shell of cells compete to divide backwards into the interior of the population, with the necessary space for new daughter cells assumed to be provided by apoptosis deep in the interior. This backwards evolution is continued until a population front with radius $R_0 - nv\tau_g$ is created. Finally, we repeat the first step and re-evolve the initial population radius outward up to a radius $R_0 + nv\tau_g$. These backwards and forward sweeps are repeated over and over to model a turn-over of cells at the tumor surface, taken to be the spherical shell of cells with fixed radius R_0 and thickness $v\tau_g$.

4.3 Population Genetics at Compact Population Fronts

When the population front is large compared to the characteristic size of an initial mutant cluster, the coarse-grained green cell (mutant) density $f \equiv f(\mathbf{x},t)$ for a uniform, *flat* front (upon exploiting dimensional reduction) obeys the Langevin equation

$$\partial_t f = D\nabla^2_{\mathbf{x}} f + s\tau_g^{-1} f(1-f) + \sqrt{2\Delta\tau_g^{-1} f\,(1-f)}\,\xi(\mathbf{x},t), \tag{4.4}$$

where s is the selective advantage of the mutant, D is a spatial diffusivity, τ_g is the generation time, Δ is a genetic drift strength, and ξ is a Gaussian, white noise $[\langle \xi(\mathbf{x},t)\xi(\mathbf{x}',t')\rangle = \delta(\mathbf{x}-\mathbf{x}')\delta(t-t')]$ interpreted in the Îto sense [13, 15]. The strength Δ of the genetic drift scales like $\Delta \sim N^{-1}$, where the effective population size N is approximately the number of organisms per unit cell width that locally compete to divide at the population frontier. In the context of a stepping stone model using Moran evolutionary dynamics, $\Delta = 1/N$ exactly [13]. If we remove the genetic drift term in Eq. (4.4) by setting $\xi(\mathbf{x},t) = 0$, we recover the Fisher-Kolmogorov-Petrovsky-Piscunov equation [13, 91]. This equation allows traveling-wave solutions that describe the deterministic sweep of a mutant sector [92]. The genetic drift induced by the noise term can strongly influence these travelling-wave dynamics [17, 129] and can extinguish an embryonic mutant sector. Using simulations and field-theoretic techniques, we will analyze Eq. (4.4) in the strong noise limit and appropriately modify it to take into account inflating population fronts in Sections 4.4 and 4.5. However, it is instructive to first consider the travelling-wave solutions in radial range expansions [corresponding to $\Theta = 1$ in Eq. (4.1)] when selection dominates the genetic drift [92, 127].

If a mutation with some $s > 0$ survives extinction at early times, it will form a sector as shown in Fig. 4.1.2 and schematically in Fig. 4.3.1 for three dimensions. We assume $s \ll 1$ and/or a strong surface tension, so that we may ignore any bulges created by the green strain and to make contact with our simulations. The boundary of this sector is a genetic Fisher wave in which the mutant green strain "invades" the red wild-type region azimuthally along the frontier with some speed $v_\perp$. In a two-dimensional circular expansion, the two sector boundaries will form (provided they escape genetic drift), on average, two counterrotating logarithmic spirals [92]. In three dimensions, this sector shape is a "logarithmic cone," i.e. the solid of revolution formed by a logarithmic spiral [127]. The average angular spread $\Delta\theta_{\mathrm{2D}}(t) = \theta_{\mathrm{2D}}(t) - \theta_0$ of this spiral for a circular expansion is given by

$$\Delta\theta_{\mathrm{2D}}(t) = w_{\mathrm{2D}}(s) \ln\left[1 + \frac{t}{t^*}\right], \tag{4.5}$$

where θ_0 is the initial sector size and $w_{\mathrm{2D}}(s) = v_\perp/v$ is a rescaled sector boundary speed (Fisher wave speed) that depends on the selection coefficient s and is influenced by the strength of the genetic drift. In general, $w_{\mathrm{2D}}(s) \to 0$ as $s \to 0$. In our model, genetic drift is strong and $w_{\mathrm{2D}}(s) \sim s$. This two-dimensional scaling, as discussed in more detail in Chapter 3, can be derived by mapping the sector boundaries to random walks. For our particular model, $w_{\mathrm{2D}}(s) \approx 1.2as/(v\tau_g)$, where $v\tau_g \approx a$ is the distance the front moves in one generation and a is the effective lattice spacing discussed in the previous section (see Chapter 3). The mapping of sector boundaries to random walks allows us to generalize our analysis to sectors with bulges, as well. Indeed, we expect the green sector to bulge out because it will grow out radially faster than the red strain. However, provided the front has some effective line tension and does not roughen, the sector boundaries will still perform biased random walks with some effective rescaled bias speed $w_{\mathrm{2D}}(s) \sim s$ in the presence of bulges. Hence, our analysis of the sector survival probability in Section 4.4 will apply for this case, as well.

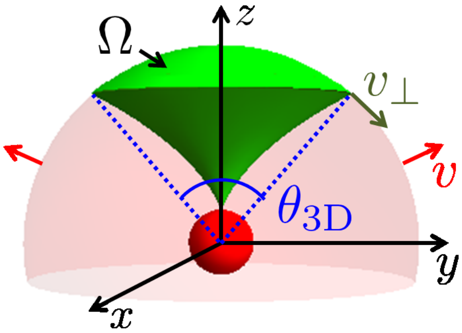


Figure 4.3.1: Schematic of the deterministic motion of a mutant sector arising at the surface of a spherical population of red cells. The green sector invades the red population with a selection-dependent lateral velocity $v_\perp = v w_{\mathrm{3D}}(s)$, where v is the radial front speed. The sector forms a logarithmic cone, formed by rotating a logarithmic spiral curve parameterized by the angle θ_{3D} around the z-axis. In our simulations, we ignore possible bulges in the population front and and neglect the enhanced outward expansion of the green sector relative to the red domain expansion velocity v. Hence, the green sector forms a spherical cap with some solid angle Ω at the spherical population frontier.

We can compute the analogous angle θ_{3D} for a spherical expansion by calculating the solid angle $\Omega = \Omega(t)$ covered by the mutant sector at time t, as shown in Fig. 4.3.1 [127]. Then, the angle of the shape of the edge of the logarithmic cone is given by the relative angle $\Delta\theta_{\mathrm{3D}}(t) = \theta_{\mathrm{3D}} - \theta_0$:

$$\Delta\theta_{\mathrm{3D}}(t) = 2\left[\arccos\left(1 - \frac{\Omega}{2\pi}\right) - \arccos\left(1 - \frac{\Omega_0}{2\pi}\right)\right] = w_{\mathrm{3D}}(s)\ln\left[1 + \frac{t}{t^*}\right], \qquad (4.6)$$

where Ω_0 is the initial solid angle covered by the mutant sector and $w_{\mathrm{3D}}(s) = v_\perp(s)/v$.

We check this result with simulations in Fig. 4.3.2. We find that Eq. (4.6) accurately describes the average mutant sector shape, as shown by the linear fits through the data in the main plot. Note that when $s = 0$, the logarithmic cone sector boundary collapses and the sector boundary approaches a constant solid angle. When many mutations with $s > 0$ are present, these logarithmic cones of mutations can collide to form interesting limiting shapes discussed in detail in Ref. [127] (see also Ref. [130]). Upon using Eq. (4.6) to extract the rescaled lateral Fisher wave velocity $w_{\mathrm{3D}}(s)$, we find $w_{\mathrm{3D}}(s) \approx \sqrt{s}$ to an excellent approximation, as illustrated in the inset of Fig. 4.3.2. Note that this square-root scaling is markedly different from the circular expansion case. Three-dimensional inflating range expansions are thus consistent with the classic Fisher equation result that the wave-speed should approach $v_{\perp} = 2\sqrt{Ds}$ when $\Delta = 0$ in Eq. 4.4 [91]. We expect that the coefficient of the $\sqrt{s}$ term depends on the noise strength, which is fixed in our simulations. Note that $v_{\perp}(s)$ for small s is much bigger than the difference between the mutant and wild-type radial expansion velocities (of order s), thus justifying our neglect of the bulge in three dimensions. This scaling result is consistent with previous studies of noisy Fisher waves in higher dimensions [17, 131].

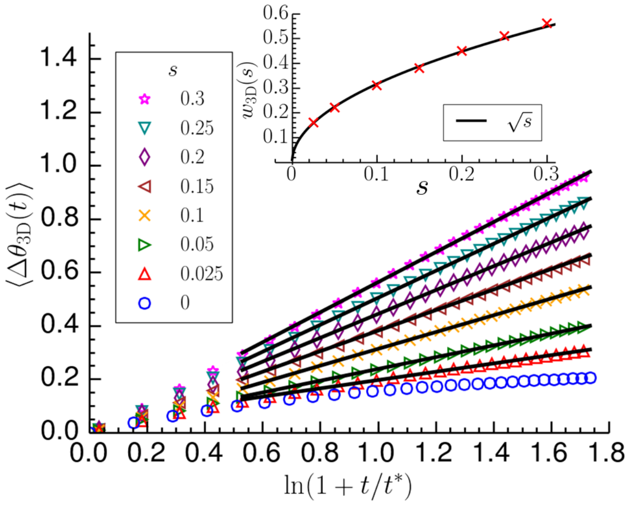


Figure 4.3.2: Simulated mutant clusters in a spherical range expansion ($R_0 = 30$ cell widths) with an initial solid angle $\Omega_0 \approx 0.20$ steradians. We plot the angle $\theta_{3D}(t)$ delineating the edge of the conical shape formed by the mutant cluster edge, averaged over 38400 surviving sectors. After an initial transient, mutant sectors form logarithmic cones for any selective advantage $s > 0$, as illustrated by the linear fits through the data (black lines). In the inset, we plot the slope of these lines, $w_{3D}(s)$. (Note that the case $s = 0$ at the bottom is exceptional.) We find, to an excellent approximation, $w_{3D}(s) \approx \sqrt{s}$.

Having established that even the deterministic spread of a mutant sector is influenced by the genetic drift noise, we now focus on an important, biologically relevant consequence of Eq. (4.4): What is the probability that a single green mutant cell establishes a surviving cluster at long times, given an initial green cell distribution $f(\mathbf{x}, t = 0)$ at the population frontier? In the strong genetic drift limit, Doering et al. mapped Eq. (4.4) in one dimension to a reaction-diffusion model [129] and found an exact result for the survival probability along a flat, infinitely long linear front with initial condition $f(\mathbf{x}, t = 0)$:

$$P[s \geq 0, \Delta; f(\mathbf{x}, t = 0)] = 1 - \exp\left[-\frac{s}{\Delta}\int \mathrm{d}\mathbf{x}\, f(\mathbf{x}, t = 0)\right], \tag{4.7}$$

where the integration is along all positions $\mathbf{x}$ along the front. We expect Eq. (4.7) to hold for three-dimensional expansions, as well, where the integration is now over a two-dimensional surface. This hypothesis was tested for a more sophisticated range expansion model in Ref. [132]. We will show that Eq. (4.7) can be derived via a field theoretic technique in Section 4.5. Note that for $s \ll \Delta^{-1} N_0$ [where $N_0 = \int d\mathbf{x}\, f(\mathbf{x}, t = 0)$ is the initial number of mutated cells], the survival probability vanishes linearly with s: $P \sim s\Delta^{-1} N_0$. In contrast, we will find in Sections 4.4 and 4.5 that for inflating expansions in two and three dimensions, this probability instead approaches a non-zero value as $s \to 0$.

The $s < 0$ case, although not the focus of this chapter, is interesting and subtle. In the deterministic regime, the logarithmic spirals in two dimensions for $s < 0$ are counterrotating, forming a tongue-like, or flower petal shape. Similarly, in three dimensions, the boundary of the logarithmic cone spirals inward instead of outward. Hence, we expect that the boundaries of any surviving mutant cluster will eventually be pinched off due to this deterministic motion for $s < 0$. Our theory predicts that the only way a mutant sector with $s < 0$ can survive at long times is if it wraps all the way around the circumference of the population. Otherwise, the long time survival probability vanishes: $P_\infty(s < 0) = 0$. However, the eventual extinction of a sector might take a very long time, due to the logarithmically slow deterministic dynamics. Since our simulations may only probe finite times, the convergence of the survival probability to its long-time steady-state value is extremely slow for $s < 0$. We briefly examine some of these issues in Appendix C.3.

4.4 Two-Dimensional Range Expansions

Mutant sectors in circular expansions can be treated easily by mapping the sector boundaries to random walks ([13], Chapter 3). Then, for treadmilling expansions ($\Theta = 0$), the random walk theory [3, 9] yields the survival probability

$$P_{\infty}^{\text{treadmill}}(s, a, R_0) = \frac{1 - e^{-sN_0/\Delta}}{1 - e^{-2\pi s R_0/(a\Delta)}}, \tag{4.8}$$

where the finite circumference of the population front is taken into account and individual cells have size a. Note the resemblance to the classic Kimura formula for the fixation probability of a mutation with a selective advantage in a well-mixed population with initial frequency $f_0 = a/(2\pi R_0)$ [11]. When $s \to 0^+$, we find a non-zero survival probability. Upon expanding Eq. (4.8) for small s, we have

$$P_{\infty}^{\text{treadmill}}(s, a, R_0) = \frac{a}{2\pi R_0} + \frac{1}{2}\left[1 - \frac{a}{2\pi R_0}\right]\frac{s}{\Delta} + \mathcal{O}[(s/\Delta)^2] \qquad (s > 0). \tag{4.9}$$

Note that when $s < 0$ in Eq. (4.8), the survival probability is exponentially suppressed,

$$P_{\infty}^{\text{treadmill}}(s, a, R_0) \approx e^{-2\pi |s| R_0/(a\Delta)} \qquad (s < 0) \tag{4.10}$$

for $R_0 \gg a\Delta/|s|$.

Let's now consider a more general $\Theta > 1/2$ case. We set $\Delta = 1$ for simplicity and look at range expansions with a single cell frontier width. Note that $\Delta \approx 1$ in our two-dimensional range expansion simulations. Upon generalizing the analysis of Ref. [11], we find that the mutant sector angle $\theta_{2\mathrm{D}}(t)$ for an arbitrary growth exponent $\Theta > 1/2$

obeys the stochastic differential equation

$$\frac{d\theta_{\mathrm{2D}}}{d\tau} = V_{\mathrm{2D}}(1-\tau)^{\frac{\Theta}{2\Theta-1}} + \sqrt{2D_{\mathrm{2D}}}\,\eta(\tau), \tag{4.11}$$

where $V_{\mathrm{2D}} \approx ast^*/[\tau_g R_0(2\Theta-1)]$ is a selection-dependent bias, $D_{\mathrm{2D}} \approx a^2 t^*/[\tau_g R_0^2(2\Theta-1)]$ is a sector boundary diffusion constant, and τ is a unit-less time given by

$$\tau = 1 - \left(1 + \frac{t}{t^*}\right)^{1-2\Theta}. \tag{4.12}$$

The function $\eta(\tau)$ in Eq. (4.11) is a Gaussian white noise such that $\langle \eta(\tau)\eta(\tau')\rangle = \delta(\tau-\tau')$. Equation (4.11) describes a simple diffusive motion for the sector boundary in the conformal coordinate τ, with a time-dependent bias. Note that for any $\Theta > 1/2$, the entire evolution of the sector is compressed into a finite dimensionless time interval $\tau \in [0,1)$. The consequences of this compression are explored in more detail in Chapter 3 and [109, 133, 134]. For $\Theta \leq 1/2$, the inflation becomes irrelevant and a neutral mutation occurring at a large front ($R_0 \gg \sqrt{D_{\mathrm{2D}} t^*}$), for example, will always eventually go extinct. The $\Theta = \Theta_c = 1/2$ case is marginal and a neutral mutation is barely unable to survive. We will show that the $\Theta = \Theta_c = 1/2$ case is also marginal for three-dimensional expansions in Section 4.5.

The survival probability can be extracted from Eq. (4.11) by analyzing the first-passage probability that the sector angle θ_{2D} vanishes (see Ref. [3] for a review of first-passage processes). Unlike the $\Theta = 0$ case, we will have to work in the limit of large population fronts $R_0 \gg \sqrt{D_{\mathrm{2D}} t^*}$ such that we can ignore any finite size corrections. We will compare our results to simulations and to the exactly soluble $\Theta = 0$ case to check the approximation. Following a similar analysis as in Chapter 3 for the special case

$\Theta = 1$ and in the large front limit, we find a long-time survival probability that depends on two scaling variables $x_{2\text{D}}$ and $\kappa_{2\text{D}}$, given by

$$\begin{cases} x_{2\text{D}} = \dfrac{R_0\phi_0}{a}\sqrt{\dfrac{\tau_g}{t^*}} \xrightarrow[\Theta=1]{} \phi_0\sqrt{\dfrac{R_0 v\tau_g}{a^2}} \\ \kappa_{2\text{D}} = s\sqrt{\dfrac{t^*}{\tau_g}} \xrightarrow[\Theta=1]{} s\sqrt{\dfrac{R_0}{v\tau_g}} \end{cases} . \tag{4.13}$$

where a is the average cell diameter. Note that we can only make the replacement $t^* = R_0/v$ in Eq. (4.13) when $\Theta = 1$. In our simulation model with $\Theta = 1$ (see Section 4.2), we set our length units so that $v\tau_g = 1 \approx a$, where a is an average cell diameter. There is an additional complication, discussed in detail in Chapter 3, related to mapping the simulation model to a random walk. The mapping introduces an additional dimensionless parameter and we must take $s \to \gamma s$ in Eq. (4.13) to match onto our simulations, where $\gamma \approx 1.2$.

The survival probability scaling function can be calculated using an adiabatic approximation, as was done in Chapter 3 for $\Theta = 1$. Although the survival probability depends on many parameters, i.e. $P_{2\text{D}} \equiv P_{2\text{D}}(R_0, \phi_0, t^*, \tau_g, a, \Theta, t)$, we find that the full time-dependent solution for $\Theta > 1/2$ in the adiabatic approximation depends on just two parameter combinations [see Eq. (4.13)] and on the growth exponent Θ:

$$\begin{aligned} P_{2\text{D}} &= P_{2\text{D}}(x_{2\text{D}}, \kappa_{2\text{D}}, \Theta, \tau) \\ &\approx 1 - \int_0^{\tau} \mathrm{d}z \, \frac{x_{2\text{D}}}{2}\sqrt{\frac{2\Theta - 1}{\pi z^3}} \exp\left\{ -\frac{\left[\kappa_{2\text{D}} z (1-z)^{\frac{\Theta}{1-2\Theta}} + x_{2\text{D}}(2\Theta - 1)\right]^2}{4z(2\Theta - 1)} \right\}. \end{aligned} \tag{4.14}$$

We can set $\tau = 1$ (i.e. $t \to \infty$) in Eq. (4.14) to find the ultimate survival probability

$P_\infty(x_{\mathrm{2D}}, \kappa_{\mathrm{2D}}, \Theta > 1/2)$. Although the integral does not appear to be expressible in terms of elementary functions, we numerically integrate it to obtain the scaling function shown in Fig. 4.4.1 for $\Theta = 1$. The predicted scaling form matches our simulation results well. There are no fitting parameters and the lattice spacing $a \approx 0.9 v\tau_g$ can be measured independently (see Chapter 3). We also show the treadmilling $\Theta = 0$ result in Fig. 4.4.1 for a fixed $R_0 = 50a$ to illustrate the large enhancement in the survival probability due to inflation.

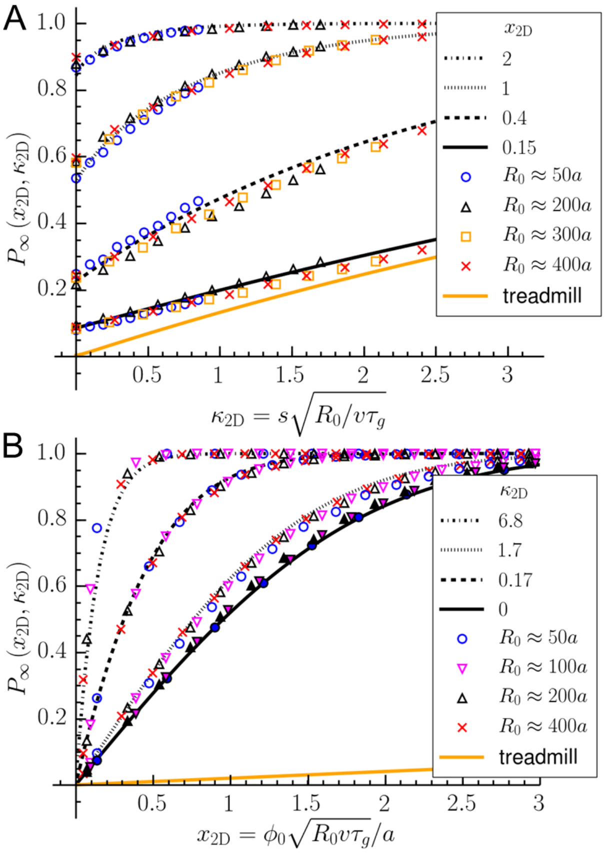


Figure 4.4.1: The long-time ($t \to \infty$, $\tau \to 1$) survival probability scaling function P_∞ of a green mutant sector [as in Fig. 4.1.2A] for circular range expansions with $\Theta = 1$, as a function of the two dimensionless variables x_{2D} and κ_{2D} displayed in Eq. (4.13). Clusters of contiguous mutants subtend an initial angle ϕ_0 along the population front of an otherwise all red population with initial radius R_0. Different symbols lying on the same line correspond to data sets with different R_0. In part (A), ϕ_0 is tuned to yield a fixed x_{2D}. In part (B), ϕ_0 is varied to find the probability at different x_{2D} at a fixed κ_{2D}. The solid symbols on the bottom curve in part (B) correspond to neutral range expansions with $s = 0$. The lines show the analytic prediction using the adiabatic approximation in Eq. (4.14) for $\Theta = 1$ (accurate for $0 < s \ll 1$ and exact for $s = 0$). In simulations (points), P_∞ is estimated from the survival probability at a long (but finite) time $t \gtrsim 1200$ generations. Hence, the survival probability, especially at large R_0 and small s, has not yet converged to P_∞, explaining some of the deviation of the data from the analytic prediction. Solid orange lines show the survival probability for a treadmilling expansion with $R_0 = 50a$. The orange line does go to a non-zero value as $\kappa_{2D} = 0$ in part (A), but it is indistinguishable from zero in the plot. Note the greatly enhanced survival probability caused by inflation in part A, as compared to treadmilling systems.

Note in Fig. 4.4.1A that as we approach a neutral mutation $\kappa_{2D} \to 0$, we find a non-zero survival probability. This is markedly different from the infinite flat front result in Eq. (4.7), which vanishes as $s \to 0$. Indeed, for a single initial mutant cell for with R_0 ($x_{\mathrm{2D}} \approx \sqrt{a/R_0}$), the limit is

$$P_\infty(x_{\mathrm{2D}}, \kappa_{\mathrm{2D}} = 0) = \operatorname{erf}\left[\frac{x_{\mathrm{2D}}}{2}\right] \approx \sqrt{\frac{a}{\pi R_0}}. \tag{4.15}$$

It is instructive to compare this result to the much smaller survival probability $a/(2\pi R_0)$ of a mutation occurring at the surface of a two-dimensional *treadmilling* tumor. Note from Fig. 4.4.1A that the survival probability as $\kappa_{\mathrm{2D}} \to 0$ is much larger for an inflating rather than a treadmilling tumor. Thus, for small selective advantages, inflation plays a much more important role in rescuing the mutation than one might expect from a naive analysis of a mutation sweeping a population with a finite size. We will show in the next section that there is an analogous scaling function $P_{\mathrm{3D}}(x_{\mathrm{3D}}, \kappa_{\mathrm{3D}})$, with appropriate *three-dimensional* analogues of the dimensionless scaling variables x_{3D} and κ_{3D}. We will again find that finite size effects are less important than inflation at small s.

4.5 Three-Dimensional Range Expansions

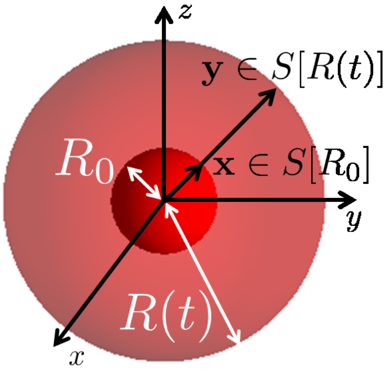


Figure 4.5.1: Coordinate system for studying spherical expansions. At time t when the expansion has radius $R(t)$, we can map any point on the spherical population front, $\mathbf{y} \in S[R(t)]$, to a corresponding point $\mathbf{x} \in S[R_0]$ on the initial population front with radius R_0, along a radial line connecting $\mathbf{y}$ to the center.

A convenient coordinate system for three-dimensional range expansions is the position $\mathbf{x} \in S(R_0)$ on the initial spherical population front $S(R_0)$. Any positions $\mathbf{y} \in S[R(t)]$ on the spherical front $S[R(t)]$ of radius $R(t) = R_0(1+t/t^*)^\Theta$ at any time $t > 0$ can be traced backward to a unique $\mathbf{x} = R_0\mathbf{y}/R(t)$ on the initial front, as shown in Fig. 4.5.1. Thus, we can fully specify any position in the range expansion by a single position $\mathbf{x} \in S(R_0)$ and a time t. We then interpret the evolution of the range expansion as a process acting on the original population front $S[R_0]$ with a superimposed inflation that locally dilates the front perpendicular to the local expansion direction. The corresponding Langevin equation for the green cell density $f \equiv f(\mathbf{x}, t)$ with this inflationary effect reads

$$\partial_t f = \frac{D}{(1+t/t^*)^{2\Theta}} \nabla^2_{\mathbf{x}} f + s\tau_g^{-1} f(1-f) + \sqrt{\frac{2\Delta f\,[1-f]}{\tau_g(1+t/t^*)^{2\Theta}}}\,\xi(\mathbf{x}, t), \tag{4.16}$$

where $\xi(\mathbf{x}, t)$ is again a Gaussian white noise [as in Eq. (4.4), interpreted via the Îto calculus] and t^* is a cross-over time. Equation (4.16) is a non-linear stochastic partial differential equation with a spatial dilation. Such equations are of interest not only in this biological context, but also in cosmology, interface growth, and other areas of condensed matter physics [135]. In the case of linear inflation ($\Theta = 1$), we have $t^* = R_0/v$, where v is the population front speed. In this case, Eq. (4.16) is a straightforward extension of the stochastic differential equation that describes *circular* range expansions in Ref. [13] and in Chapter 3. We now ignore the finite size of the sphere and extend our coordinate $\mathbf{x} \in S(R_0)$ to the infinite two-dimensional space $\mathbb{R}^2$. This approximation will work as long as our sector sizes are small compared to the total surface area of the front. We move to the field-theoretic representation of Eq. (4.16), suitable for application of the methods of path integrals and statistical mechanics, and look for the ultimate survival probability P_∞. For a review of similar ideas from the perspective of population genetics, see Ref. [136]. Readers unfamiliar with these formal developments may wish to pass on to our final conclusions in Sections 4.5.1 and 4.5.2 which have been subjected to extensive numerical checks.

Equation (4.16) can be converted to a field-theoretic action by employing standard methods, pioneered over 40 years ago by Martin, Siggia, and Rose [137], and extended to non-linear stochastic equations such as Eq. (4.4) and Eq. (4.16) by Janssen [138] and de Dominicis [139]. We introduce a response function $\tilde{f} = \tilde{f}(\mathbf{x}, t)$ to account for the noise term in Eq. (4.16). This procedure yields the response functional

$$\begin{aligned} \mathcal{J}[\tilde{f}, f] = \int \mathrm{d}t\, \mathrm{d}^2\mathbf{x} \left\{ \tilde{f} \left[\frac{\partial f}{\partial t} - \frac{D}{(1 + t/t^*)^{2\Theta}} \nabla^2 f - s\tau_g^{-1} f(1 - f) \right.\right. \\ \left.\left. - \frac{\Delta}{\tau_g (1 + t/t^*)^{2\Theta}}\, \tilde{f} f(1 - f) \right] \right\}, \end{aligned} \tag{4.17}$$

which, as shown below, allows us to treat inflationary dynamics on the sphere using the methods of statistical mechanics. The noise term (proportional to Δ) is now on equal footing with the other contributions, making it easier to analyze than in the Langevin equation (4.16). When such functionals are exponentiated, they determine the probability of particular path histories associated with the underlying stochastic differential equation [136, 137, 138, 139]. We introduce an initial density $f_0(\mathbf{x})$ of green cells by adding an additional term to the field theoretic action in Eq. (4.17):

$$-\int \mathrm{d}^2\mathbf{x}\, f_0(\mathbf{x})\, \tilde{f}(\mathbf{x}, t=0) \approx -\int \mathrm{d}^2\mathbf{x}\, N_0\delta(\mathbf{x})\, \tilde{f}(\mathbf{x}, t=0) = -N_0\tilde{f}(\mathbf{0}, 0), \tag{4.18}$$

where the first equality assumes a concentrated aggregation of N_0 cells near the origin at time $t = 0$. For a solid angle Ω_0 of the initial patch, the number of initial cells N_0 is approximately

$$N_0 \approx \frac{\gamma\Omega_0 R_0^2}{a^2}, \tag{4.19}$$

where γ is a non-universal constant related to the surface packing of spheres in the three-dimensional Bennett model.

The response functional (4.17) can be used to calculate the survival probability for the initial patch of N_0 green cells. As shown in Appendix C.1, a variational principle applied to the exponentiated response functional (4.17) yields a spatially uniform, mean field approximation to the response function $\tilde{f}(\mathbf{x}, t) \approx \tilde{f}(t)$. With this mean field approximation, our survival probability at long times T [see Eqs. (C.9)-(C.11) in Appendix C.1] can be obtained by a limiting procedure,

$$P_\infty \approx 1 - \exp\left[\lim_{\zeta\to 0^+}\lim_{T\to\infty} N_0\tilde{f}(t=0)\right], \tag{4.20}$$

where $\tilde{f}(t)$ satisfies

$$\frac{\partial \tilde{f}}{\partial t} = -s\tau_g^{-1}\tilde{f} - \frac{\Delta}{\tau_g(1+t/t^*)^{2\Theta}}\tilde{f}^2 + \zeta\theta(T-t), \tag{4.21}$$

and where $\theta(x)$ is the step function ($\theta(x) = 1$, $x > 0$, $\theta(x) = 0$, $x < 0$). Note that the order of limits in Eq. (4.20) is crucial! We will consider $\Theta = 0$ (treadmilling), $\Theta = 1$ (linear inflation), and $\Theta = 1/2$ (borderline case) in the next subsections.

4.5.1 Treadmilling Fronts ($\Theta = 0$)

When $\Theta = 0$, the range expansion simply turns over cells at its surface, which we take for simplicity to be a one-cell-diameter thick spherical shell with radius R_0. This dynamics is relevant for "treadmilling" tumors. If we ignore the finite size of the shell (a good approximation provided $R_0 \gg a$), the survival probability can be calculated from the solution to the Eq. (4.21) with $\Theta = 0$. This equation becomes a Ricatti equation [4] with constant coefficients. It can be converted to a linear, second-order ordinary differential equation as discussed in Appendix C.2. This equation is soluble and, consistent with Ref. [140], we find a survival probability

$$P_\infty^{\text{treadmill}}[s, N_0, \Delta] \approx 1 - e^{-\frac{N_0}{2\Delta}(|s|+s)} \tag{4.22}$$

$$P_\infty^{\text{treadmill}}[s, N_0, \Delta] \approx \begin{cases} 1 - e^{-N_0 s/\Delta} & s \geq 0 \\ 0 & s < 0 \end{cases}, \tag{4.23}$$

where we have made use of Eq. (4.20). This result is consistent with the generalized formula of Doering et al. in Eq. (4.7). We expect corrections for both positive and negative s when the finite size of the population front is taken into account, as in the treadmilling circle result in Eq. (4.8).

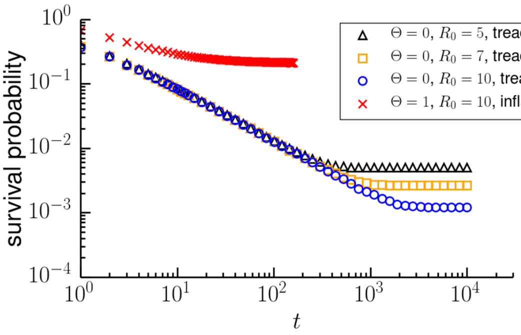


Figure 4.5.2: A comparison of treadmilling ($\Theta = 0$) and linearly inflating ($\Theta = 1$) expansion survival probabilities for single cell, neutral mutations at different initial radii R_0 (in units of the cell diameter a). Note that the survival probability for the treadmilling cases decay much faster with the time t (measured in generations) and saturate at much smaller values. The plateaus at long times are consistent with $P_\infty^{\text{treadmill}} \sim (a/R_0)^2$, as in Eq. (4.24). The survival probability for an inflating front saturates earlier at a much higher value, which we explore in the next subsection. Note the approximately 200-fold enhancement in the survival probability of inflating sectors for $R_0 = 10a$.

The neutral case $s = 0$ is easy to analyze even for a finite size population front such as a treadmilling sphere. We know from studies of voter model dynamics [101, 141] that the descendants of one of the initial cells will eventually sweep the entire population. Hence, the survival probability of a single mutant cell out of a total of N_{tot} cells at the frontier is

$$P_\infty^{\text{treadmill}}[s = 0, N_0 = 1] = \frac{1}{N_{\text{tot}}} \approx \frac{a^2}{4\pi R_0^2} \sim \left(\frac{a}{R_0}\right)^2. \tag{4.24}$$

We shall see in the next subsection that this treadmilling survival probability will be much smaller than the survival probability for an inflating tumor for large initial radii $R_0 \gg a$. This expectation is already evident upon comparing a treadmilling and an inflating front in our simulations. Note the dramatic qualitative difference in the survival probabilities with increasing time for $R_0/a = 10$ (more than two orders of magnitude!)

in Fig. 4.5.2.

4.5.2 Linearly Inflating Fronts ($\Theta = 1$)

We now solve Eq. (4.21) with $\Theta = 1$. The details are discussed in Appendix C.2, where we find

$$\tilde{f}(t=0) = -\frac{1}{\Delta\tau_g^{-1}t^*[1 + s\tau_g^{-1}t^* e^{s\tau_g^{-1}t^*}\mathrm{Ei}(-s\tau_g^{-1}t^*)]}, \tag{4.25}$$

where $\mathrm{Ei}(x)$ is the exponential integral function [4]. Hence, the survival probability from (4.20) is

$$\begin{aligned} P_\infty(N_0, s, \Delta, t^*, \tau_g) &= 1 - \exp\left[-\frac{N_0\theta(s)}{\Delta\tau_g^{-1}t^*[1 + s\tau_g^{-1}t^* e^{s\tau_g^{-1}t^*}\mathrm{Ei}(-s\tau_g^{-1}t^*)]}\right] && (4.26)\\ &= P_\infty\left(x_{\mathrm{3D}} = \frac{N_0\tau_g}{\Delta t^*}, \kappa_{\mathrm{3D}} = s\tau_g^{-1}t^*\right) && (4.27)\\ &= 1 - \exp\left[-\frac{x_{\mathrm{3D}}\theta(\kappa_{\mathrm{3D}})}{1 + \kappa_{\mathrm{3D}}e^{\kappa_{\mathrm{3D}}}\mathrm{Ei}(-\kappa_{\mathrm{3D}})}\right], && (4.28) \end{aligned}$$

where $t^* = R_0/v$ as usual is the time to inflate to twice the initial radius and we have identified the two key dimensionless parameters x_{3D} and κ_{3D}:

$$\begin{cases} x_{\mathrm{3D}} = \dfrac{N_0\tau_g}{\Delta t^*} = \dfrac{N_0 v\tau_g}{\Delta R_0} \\[2ex] \kappa_{\mathrm{3D}} = \dfrac{st^*}{\tau_g} = \dfrac{sR_0}{v\tau_g} \end{cases}, \tag{4.29}$$

where N_0 is the initial number of mutant cells. Note the similarity between Eq. (4.29) and the analogous dimensionless variables for circular inflation in Eq. (4.13). Finally, note that $P_\infty = 0$ when $s < 0$. Hence, since $P_\infty > 0$ for $s = 0$, there is a jump discontinuity in the survival probability at the origin. This discontinuity occurs because we have assumed that our inflating population front is effectively infinite relative to the size of the inflating sector. Hence, when $s < 0$, a mutant sector will always be able to shrink back deterministically to a small enough size that it can be extinguished via genetic drift. This may happen very slowly, however, because the deterministic motion of the sector is logarithmic in time, as discussed in Section 4.3. We also expect that, similarly to the two-dimensional treadmilling result in Eq. (4.8), corrections to the ultimate survival probability in the $s < 0$ regime are exponentially suppressed by the population front size. Evaluation of the (small!) survival probability for negative s for finite-sized fronts is a subject for future investigation.

It is instructive to study biologically relevant limits of Eq. (4.28). For example, if we let $t^* \to \infty$ so that inflation becomes negligible, we recover the linear front ($\Theta = 0$) result, consistent with Eq. (4.7) and Eq. (4.23):

$$P_\infty(x_{3\mathrm{D}}, \kappa_{3\mathrm{D}} \gg 1) = 1 - \exp\left[-x_{3\mathrm{D}}\kappa_{3\mathrm{D}} + \mathcal{O}(x_{3\mathrm{D}})\right] = 1 - e^{-sN_0/\Delta}. \tag{4.30}$$

Another important limit is of course $s \to 0^+$ (so that $\kappa_{3\mathrm{D}} \to 0^+$). We find

$$P_\infty(x_{3\mathrm{D}}, \kappa_{3\mathrm{D}} = 0) \approx 1 - \exp\left[-x_{3\mathrm{D}}\right] = 1 - e^{-\frac{N_0\tau_g}{t^*\Delta}}. \tag{4.31}$$

Note that in our simulations, $t^*/\tau_g \approx R_0/a$. Thus, for a single mutation ($N_0 = 1$) and a small selection coefficient, we find from (4.28) a linear increase in the survival probability as a function of $s > 0$:

$$P_\infty\left(N_0 = 1, R_0, s \ll 1\right) \approx \frac{a}{\Delta R_0} - \ln\left[e^{\gamma_E}\frac{R_0}{a}s\right]\frac{s}{\Delta} + \mathcal{O}(s^2 \ln^2 s), \tag{4.32}$$

where $\gamma_E \approx 0.577$ is Euler's constant. In our simulations, the effective population size $N = \mathcal{O}(1)$, so $\Delta \sim 1/N$ is of order unity as well. Note that the inverse scaling of the zeroeth order term $[\mathcal{O}(s^0)]$ in Eq. (4.32) with the initial population radius is markedly different from the treadmilling result in Eq. (4.24). Indeed, the inflation will rescue the mutation with much higher probability than the probability the mutation fixes in a treadmilling expansion due to the finite tumor size. Also, the logarithm in the first order term $[\mathcal{O}(s)]$ generates a diverging slope in P_∞ as we send $s \to 0$. Note that this slope becomes finite for a flat front ($t^* \to \infty$).

We check the approximation leading to our results with simulations in Fig. 4.5.3. There are no fitting parameters, except the genetic drift strength Δ. We find $\Delta \approx 0.6$, confirming our expectation that the genetic drift strength Δ is of order unity in our simulations. The good agreement between the simulation results and the field theoretic solution justifies our approximation of an infinite population front. Indeed, from our scaling analysis in the previous section, we expect these particular finite size corrections to be small. We are inherently limited by our finite simulation time, which allows us to only simulate clusters which have grown to radii of at most 170 cell diameters.

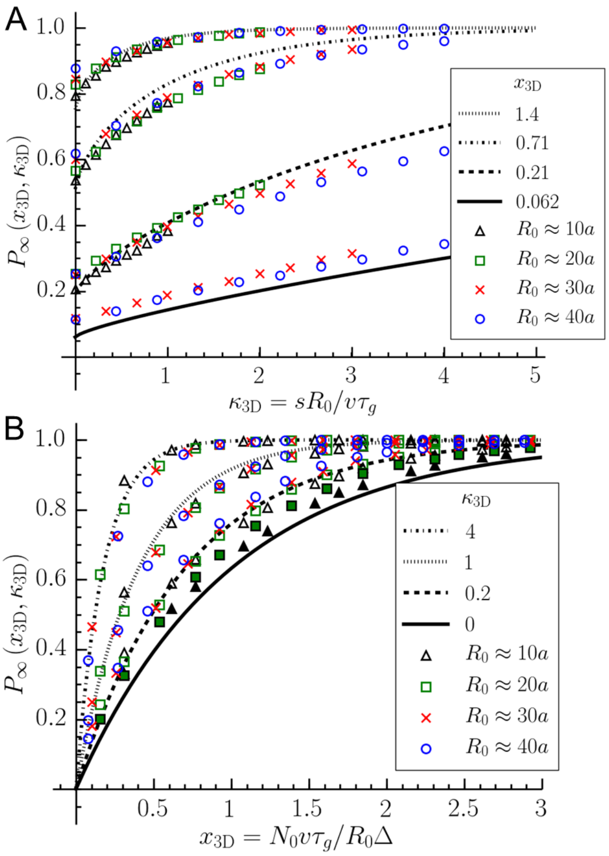


Figure 4.5.3: The survival probability scaling function of a green mutant sector like that in Fig. 4.1.2B for spherical range expansions. Clusters of mutants with N_0 initial cells start on the periphery of an all red population with initial radius R_0 cell widths. Different symbols lying on the same line correspond to data sets with different N_0 and R_0, but a fixed x_{3D} in part (A) and a fixed κ_{3D} in part (B). All range expansions were evolved up to a total cell cluster with a 170 cell diameter radius. The solid symbols in part (B) correspond to neutral range expansions with $s = 0$. The lines show the analytic prediction in Eq. (4.26). Note that at large values of R_0 and small values of κ_{3D} in part (A), the survival probability calculated in the simulations has not yet converged to the steady-state value, so the data points lie above the analytic prediction (lines).

4.5.3 Marginally Inflating Fronts ($\Theta = 1/2$)

It is of interest to find the critical value $\Theta = \Theta_c$ at which inflation is *barely* unable to rescue a neutral mutation from a very low survival probability due to the finite population frontier size. The neutral dynamics associated with $s \to 0^+$ belong to the voter model universality class. Hence, Θ_c is easy to calculate because we know that in this case the average cluster size, $\sqrt{\langle X(t)^2 \rangle}$ (where $X(t)$ is the linear extent of the mutant cluster), scales according to $\sqrt{\langle X(t)^2 \rangle} \sim t^{1/z}$, where $z = 2$ is the dynamical critical exponent of the voter model class [142]. Hence, the linear size of the cluster spreads diffusively and we expect that $\Theta_c = 1/2$. Equation (4.21) turns out to be soluble in this case and we now find a survival probability that reads

$$P_{\infty}^{\Theta=1/2}(x_{3\mathrm{D}}, \kappa_{3\mathrm{D}}) = 1 - \exp\left[\frac{x_{3\mathrm{D}}\theta(\kappa_{3\mathrm{D}})}{e^{\kappa_{3\mathrm{D}}}\mathrm{Ei}(-\kappa_{3\mathrm{D}})}\right], \tag{4.33}$$

where $\theta(\kappa_{3\mathrm{D}}) = \theta(s)$ is again the Heaviside step function. This result also reduces to the linear front solution when $t^* \to \infty$. However, unlike linear inflation, as $\kappa_{3\mathrm{D}} \to 0$, we find the peculiar behavior

$$\begin{aligned} P_{\infty}^{\Theta=1/2}(x_{3\mathrm{D}}, \kappa_{3\mathrm{D}} \ll 1) &\approx 1 - \exp\left[\frac{x_{3\mathrm{D}}\theta(\kappa_{3\mathrm{D}})}{\ln(\kappa_{3\mathrm{D}} e^{\gamma_E})}\right] && (4.34) \\ &\approx \frac{x_{3\mathrm{D}}\theta(\kappa_{3\mathrm{D}})}{\ln(\kappa_{3\mathrm{D}} e^{\gamma_E})}, && (4.35) \end{aligned}$$

where we assume $\kappa_{3\mathrm{D}} \ll e^{-x_{3\mathrm{D}}}$ in the second line. This probability does, in fact, vanish as $\kappa_{3\mathrm{D}} \to 0$. However, it does this logarithmically slowly! We plot this unusual marginal scaling function in Fig. 4.5.4.

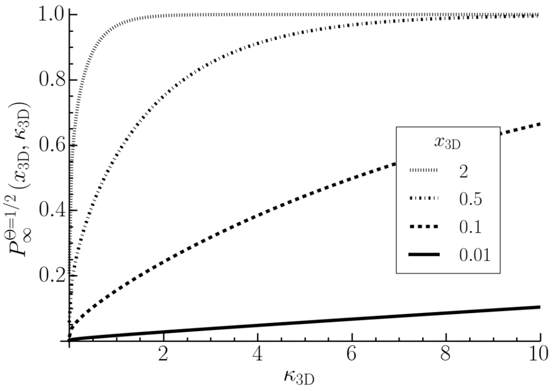


Figure 4.5.4: The survival probability for a mutation living on the surface of a spherical expansion with a $\Theta = 1/2$ growth exponent [eq. (4.1)]. The survival probability vanishes as $\kappa_{\rm 3D} \to 0$, but logarithmically slowly.

4.5.4 Rough Fronts

We focus here just on the case $\Theta = 1$ and discuss first the fate of neutral mutations. If a circular or spherical range expansion has a negligible surface tension at the frontier, then the population front will roughen over time as the cells stochastically divide and push each into virgin territory. This roughening can affect the evolutionary dynamics. For example, in two-dimensional expansions, the motion of genetic sector boundaries becomes super-diffusive [9]. Also, when selection is present, the genetic sectors of the selectively advantageous strain will create bulges, modifying the interface roughening itself [94]. However, even with rough fronts, we expect inflation to act in a similar way. If the average population radius still grows linearly in time, $R(t) \sim t$, then any characteristic length (e.g., the size of a genetic sector) will get inflated linearly in time, just as in the case of smooth fronts. We expect inflation to rescue a mutation from genetic drift if it survives until the crossover time $t^* = R_0/v$. Hence, for a neutral mutation, if the survival probability decays according to $P(t) \sim t^{-\beta}$ at early times, then the ultimate survival probability will scale according to $P_\infty \sim P(t^*) \sim (a/R_0)^\beta$. For rough fronts in two-dimensions, we expect $\beta = 2/3$ (compare to $\beta = 1/2$ for uniform fronts), so the decay is more rapid. In two dimensions, these scaling arguments for neutral mutations can be made rigorous and a precise mapping between expansions with inflating and linear rough fronts can be made by exploiting the mapping between sector boundaries and random walks [109].

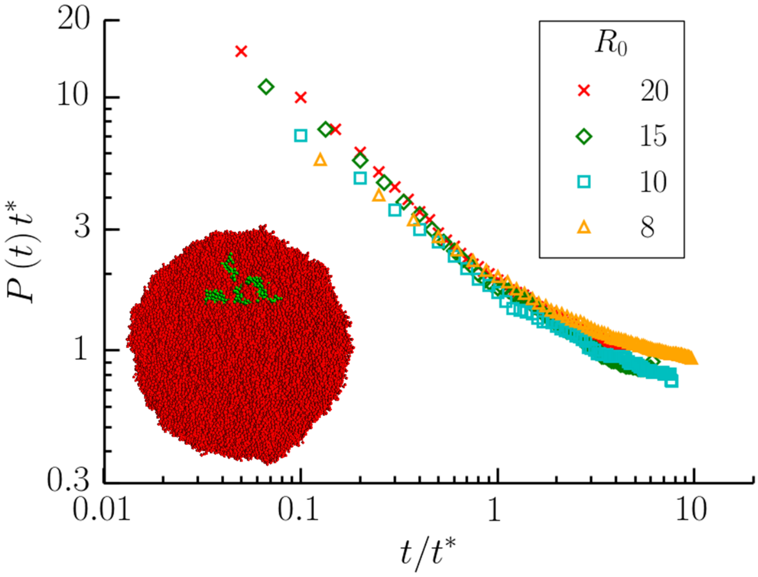


Figure 4.5.5: The survival probability $P(t)$ of a cluster formed from a single neutral green cell at an initially uniform front of a spherical population with radius R_0. The spherical population grows out with a rough front due to the stochastic birth dynamics. The survival probabilities for different R_0 are collapsed using a simple scaling with the cross-over time $t^* = R_0/v$, where $v = 1$ in the simulations. Note that for $t \gg t^*$, the data is consistent with survival probabilities that approach a constant such that $P_\infty = P(t \to \infty) \propto 1/t^* \sim a/R_0$. The inset has an example of a surviving cluster for $R_0 = 10$ cells after about 30 generation times to $R(t = 30) \approx 40$.

In three dimensions, the situation is more complicated because the universal properties of the interface roughening are still being investigated and the coupling of the roughness to the genetic sector dynamics is not obvious. An example of a fragmented, surviving green cluster is shown in the inset of Fig. 4.5.5. Despite these complications, the simulation results in Fig. 4.5.5 suggest that the coupling between frontier roughness and sector dynamics is weak. At short times $t \ll t^*$, the survival probability of a neutral mutation decays approximately as it would in a linear (non-inflating), uniform front, i.e., $P(t) \sim t^{-1}$ (neglecting logarithmic corrections), with perhaps a small correction to the exponent. Then, for $t \gg t^*$, inflation prevents the extinction of all surviving mutations and we find an enhanced survival probability $P_\infty \propto 1/t^* \sim a/R_0$ [compared to the treadmilling result $P_\infty \sim a^2/(4\pi R_0^2)$], just as for the uniform front case. Our simulations are indeed consistent with neutral mutations are rescued with a probability inversely proportional to the initial population radius. These findings are illustrated by the data collapse in Fig. 4.5.5.

What about selection? Anderson and Hallatschek have investigated the two-dimensional case using simulations and experiments, and have found properties of the survival probability scaling function by mapping the genetic sector boundary dynamics at the rough frontier to superdiffusive random walks [143]. In three dimensions, we can do a similar analysis as in the uniform front and find P_∞ for various R_0 and initial green cell numbers N_0. However, our rough front models generate sphere clusters with radii of at most 80 cell diameters. Hence, it is computationally challenging to get the survival probability to converge to a steady-state value for small s and large R_0. Our limited attempts suggest that $\kappa_{3\mathrm{D}} = sR_0/v\tau_g$ is once again an appropriate scaling combination to perform a data collapse for a fixed $x_{3\mathrm{D}} = N_0 v\tau_g/R_0\Delta$. However, a more thorough investigation is necessary and an analytic form for the conjectured scaling function is an open question.

4.6 Conclusions

We tried to show how geometry plays a crucial role in the evolutionary dynamics of spherical range expansions, just as it does for circular ones. Our focus has been on linearly inflating and treadmilling range expansions, both of which are relevant for tumor evolution. Inflating population fronts yield an enhanced survival probability of mutations at range expansion frontiers in two and three dimensions. We have shown how this survival probability scales with the population front radius R_0 when an isolated mutation occurs. The long time survival probability of a mutation decreases as $P_\infty \sim \sqrt{a/R_0}$ in two dimensions and $P_\infty \sim a/R_0$ in three dimensions. The corresponding probabilities for *treadmilling* expansions are instead $P_\infty \sim a/R_0$ and $P_\infty \sim (a/R_0)^2$ for two and three dimensions, respectively. Hence, inflation rescues mutations with a much higher probability than might be expected if we only consider the finite size of the initial population front. For smooth frontiers, we have also gone beyond simple scaling and calculated analytically the scaling function for the survival probability in three dimensions and recaptured some previous results in Chapter 3 for this function in the two-dimensional case, generalizing them to an arbitrary power-law growth of the population radius in time [Eq. (4.1)]. Finally, we looked at the fascinating case of marginal inflation where the population radius increase is just barely unable to rescue a neutral mutation from extinction via genetic drift.

Although a detailed investigation is beyond the scope of this chapter, we have also discussed range expansions with undulating frontiers. Rough frontiers for spherical expansions only change the effect of inflation on the survival probability slightly; larger changes occur in two dimensions. However, simulations of larger expansions would be necessary to confirm this. An interesting avenue for future research, initiated by Anderson and Hallatschek [143], is to calculate scaling functions for the survival probability of mutations on a rough population front with a non-zero selective advantage $s > 0$ in either two or three dimensions. It would also be interesting to study the $s < 0$ regime in more detail. Since the convergence of the survival probability to its infinite time value is so slow in this regime (see Appendix C.3), a time-dependent theory would be necessary to find the survival probability even at very long times.

Chapter 5

Asymmetric Mutualism in Two- and Three-Dimensional Range Expansions

This chapter appears in its entirety in M.O. Lavrentovich & D.R. Nelson, *Phys. Rev. Lett.* **112**, 138102 (2014)

5.1 Introduction

When a population colonizes new territory, the abundance of unexploited resources allows the descendants of the first few settlers to thrive. These descendants invade the new territory and form genetically distinct regions or sectors at the population frontier. If the frontier population is small, the birth and death of individuals create large fluctuations in the sector sizes. These fluctuations, called genetic drift, cause some settler lineages to become extinct as neighboring sectors engulf their territory. Over time, this sector coarsening process dramatically decreases genetic diversity at the frontier [13] (also see Chapter 3).

Interactions between the organisms can modify the coarsening process. For example, cooperative interactions, in which genetic variants in close proximity confer growth benefits upon each other, can lead to the founders' progeny remaining intermingled. Then, coarsening does not occur, and the consequent growth pattern is called a "mutualistic phase" [24]. Cooperative interactions are commonly found in nature: microbial strains exchange resources [144], ants protect aphids in exchange for food [145], and different species of mammals share territory to increase foraging efficiency [146]. Recently, a mutualistic phase was experimentally realized in partner yeast strains that feed each other [22, 23]. These experiments require an understanding of *asymmetric* interactions where species do not benefit equally from cooperation. Antagonistic interactions may also occur, e.g., between bacterial strains secreting antibiotics against competing strains [147]. These interactions and the mutualistic phase also play prominent roles in theories of nonequilibrium statistical dynamics [24, 95, 148, 149, 150, 151, 152].

We explore here asymmetric cooperative and antagonistic interactions in two- and three-dimensional range expansions. Two-dimensional expansions ($d = 1 + 1$) occur when the population grows in a thin layer, such as in a biofilm or on a Petri dish. Three-dimensional expansions ($d = 2 + 1$) occur, for example, at the boundaries of growing avascular tumors [125, 153, 154]. We model both flat and rough interfaces at the frontier, the latter being an important feature of many microbial expansions [2].

We arrive at the following biologically relevant results: Three-dimensional range expansions support mutualism more readily than planar ones, and a mutualistic phase occurs for *any* symmetric cooperative benefit. Conversely, two-dimensional expansions require a critical benefit [24, 148]. In addition, we find that the frontier roughness is strongly enhanced at the onset of mutualism for asymmetric interactions. Finally, we find that frontier roughness allows for a mutualistic phase over a wider range of cooperative benefits.

5.2 Flat Front Models

We consider two genetic variants, labelled black and white. Cells divide only at the population frontier [13]. Such expansions occur when nutrients are absorbed before they can diffuse into the interior of the population, inhibiting cell growth behind the population front. This can occur in tumor growth [125] and in microbial expansions at low nutrient concentrations (see Chapter 2).

According to a continuum version of a stepping stone model [24, 88], the coarse-grained fraction of black cells $f \equiv f(\mathbf{x}, t)$ at some position $\mathbf{x}$ along a *flat* population front at time t obeys

$$\partial_t f = D\nabla^2 f + \tau_g^{-1} f(1-f)\left[s\left(\frac{1}{2} - f\right) + \frac{r}{2}\right] + \eta, \tag{5.1}$$

where D is a diffusivity, $\eta(\mathbf{x}, t)$ an Îto noise term [13] with $\langle \eta(\mathbf{x}, t)\eta(\mathbf{x}', t)\rangle = 2\ell^{d_s}\tau_g^{-1} f(1-f)\delta(\mathbf{x}-\mathbf{x}')\delta(t-t')$, τ_g a generation time, d_s the spatial dimension, and ℓ an effective lattice spacing. Also, $r = \alpha - \beta$ and $s = \alpha + \beta$, where α and β represent the increase in growth rates over the base rate per generation of the black and white species, respectively, in the presence of the other species. Equation (5.1) describes the behavior of two-dimensional expansions of mutualistic strains of yeast [22] and is expected to characterize many different range expansions [13, 24]. At $r = s = 0$, Eq. (5.1) reduces to the Langevin equation of the voter model [151, 152]. A signature of the mutualistic phase is a non-zero average density $\langle \rho \rangle$ of black and white cell neighbors (genetic sector interfaces) at the front at long times. To construct phase diagrams for our different models in Fig. 5.2.1, we calculate $\langle \rho \rangle$ as a function of the cooperative benefits $\alpha, \beta \geq 0$ and antagonistic interactions $\alpha, \beta < 0$.

We propose a microscopic model for flat fronts in the spirit of Grassberger's cellular automaton [155], which obeys Eq. (5.1) under an appropriate coarse-graining, as we verify in the following. Domain wall branching, shown in Fig. 5.2.2(a), is required for a mutualistic phase in $d = 1 + 1$ dimensions. Our model allows branching by having three cells compete to divide into a new spot on the frontier during each time step. To approximate cell rearrangement at the frontier, we assume that the order of competing cells in each triplet is irrelevant. The update rules are:

$$\begin{cases} p\begin{pmatrix} \circ\,\bullet\,\circ \\ \Downarrow \\ \bullet \end{pmatrix} = p\begin{pmatrix} \bullet\,\circ\,\circ \\ \Downarrow \\ \bullet \end{pmatrix} = p\begin{pmatrix} \circ\,\circ\,\bullet \\ \Downarrow \\ \bullet \end{pmatrix} = \dfrac{1}{3} + \alpha \\ p\begin{pmatrix} \bullet\,\circ\,\bullet \\ \Downarrow \\ \circ \end{pmatrix} = p\begin{pmatrix} \circ\,\bullet\,\bullet \\ \Downarrow \\ \circ \end{pmatrix} = p\begin{pmatrix} \bullet\,\bullet\,\circ \\ \Downarrow \\ \circ \end{pmatrix} = \dfrac{1}{3} + \beta \\ p\begin{pmatrix} \bullet\,\bullet\,\bullet \\ \Downarrow \\ \bullet \end{pmatrix} = p\begin{pmatrix} \circ\,\circ\,\circ \\ \Downarrow \\ \circ \end{pmatrix} = 1, \end{cases} \tag{5.2}$$

where $-1/2 \leq \alpha, \beta \leq 2/3$. The rules for all other combinations follow from probability conservation.

Positive α (β) biases the propagation of a black (white) cell into the next generation, due to beneficial goods (e.g. an amino acid in short supply by the partner species [22]) generated by two nearby cells of the opposite type. Negative α and β represent the effect of cells inhibiting the growth of neighboring competing variants.

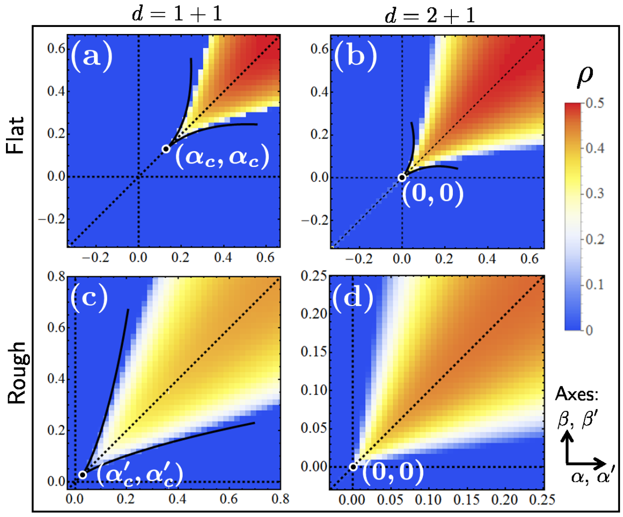


Figure 5.2.1: (Color online) The density $\langle \rho \rangle$ of genetic sector interfaces along the population front at long times t (measured in generations), averaged over 10^3 runs of the flat [(a) and (b)] and rough [(c) and (d)] front range expansion models for $d = 1+1$ [(a) and (c)] and $d = 2+1$ [(b) and (d)] (system sizes L and L^2 cells, respectively) for cooperative benefits α and β defined in Eq. (5.2) [α' and β' for rough fronts in Eq. (5.5)]. The parameters are: (a) $t = 1.5 \times 10^5$, $L = 8 \times 10^3$; (b) $t = 2 \times 10^3$, $L^2 = 500^2$; (c) $t = 2.5 \times 10^3$, $L = 10^3$; (d) $t = 500$, $L^2 = 50^2$; There is a mutualistic phase in the $\alpha, \beta > 0$ ($\alpha', \beta' > 0$) quadrant in all panels. A DP2 point occurs at $\alpha = \beta = \alpha_c = 0.1242(5)$ in (a) and a "rough DP2" point at $\alpha' = \beta' = \alpha'_c = 0.0277(2)$ in (c). The solid lines in (a), (b) and (c) show the DP transition line shapes near the bicritical points (see Appendix D.1). The dotted lines indicate the loci $\alpha = 0$, $\beta = 0$, and $\alpha = \beta$.

For $d = 1+1$, the model is implemented on a square lattice (with one space and one time direction) with periodic boundary conditions in the spatial direction. During each generation (one lattice row along the spatial direction), the states of all triplets of adjacent cells are used to determine the state of the middle cell in the next generation using Eq. (5.2). For $d = 2+1$, we stack triangular lattices of cells (representing successive generations) in a hexagonal close-packed three dimensional array. Each cell sits on top of a pocket provided by three cells in the previous generation, so Eq. (5.2) generalizes immediately (see Appendix D.1).

These simple flat front models generate the rich phase diagrams of Fig. 5.2.1(a) and (b). The $d = 1 + 1$ diagram in Fig. 5.2.1(a) resembles the stepping stone model result [24]. One feature is a DP2 point, located at $(\alpha, \beta) = (\alpha_c, \alpha_c) \approx (0.1242, 0.1242)$ in our model. Applying a symmetry-breaking coefficient $r \equiv \alpha - \beta \neq 0$ biases the formation of either black ($r > 0$) or white ($r < 0$) cell domains, and the DP2 transition crosses over to DP transitions along a symmetric pair of critical lines $s_c(r) \equiv \alpha_c(r) + \beta_c(r)$ for $r < 0$ and $r > 0$. As in typical cross-over phenomena [156], the phase boundaries near the DP2 point are given by $r \sim \pm[s_c(r) - s_c(0)]^{\phi}$, where $s_c(0) = 2\alpha_c \approx 0.2484$ and ϕ is a cross-over exponent [157]. We find $\phi \approx 1.9(1)$, consistent with studies of related models [158] (see Appendix D.1). Hence, we confirm that our model is in the same universality class as Eq. (5.1). Thus, many features of our nonequilibrium dynamical models near the transition lines (e.g., the power laws governing the phase boundary shape) will also appear in the various range expansions describable by Eq. (5.1).

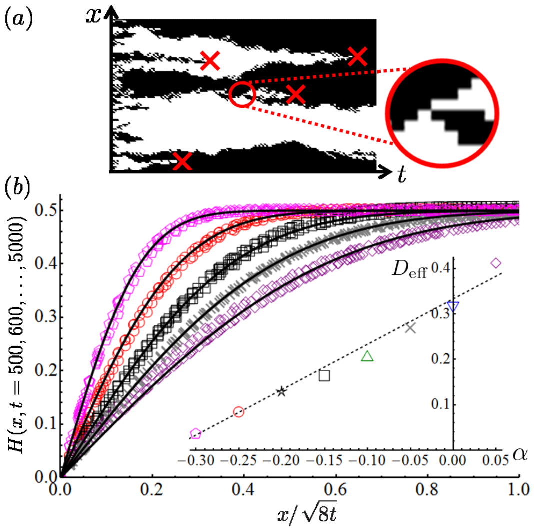


Figure 5.2.2: (Color online) (a) A sample evolution of the flat front model for $d = 1+1$ at $\alpha = \beta = 0$ (front size $L = 100$). The rules in Eq. (5.2) allow interface branching (circled event), facilitating mutualistic mixing. The coarse-grained dynamics for $\alpha = \beta < \alpha_c$ are dominated by pair annihilations of interfaces (marked by crosses) [149, 150]. (b) We collapse the heterozygosity $H(x, t)$, averaged over 1600 runs with $L = 10^3$, at different times (points in main plot) and fit to Eq. (5.3) (solid lines) to find D_{eff} in the inset. The dashed line indicates the prediction $D_{\text{eff}} = 1/3 + \alpha$ (see Appendix D.1).

We now study the approach to the DP2 point along the $\alpha = \beta$ line for $\alpha < \alpha_c$. As α increases from $-1/3$ to $\alpha_c > 0$, domain boundaries between white and black sectors diffuse more vigorously. To check that the entire line $-1/3 < \alpha < \alpha_c$ with $\alpha = \beta$ is dominated at long wavelengths by the annihilation of domain wall pairs [see Fig. 5.2.2(a)], we study the heterozygosity correlation function $H(x, t) = \langle f(x+y, t)[1-f(y, t)] + f(y, t)[1-f(x+y, t)]\rangle$, where $\langle \ldots \rangle$ is an ensemble average and an average over points y along the front [13]. For a random initial condition of black and white cells in equal proportion, $H(x, t)$ can be fit to

$$H(x, t) = \frac{1}{2}\,\text{erf}\left[\frac{x}{\sqrt{8 D_{\text{eff}} t}}\right], \tag{5.3}$$

where the fitting parameter D_{eff} is the effective diffusivity of the domain walls [13]. The dependence of D_{eff} on α away from the DP2 point is consistent with a simple random walk model of domain walls (see Appendix D.1), which predicts $D_{\text{eff}} \approx 1/3 + \alpha$ [inset of Fig. 5.2.2(b)]. As we approach the DP2 point ($\alpha \to \alpha_c^-$) and domain wall branching becomes important, we observe violations of Eq. (5.3), consistent with field theoretic studies [149, 150].

When $d = 2 + 1$, two-dimensional domains at the voter model point $\alpha = \beta = 0$ lack a surface tension and readily "dissolve" [16, 152]. Our simulations show that these dynamics allow for a mutualistic phase for all $\alpha = \beta > 0$, with a remarkable pinning of the corner of the "wedge" of mutualism in Fig. 5.2.1(b) to the origin. A similar phenomenon occurs in branching and annihilating random walks, where an active phase exists for any non-zero branching rate for $d = 2 + 1$ [150]. However, our model is equivalent to the random walk model only for $d = 1 + 1$ [151], and the potential connection in higher dimensions is subtle. We now describe how we find the shape of the mutualistic wedge for $d = 2 + 1$.

When $r = 0$, we find a voter model transition at $s = s_c(r = 0) = 0$, and *any* $s > 0$ pushes the system into a mutualistic phase with a non-zero steady-state domain interface density. A perturbation $r \neq 0$ pushes the system away from the voter model class by suppressing interface formation and induces a DP transition at some $s_c(r) > 0$. Upon exploiting cross-over results for a similar perturbation in Ref. [140], we find phase boundaries given by $r \sim \pm\, s_c(r)/\ln[s_c(r)/s_0]$, where $s_0 \approx 0.551$ is a non-universal constant found by fitting. The resulting curves, plotted in Fig. 5.2.1(b), agree well with simulations (see Appendix D.1).

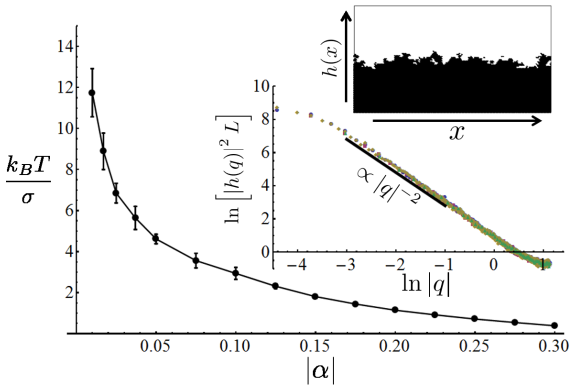


Figure 5.2.3: (Color online) The main plot shows the effective, dimensionless inverse line tension k_BT/σ in $d = 2 + 1$ dimensions [see Eq. (5.4)] for negative α (line guides the eye). Results are from the final interface position $h(x)$ in simulations of an initially flat interface of length $L = 512$ evolved for $t \gtrsim 7000$ generations. The interface has overhangs and holes, and $h(x)$ is the average position for each x. The upper right inset is a sample interface for $L = 128$ and $t = 500$. In the remaining inset, we confirm that Eq. (5.4) correctly predicts the scaling with L by collapsing the Fourier-transformed height $h(q)$, averaged over 160 runs, for $L = 128$, 256, 512, 1024.

When $\alpha = \beta < 0$ for $d = 2 + 1$, we find dynamics similar to a kinetic Ising model with a non-conserved order parameter quenched below its critical temperature with an interface density decay $\langle\rho\rangle \sim t^{-1/2}$ [16] (see Appendix D.1). A local "poisoning" effect penalizes domain wall deformations, creating an effective line tension σ between domains. To find σ, we evolve initially flat interfaces of length L to an approximate steady-state. Fluctuations in the interface position $h(x)$ are characterized by its Fourier transform $h(q)$. Upon averaging over many realizations, we expect that, in analogy with capillary wave theory [156],

$$\left\langle|h(q)|^2\right\rangle = \frac{k_BT}{\sigma L q^2}, \tag{5.4}$$

where T is an effective temperature. Figure 5.2.3 shows that the dimensionless line tension σ/k_BT increases as α becomes more negative and that Eq. (5.4) gives the correct prediction for $\langle|h(q)|^2\rangle$. As we approach the voter model point ($\alpha \to 0^-$), σ/k_BT vanishes with an apparent power law $\sigma/k_BT \sim |\alpha|^{0.61}$. However, models with stronger noise might have a voter-model-like coarsening for $\alpha = \beta < 0$, instead [159].

5.3 Rough Front Models

We model range expansions with rough frontiers using a modified Eden model which tracks cells with at least one empty nearest or next-nearest neighbor lattice site [94]. Each such "active" cell i has a birth rate

$$b_i = \frac{1}{3} + \alpha' N_w(i) \qquad \text{or} \qquad b_i = \frac{1}{3} + \beta' N_b(i), \tag{5.5}$$

if the cell is black or white, respectively. We set the background birth rates (i.e. for populations of all-black or all-white cells) to $1/3$ to make contact with a neutral flat front model. $N_{b,w}(i)$ denote the number of black and white nearest neighbors of cell i, respectively.

At each time step, we pick an active cell i to divide into an adjacent, empty lattice site with probability b_i/b_{tot}, where b_{tot} is the sum of the active cell birth rates. For $d = 1+1$, cells can divide into next-nearest neighbor spots to allow for domain boundary branching (see Appendix D.2). When computing quantities such as the interface density, we wait for the undulating front to pass and then take straight *cuts* through the population parallel to the initial inoculation. The distance of the cuts from the initial inoculation defines our time coordinate.

At the voter model point $\alpha' = \beta' = 0$, the roughness of the front is insensitive to the evolutionary dynamics and genetic domain walls inherit the front fluctuations [2, 93]. The average interface density satisfies $\langle \rho(t)\rangle \sim t^{-2/3} \sim t^{-1/\tilde{z}}$ [93], where $\tilde{z} = 3/2$ represents the dynamical critical exponent associated with the KPZ equation [160], or equivalently, the noisy Burgers equation [161]. We find that the interface density obeys this scaling for *all* $\alpha' = \beta' < \alpha'_c$ (see Appendix D.2). Rough fronts yield novel critical behavior at the DP2 point for $d = 1 + 1$: The cross-over exponent governing the shape of the phase diagram in Fig. 5.2.1(c) decreases considerably to $\phi' \approx 1.27(15)$, from $\phi \approx 1.9(1)$ for flat fronts. This change leads to a wider mutualistic wedge near the DP2 point. In addition, we find a power law decay of the interface density, $\langle \rho(t)\rangle \sim t^{-\theta'_{\mathrm{DP2}}}$, with a dramatically different critical exponent $\theta'_{\mathrm{DP2}} \approx 0.61(1)$ compared to $\theta_{\mathrm{DP2}} \approx 0.285(5)$ for flat fronts [103] (also see Appendix D.3). For $d = 2 + 1$, we did not have enough statistics to precisely determine the phase diagram shape. However, the DP2 point again appears to move to the origin.

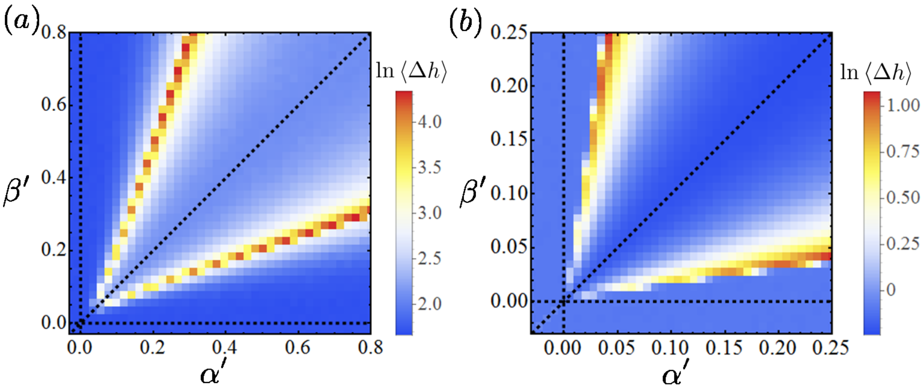


Figure 5.3.1: (Color online) The fluctuation $\langle \Delta h(t)\rangle$ of the rough frontier (averaged over 10^3 runs) of a range expansion with: (a) $d = 1+1$ with $L = 10^3$ cells at time $t = 2.5 \times 10^3$ generations; (b) $d = 2 + 1$ with $L = 50^2$ cells at time $t = 500$ generations. The black dotted lines indicate the loci $\alpha' = 0$, $\beta' = 0$, and $\alpha' = \beta'$.

The front roughness is a remarkable barometer of the onset of mutualism. We characterize the roughness by calculating the interface height $h(\mathbf{x}, t)$ and its root mean square fluctuation $\langle \Delta h(t)\rangle \equiv \sqrt{\langle (h(\mathbf{x}, t) - \langle h(\mathbf{x}, t)\rangle)^2\rangle}$, where $\langle \ldots \rangle$ is both an ensemble average and an average over all points $\mathbf{x}$ at the front. Front fluctuations are greatly enhanced along the pair of DP transition lines for $d = 2 + 1$ and $d = 1 + 1$, as shown in Fig. 5.3.1. At long times, the roughness saturates due to the finite system size (see Ref. [94] and also Appendix D.3).

5.4 Conclusions

To summarize, we found that a mutualistic phase is more accessible in three-dimensional than in two-dimensional range expansions. Also, antagonistic interactions between genetic variants in three dimensions create an effective line tension between genetic domains. The line tension vanishes at a neutral point where the variants do not interact and where the mutualistic phase "wedge" gets pinned to the origin. In addition to the power laws governing the phase diagram shapes in Fig. 5.2.1(a, b, c), we found a striking interface roughness enhancement at the onset of mutualism. These results should apply to a wide variety of expansions because they are insensitive to the microscopic details of our models along transition lines, where we expect universal behavior at large length scales and long times. The existence of universality has been established for flat fronts [95] and a recent study points to a rough DP universality class [94].

It would be interesting to compare two- and three-dimensional range expansions of microorganisms [22, 23] to test the predicted pinning of the mutualistic phase to the $\alpha = \beta = 0$ point. In two dimensions, these expansions are readily realized in Petri dishes [78, 92]. In three dimensions, one may, for example, grow yeast cell *pillars* on a patterned Petri dish with an influx of nutrients at one end of the column, or study the frontier of a growing spherical cluster in soft agar [162].

Appendices

Appendix A

Chapter 2 Appendices

A.1 Detailed Effective Medium Calculation

Our effective medium calculation of the response function Σ will closely follow Cukier and Freed's analysis of the perfectly absorbing cell case [64]. We generalize their work by allowing for partially absorbing cells via the radiation boundary condition [Eq. (2.13)]. Cukier also considered this case [65], but only to first order in scattering in the effective medium. We extend his argument to include the pair distribution function of the cells, which leads to an important correction term.

We begin by solving for the Lagrange multiplier functions $\sigma_\alpha(\mathbf{r}_i)$ in Eq. (2.12) in terms of an "effective medium propagator" $G_W(\mathbf{r}, \mathbf{r}')$. Upon assuming the effective nutrient transport properties are homogeneous over the cell cluster, the propagator is translationally invariant ($G_W(\mathbf{r},\mathbf{r}') \equiv G_W(\mathbf{r}-\mathbf{r}')$) and, in Fourier space,

$$\hat{G}_W(\mathbf{q}) \equiv \left[D_0 q^2 + W(q)\right]^{-1}, \tag{A.1}$$

where W is some suitable approximation to the self-energy Σ that tells us how nutrients diffuse through a medium consisting of the nutrient sinks and their interstitial space.. To simplify the analysis, we introduce an operator notation for all of our convolution integrals. For example, $[W\Psi]_{\mathbf{r}} \equiv \int W(\mathbf{r}-\mathbf{r}')\Psi(\mathbf{r}')\,\mathrm{d}\mathbf{r}'$. Subtracting the convolution $W\Psi$ from both sides of the equation for $\Psi(\mathbf{r})$ [Eq. (2.12)] yields

$$\begin{aligned}\Psi(\mathbf{r}) &= -\int \mathrm{d}\mathbf{r}'\, G_W(\mathbf{r}-\mathbf{r}')\left[\sum_{i=1}^{N}\tilde{\sigma}_i(\mathbf{r}') + s(\mathbf{r}') - [W\Psi]_{\mathbf{r}'}\right]\\ &\equiv -G_W\left[\sum_{i=1}^{N}\tilde{\sigma}_i + s - W\Psi\right],\end{aligned} \tag{A.2}$$

where $\tilde{\sigma}_i(\mathbf{r}) = \int \mathrm{d}\Omega_i\,\delta(\mathbf{r}-\mathbf{R}_i)\sigma_i(\mathbf{r}_i)$, and $G_W[\ldots] \equiv \int G_W(\mathbf{r}-\mathbf{r}')[\ldots]_{\mathbf{r}'}\,\mathrm{d}\mathbf{r}'$ is another example of the convolution described above.

To solve for σ_i via Eq. (A.2), denote the inverse of the G_W operator on the ith cell surface by $K_i^{-1}(\mathbf{r}_i,\mathbf{r}_i')$, where $\mathbf{r}_i$ and $\mathbf{r}_i'$ are two vectors from the cell origin to the cell surface (see Fig. 2.2.3(c)). Then, since Eq. (A.2) is an equation for $\Psi(\mathbf{r})$ for all points $\mathbf{r}$ in the cluster, $\Psi(\mathbf{r})$ in Eq. (A.2) is evaluated at an arbitrary point $\mathbf{r} = \mathbf{R}_i$ on the i-th cell surface. In the perfectly absorbing case Cukier and Freed [64] considered, the boundary condition $\Psi(\mathbf{R}_i) = 0$ is then used to solve Eq. (A.2) for $\sigma_i(\mathbf{r}_i)$. The analogous condition for partially absorbing cells is Eq. (2.13). Using the partially absorbing boundary condition yields

$$\sigma_i(\mathbf{r}_i) = -\int \mathrm{d}\Omega_i'\, g_i(\mathbf{r}_i, \mathbf{r}_i') \left[(1-\mathcal{Q}_i')G_W\left(s - W\Psi + \sum_{j\neq i}\tilde{\sigma}_j\right)\right]_{\mathbf{r}'=\mathbf{R}_i'}. \tag{A.3}$$

where $g_i \equiv \left[1 - K_i^{-1}\mathcal{Q}_i G_W\right]^{-1} K_i^{-1}$ is an operator defined on the surface of the ith sphere and does not depend explicitly on the sphere center $\mathbf{R}_i^0$. $\mathcal{Q}_i'$ is the gradient operator with respect to the $\mathbf{r}'$ coordinate. Equation (A.3) corresponds to Eq. 2.9 in Cukier's analysis [65]. Note that in the operator notation, g_i multiplying a function implies a convolution over the ith sphere surface: $g_i[\ldots] \equiv \int \mathrm{d}\Omega_i' g_i(\mathbf{r}_i, \mathbf{r}_i')[\ldots]_{\mathbf{r}=\mathbf{R}_i'}$.

A useful "scattering" operator used in the following is

$$T_i(\mathbf{r}, \mathbf{r}') \equiv \int \delta(\mathbf{r} - \mathbf{R}_i) g_i(\mathbf{r}_i, \mathbf{r}_i')(1 - \mathcal{Q}_i')\delta(\mathbf{r}' - \mathbf{R}_i')\, \mathrm{d}\Omega_i\, \mathrm{d}\Omega_i'. \tag{A.4}$$

The operator T_i describes the scattering of a nutrient off of the surface of the ith cell. We now substitute Eq. (A.3) into the $\tilde{\sigma}_i$ term in Eq. (A.2) and solve Eq. (A.2) for Ψ by iteration:

$$\begin{aligned}
\Psi &= G_W\left[\sum_{i=1}^{N} T_i G_W\left(s - W\psi + \sum_{j\neq i}\tilde{\sigma}_j\right) - s + W\Psi\right] \\
&= \left[1 - G_W\overline{T} + \sum_{i,j\neq i} G_W T_i G_W T_j - \ldots\right] G_W(W\Psi - s) \\
&= -\left[1 + G_W(\overline{T} - W)\right]^{-1} G_W s, \qquad\qquad \text{(A.5)}
\end{aligned}$$

where $\overline{T} \equiv \sum_i T_i$. In Eq. (A.5), it is important to incorporate an exclusion condition for consecutive sums in the expansion of $[1 + G_W(\overline{T} - W)]^{-1}$, e.g. $\overline{T}G_W\overline{T} = \sum_{i,j\neq i} T_i G_W T_j$

(see [64, 65] for more details). Note also that Eq. (2.16) (main text) implies that the ensemble averaged field $\psi = \langle\Psi\rangle$ in Fourier space satisfies

$$\begin{aligned} -D_0 q^2\psi - W\psi &= \Sigma\psi - W\psi + s \\ \psi &= -G_W(\Sigma - W)\psi - G_W s \\ \psi &= -\left[1 + G_W(\Sigma - W)\right]^{-1} G_W s. \end{aligned} \tag{A.6}$$

The field ψ in Eq. (A.6) must be equal to the ensemble average of Eq. (A.5). We combine the two expressions and set $W = \Sigma$ to find a self-consistent equation for Σ. After some algebraic manipulations (see Cukier's analysis [65] for details), the equation can be expanded in a series in $\overline{T}$. The expansion up to second order is

$$\Sigma \approx \left\langle\overline{T}\right\rangle - \sum_{i,j\neq i}\langle T_i G_\Sigma T_{j\neq i}\rangle + \left\langle\overline{T}\right\rangle G_\Sigma \left\langle\overline{T}\right\rangle. \tag{A.7}$$

It makes sense to expand in $\overline{T}$ because this operator describes a single interaction of the nutrient with any of the cells. The higher order terms in the expansion describe multiple scattering events, which we expect to be less probable. We follow Cukier's analysis [65] in keeping only the first term on the right hand side of Eq. (A.7).

The various operators in Eq. (A.7) are computed by moving to Fourier space and exploiting expansions in spherical harmonics. For example, $G_W(\Omega_i, \Omega_i') \equiv G_W(\mathbf{r}_i - \mathbf{r}_i')$ [given by Eq. (A.1)] is expanded in Fourier modes $e^{i\mathbf{q}\cdot(\mathbf{r}_i - \mathbf{r}_i')}$, which are then rewritten in terms of spherical harmonics via the spherical wave expansion of the plane wave [163].

We find

$$G_W(\Omega_i, \Omega_i') = \sum_{\ell,m} \gamma_\ell Y_{\ell m}(\Omega_i') Y_{\ell m}^*(\Omega_i), \tag{A.8}$$

where $l = 0, 1, \ldots$, $m = -\ell, -\ell + 1, \ldots, \ell$, and

$$\gamma_\ell \equiv \int_0^\infty \frac{2q^2 \mathrm{d}q}{\pi} \frac{j_\ell(qa)^2}{D_0 q^2 + W(q)}. \tag{A.9}$$

The j_ℓ functions are the spherical Bessel functions of the first kind. Cukier [65] describes this procedure in more detail. Another important result (Cukier's Eq. A15) is

$$g_i(\mathbf{r}_i, \mathbf{r}_i') = \sum_{\ell,m} [1 - \gamma_\ell^{-1} \zeta_\ell]^{-1} \gamma_\ell^{-1} Y_{\ell m}(\Omega_i) Y_{\ell m}^*(\Omega_i'), \tag{A.10}$$

where

$$\zeta_\ell \equiv \frac{2a}{\pi \nu} \int_0^\infty \mathrm{d}q \, \frac{q^3 j_\ell'(qa) j_\ell(qa)}{D_0 q^2 + W(q)}. \tag{A.11}$$

Now that we have an expression for g_i, it is possible to compute $\langle \overline{T} \rangle$. The bracket ensemble average will require an averaging over all cell positions, so let us assume that a single cell center is distributed uniformly over the cluster volume V, so that $P(\mathbf{R}_i^0) = V^{-1}$ for all $i = 1, \ldots, N$. After averaging, the operator $\langle \overline{T} \rangle$ acts on an arbitrary function $f(\mathbf{q})$

in Fourier space as follows:

$$
\begin{aligned}
\langle \overline{T} \rangle \, f(\mathbf{q}) &= c \int e^{-i\mathbf{k}\cdot\mathbf{r}_0} g_0(\mathbf{r}_0, \mathbf{r}_0')(1 - i\nu^{-1}\mathbf{r}_0' \cdot \mathbf{q}') e^{i\mathbf{q}'\cdot\mathbf{r}_0'} e^{i(\mathbf{q}'-\mathbf{q})\cdot\mathbf{r}} f(\mathbf{q}') \, \mathrm{d}\Omega_0 \, \mathrm{d}\Omega_0' \, \frac{\mathrm{d}\mathbf{q}'}{(2\pi)^3} \, \mathrm{d}\mathbf{r} \\
&= 4\pi c \sum_\ell \kappa_\ell (2\ell+1) j_\ell(aq) \left[j_\ell(aq) - \nu^{-1} aq j_\ell'(aq) \right] f(\mathbf{q}) \\
&\equiv \mathcal{T}(\mathbf{q}, \mathbf{q}) f(\mathbf{q}), \qquad \text{(A.12)}
\end{aligned}
$$

where $c = N/V$ is the concentration of cells and $\kappa_\ell \equiv [1 - \gamma_\ell^{-1}\zeta_\ell]^{-1}\gamma_\ell^{-1}$. In the last line of Eq. (A.12), we have introduced $\mathcal{T}(\mathbf{q}, \mathbf{q}')$, a convenient scattering operator (averaged over the single cell center distribution $P(\mathbf{R}_i^0) = V^{-1}$). In Fourier space,

$$
\mathcal{T}(\mathbf{q}, \mathbf{q}') \equiv 4\pi \sum_\ell \kappa_\ell \left[j_\ell(aq') - \nu^{-1} aq' j_\ell'(aq') \right] j_\ell(aq)(2\ell+1) P_\ell(\cos\theta_q), \qquad \text{(A.13)}
$$

where θ_q is the angle between $\mathbf{q}$ and $\mathbf{q}'$ and P_ℓ are the Legendre polynomials. Only the diagonal elements $\mathcal{T}(\mathbf{q}, \mathbf{q})$ contribute to $\langle \overline{T} \rangle$. The off-diagonal terms will contribute to the second order correction term $\sum_{i,j\neq i} \langle T_i G_\Sigma T_{j\neq i} \rangle$ in Eq. (A.7).

The second order correction term requires the use of the pair distribution function $P(\mathbf{R}_i^0, \mathbf{R}_j^0)$ for the centers of pairs of cells i and j. We use the approximation Eq. (2.31) for $P(\mathbf{R}_i^0, \mathbf{R}_j^0)$ and find

$$
\sum_{i,j\neq i} \langle T_i G_\Sigma T_{j\neq i} \rangle = \frac{N(N-1)}{V^2} \int \mathcal{T}(\mathbf{q}, \mathbf{q}') \hat{G}_\Sigma(\mathbf{q}') \hat{\theta}(\mathbf{q}' - \mathbf{q}) \mathcal{T}(\mathbf{q}', \mathbf{q}) \, \frac{\mathrm{d}\mathbf{q}'}{(2\pi)^3}, \qquad \text{(A.14)}
$$

where $\hat{\theta}(\mathbf{q})$ is the Fourier transform of $\theta(|\mathbf{r}| - 2a)$:

$$
\hat{\theta}(\mathbf{q}) = -\frac{32\pi a^3 j_1(2aq)}{2aq} + (2\pi)^3 \delta^3(\mathbf{q}). \qquad \text{(A.15)}
$$

We assume N is large enough so that $N(N-1) \approx N^2$, and then substitute Eq. (A.14) into Eq. (A.7) to find that to second order in scattering in the effective medium

$$\Sigma(\mathbf{q}) = c\mathcal{T}(\mathbf{q},\mathbf{q}) + \frac{4a^3c^2}{\pi^2} \int \mathcal{T}(\mathbf{q},\mathbf{q}')\hat{G}_\Sigma(\mathbf{q}') \frac{j_1(2a|\mathbf{q}'-\mathbf{q}|)}{2a|\mathbf{q}'-\mathbf{q}|} \mathcal{T}(\mathbf{q}',\mathbf{q})\,\mathrm{d}\mathbf{q}'. \tag{A.16}$$

In principle, Eq. (A.16) could be solved for $\Sigma(\mathbf{q})$ numerically using an iterative procedure. However, the second order terms involve an integral of $\hat{G}_\Sigma(\mathbf{q}) = (D_0q^2+\Sigma(q))^{-1}$ multiplied by various Bessel functions. Due to the oscillatory nature of the Bessel functions, these kinds of integrals have poor convergence properties using standard numerical techniques and require special integration methods [164]. To avoid these complications, we make a "hydrodynamic" approximation [54, 65, 165] in which $\Sigma(\mathbf{q})$ is approximated by the first two terms of its Taylor expansion: $\Sigma(\mathbf{q}) \approx k + \delta D q^2$.

The constants γ_ℓ and ζ_ℓ have closed forms in the hydrodynamic approximation and involve integrations over spherical Bessel functions (tabulated in [4]). We first note that

$$\gamma_\ell = \frac{2}{D\pi a} \int_0^\infty \frac{j_\ell(x)^2}{x^2+\alpha^2}\, x^2\,\mathrm{d}x = \frac{\alpha i_\ell(\alpha) k_\ell(\alpha)}{Da}, \tag{A.17}$$

where $\alpha \equiv a/\xi$ is the ratio of the cell radius to the correlation length $\xi = \sqrt{D/k}$ and $i_\ell(x)$, $k_\ell(x)$ are the modified spherical Bessel functions of the first and second kinds, respectively. Next, we have

$$\zeta_\ell = \frac{2}{\pi\nu Da} \int_0^\infty \frac{j_\ell'(x_+) j_\ell(x)}{x^2+\alpha^2}\, x^3\,\mathrm{d}x = \frac{\alpha^2 i_\ell(\alpha) k_\ell'(\alpha)}{\nu Da}, \tag{A.18}$$

where the x_+ means we evaluate the derivative of the Bessel function at $x+\epsilon$ and let $\epsilon \to 0$. Using Eq. (A.17) and Eq. (A.18) to evaluate $\mathcal{T}(\mathbf{q},\mathbf{q}')$ in Eq. (A.13) and substituting $\mathcal{T}(\mathbf{q},\mathbf{q}')$ into Equation (A.16) (evaluated at $\mathbf{q}=0$) yields

$$k = c\mathcal{T}(0,0) + \frac{8a^2c^2}{\pi D}\left[\frac{4\pi D a\alpha\nu(1+\coth\alpha)}{1+\alpha+\nu}\right]^2 \\ \times \int_0^\infty \frac{j_0(x)j_1(2x)\left[j_0(x)-\nu^{-1}xj_0'(x)\right]}{x^2+\alpha^2}\,x\,\mathrm{d}x, \tag{A.19}$$

where the integral over the product of three spherical Bessel functions is performed by writing the Bessel functions in terms of elementary functions such as powers and exponentials and then using contour integration techniques. The final result reduces to the implicit equation for ξ, given in the main text as Eq. (2.32).

The diffusion coefficient D can be calculated by looking at the second order term in the Taylor expansion of Eq. (A.16) around $\mathbf{q}=0$. Note that for the purposes of determining the effects on nutrient shielding in a cell cluster, the relevant quantity is the screening length ξ. A lengthy calculation of the change in the diffusion coefficient $\delta D = D - D_0$ yields

$$\begin{aligned}\frac{\delta D}{D} &= \frac{c}{2}\,\frac{d^2}{dq^2}\left[\mathcal{T}(\mathbf{q},\mathbf{q})\right]\Big|_{\mathbf{q=0}} + \frac{a^3c^2}{\pi^2}\,\frac{d^2}{dq^2}\left[\int \mathrm{d}\mathbf{q}'\,\mathcal{T}(\mathbf{q},\mathbf{q}')\hat{G}_\Sigma(\mathbf{q}')\frac{j_1(2a|\mathbf{q}'-\mathbf{q}|)}{a|\mathbf{q}'-\mathbf{q}|}\mathcal{T}(\mathbf{q}',\mathbf{q})\right]\Big|_{\mathbf{q=0}} \\ &= \frac{(1-\nu)(\tilde{\kappa}_0-\tilde{\kappa}_1)}{\nu}\left[1+\frac{6\tilde{\kappa}_0}{\alpha^2}\,\phi\right]\phi - \frac{9\tilde{\kappa}_0^2}{\alpha^4}\,\phi^2 \\ &\quad + 36\left\{\tilde{\kappa}_0 i_0(\alpha)\left[\left(\frac{1}{\alpha^2}+1-\frac{1}{3\nu}\right)k_1(2\alpha)+\frac{2k_0(2\alpha)}{3\alpha}\right]\right. \\ &\qquad\qquad \left. + \frac{2(1-\nu)k_2(2\alpha)\tilde{\kappa}_1 i_1(\alpha)}{3\nu}\right\}\tilde{\kappa}_0\left[i_0(\alpha)-\frac{\alpha i_0'(\alpha)}{\nu}\right]\phi^2 \\ &\quad + \frac{12\tilde{\kappa}_1}{\nu}\left[\tilde{\kappa}_1(\nu-1)k_1(2\alpha)i_1(\alpha)-2\nu k_2(2\alpha)\tilde{\kappa}_0 i_0(\alpha)\right]\left[i_1(\alpha)-\frac{\alpha i_1'(\alpha)}{\nu}\right]\phi^2,\end{aligned} \tag{A.20}$$

where $\alpha \equiv a/\xi$ is given by the self-consistent solution to Eq. (2.32) and $\left[\tilde{\kappa}_\ell(\alpha)\right]^{-1} \equiv \alpha i_\ell(\alpha)\left[k_\ell(\alpha) - \frac{\alpha k_\ell'(\alpha)}{\nu}\right]$ for $\ell = 0, 1$.

Appendix B

Chapter 3 Appendices

B.1 Quasi-species and Mutation-Selection Balance

An important biological motivation for studying irreversible, deleterious mutations is the molecular quasispecies theory developed by Eigen [111]. In this theory (reviewed in [112, 113]), each individual in an asexually reproducing population has a molecular sequence (i.e., a DNA or a protein sequence). During each reproduction step, the sequence gets replicated and acquires errors at some rate μ. The errors change the sequence's fitness. If the error rate is small, the population should become well adapted and its sequences will be clustered around a peak in a fitness landscape. The individuals at the peak have a unique "master sequence" corresponding to this highest possible fitness in the population, as illustrated in Fig. B.1.1. Each mutation away from the master sequence is necessarily deleterious. For a variety of fitness peaks with this general structure, there is a critical mutation rate μ_c above which the population fitness distribution becomes delocalized, and the fraction of individuals that preserve the master sequence goes to zero at long times. The critical mutation rate μ_c is called an “error threshold” [112].

The irreversible mutation case ($\mu_b = 0$) we consider in Sec. 3.3.4 corresponds to a particularly simple fitness landscape in which the fitness peak is infinitely sharp and each mutation away from the master sequence will have an identical deleterious effect with a selection coefficient s (see Fig. B.1.1 for an illustration and Ref. [112] for more details) . The green cells are the cells with the master sequence. Cells without the master sequence are red and have the same fitness because the landscape is flat away from the peak. If the master sequence is very long, then the probability that a red cell recovers the master sequence as it wanders aimlessly through sequence space is vanishingly small. We can then treat the error rate μ as an irreversible mutation rate $\mu = \mu_f$ which mutates green cells to red cells. The red cells also acquire errors with rate μ, but these errors will just move them around the flat fitness landscape.

The error threshold transition in the quasi-species model for the simple fitness landscape in Fig. B.1.1 is identical to the directed percolation transition discussed in this chapter. In population genetics, these transitions are typically studied in well mixed populations, which corresponds to a mean field solution in the directed percolation language. The mean field solution for the density of green cells $f(t)$ (fraction of cells with a master sequence) can be found from Eq. (3.9) by assuming a homogeneous (space independent) density $f(\mathbf{x}, t) \equiv f(t)$ and negligible noise. For an all green initial population ($f(t = 0) = 1$), the fraction of green cells (averaged over many experiments) is given by the solution of $d\langle f(t)\rangle/dt = (s-\mu_f)\langle f(t)\rangle - s\langle f(t)\rangle^2$:

$$\langle f(t)\rangle = \begin{cases} \dfrac{s-\mu_f}{s-\mu_f e^{-t(s-\mu_f)}} & s \neq \mu_f \\ \\ \dfrac{1}{1+\mu_f t} & s = \mu_f \end{cases}. \tag{B.1}$$

Thus, $\langle f(t)\rangle$ approaches 0 or $1-\mu_f/s$ exponentially with time for $\mu_f < s$ and $\mu_f > s$, respectively. The error threshold is given by the critical selective advantage $s_c = \mu_f$, and we evidently have a phase transition controlled by mutation-selection balance. For more details about the relationship between quasi-species and mutation-selection balance, see Ref. [166]. Note that this mean field analysis assumes a very large population size with negligible genetic drift – the sharp phase transition is destroyed by number fluctuations in the well mixed "zero-dimensional" case.

Finite spatial diffusion and noise due to genetic drift drastically change the position of the phase transition and the broad features of the dynamics of $f(\mathbf{x},t)$. Upon defining the rates μ_f and s in units of the inverse cell division time (equivalent to setting the generation time $\tau_g = 1$), we construct the phase diagram (Fig. 3.3.7) by studying the limiting fraction of green cells $\lim_{t\to\infty}\langle f(\mathbf{x},t)\rangle_{\mathbf{x}}$ with the "all green" initial condition $f(\mathbf{x},0) = 1$, where $\langle\ldots\rangle_{\mathbf{x}}$ is both an ensemble average and an average over all cell positions in the growing population at some time t. The Monte Carlo simulations used to calculate this quantity are described in Sec. 3.2. A much stronger critical selective advantage ($s_c \sim \sqrt{\mu_f}$) than that expected from mean field theory ($s_c = \mu_f$) is required to maintain the green strain in the population. These and other corrections to the mean field behavior have been extensively studied with both simulations and field theoretic methods [95, 103].

B.2 Coarse-Graining and Langevin Description

The simulations in this chapter focus on voter-type models, with one individual per lattice site. However, a coarse-graining procedure leads to a description in terms of continuous variables, similar to that reviewed for stepping stone models in Ref. [13]. We coarse-grain voter-type models by considering a set of $\Omega > 1$ cells, and rescaling our lattice constant so that these Ω cells are all contained within a spatial volume a^d. Following, e.g., Vazquez and López [167], a coarse-grained population density of green cells is given by

$$f(\mathbf{x},t) \equiv \frac{1}{\Omega} \sum_{i \in \mathcal{B}_\Omega(\mathbf{x})} \frac{1+\sigma_i(t)}{2}, \tag{B.2}$$

where the sum is over cells in a neighborhood $\mathcal{B}_\Omega(\mathbf{x})$ of the Ω cells centered on $\mathbf{x}$, and we use Ising spin variables $\sigma_i(t) = \pm 1$ to specify the occupancy of site i. Individual cells in this neighborhood change states according to the rates given by Eq. (3.4). We now randomly choose one of the Ω spins to flip in each neighborhood. This means that each spin flip at $\mathbf{x}'$ changes our field $f(\mathbf{x},t)$ by a displacement field $\pm r(\mathbf{x},\mathbf{x}')$ given by

$$\pm r(\mathbf{x},\mathbf{x}') \equiv \pm \Omega^{-1} a^d \delta(\mathbf{x}-\mathbf{x}'), \tag{B.3}$$

where the sign depends on the state of the spin at $\mathbf{x}'$. In addition, we now use the average value of the spins in neighboring cells to evaluate the summations over nearest neighbors in the flipping rates due to selection and genetic drift (see Eq. (3.6) and Eq. (3.7)). The resulting dynamics resembles the stepping stone model of population genetics [11, 13], with populations of size Ω on compact "islands" where mutation and selection act, with exchange of individuals between neighboring islands.

To study the dynamics of the the coarse-grained population density $f(\mathbf{x}, t)$, we introduce rates $\omega(f(\mathbf{x}); \pm r(\mathbf{x}, \mathbf{x}'))$ that describe a transition from a starting field configuration f to a field $f \pm r$. The new coarse-grained transition rates (in units of τ_g^{-1}) that define an underlying master equation for the field f are

$$\omega[f; +r] \equiv \omega\left[f(\mathbf{x}) \rightarrow f(\mathbf{x}) + r(\mathbf{x}, \mathbf{x}')\right] = [1 - f(\mathbf{x})]\left[\mu_b + \frac{1}{z}\sum_i f(\mathbf{x} + \boldsymbol{\delta}_i)\right], \tag{B.4}$$

and

$$\omega[f; -r] = f(\mathbf{x})\left[\mu_f + (1 - s) + \frac{s - 1}{z}\sum_i f(\mathbf{x} + \boldsymbol{\delta}_i)\right], \tag{B.5}$$

where we sum over z neighboring unit cells at vector distances $\boldsymbol{\delta}_i$ from the cell centered on $\mathbf{x}$. For a cubic lattice $\boldsymbol{\delta}_i = a\hat{\mathbf{x}}_i$, where $\hat{\mathbf{x}}_i$ are Cartesian unit vectors.

To construct the master equation for the probability distribution functional $P[f, t]$ of the coarse-grained field, we expand in $1/\Omega$. This Kramers-Moyal expansion [14, 15, 44] leads to a Fokker-Planck equation (or "Kolmogorov forward equation") for the voter model dynamics. The rates $\omega(f; \pm r)$ will vary rapidly with the displacement r. However, if we choose a Ω large enough, we can expand $\omega(f; \pm r)$ around $r = 0$ in powers of Ω^{-1} [14]. To second order in Ω^{-1}, the master equation for $P \equiv P[f(\mathbf{x}), t]$ follows from

$$\begin{aligned}\partial_t P = \int \mathrm{d}^{\mathrm{d}}\mathbf{x}' \Big\{ &\omega\left[f + r(\mathbf{x}'); -r(\mathbf{x}')\right] P\left[f + r(\mathbf{x}'), t\right] + \omega\left[f - r(\mathbf{x}'); r(\mathbf{x}')\right] P\left[f - r(\mathbf{x}'), t\right] \\ &- \left[\omega\left(f; r(\mathbf{x}')\right) + \omega\left(f; -r(\mathbf{x}')\right)\right] P[f, t] \Big\},\end{aligned} \tag{B.6}$$

which leads to

$$\partial_t P \approx -\frac{1}{\Omega}\frac{\delta}{\delta f}\left\{\left[\mu_b(1-f) - \mu_f f + sf\,(1-f) + \frac{(1-sf)a^2}{z}\nabla^2 f\right] P\right\}$$
$$+\frac{1}{2\Omega^2}\frac{\delta^2}{\delta f^2}\left\{\left[\mu_f f + (2-s)\,f(1-f) + \mu_b(1-f) + \frac{(s-f)a^2}{2z}\,\nabla^2 f\right] P\right\}. \quad \text{(B.7)}$$

Eq. (B.7) is a Fokker-Planck equation for the coarse-grained dynamics of the voter model in d-dimensions. For a number of applications, it is convenient to shift to an equivalent Langevin description [13]. Upon rescaling space to eliminate Ω and assuming $s, \mu_f, \mu_b \ll 1$, we use the Îto prescription [15] to move from Eq. (B.7) to the stochastic differential equation for $f(\mathbf{x}, t)$ given by Eq. (3.9) and Eq. (3.10) in the main text. The Langevin equation, when interpreted by the rules of the Îto calculus [15], belongs to the universality class of directed percolation when $\mu_b = 0$ [95], i.e. for "one-way" mutations to the less fit phase in the presence of number fluctuations. Within an ϵ-expansion about $d = 4$ spatial dimensions, the terms we have neglected are irrelevant in the renormalization group sense [95]. Note also the equivalence to a stochastic stepping stone model with one individual per deme [13].

Finally, we consider the connection between our radial expansion models and experiments. Both the diffusion coefficient $D_r = a_r^2/z_r\tau_g$ and the front velocity v could vary for different microbial radial expansions [78]. Although the Langevin description in Eq. (3.14) and Eq. (3.15) can accommodate this variation, our Bennett lattice model is constrained to have $D_r/(va_r) \approx a_r/(z_r v\tau_g) \lesssim 1$. A likely system that satisfies the bound is a *Pseudomonas aeruginosa* radial expansion on a Petri dish [78]. However, the bound will not be satisfied by all microbial radial expansions.

B.3 Single Sector Dynamics

This appendix supplements the discussion in Sec. 3.3.3 by further developing the theory of the dynamics of a single green cell sector evolving in an otherwise red cell population under the radial Domany-Kinzel dynamics (see Sec. 3.2) of a range expansion with a uniform circular front. We will ignore mutations ($\mu_f = \mu_b = 0$) and allow for the green cells to enjoy a selective advantage s. As discussed in the main text, the Langevin equation (Eq. (3.9)) for the coarse-grained cell density $f(\mathbf{x}, t)$ is difficult to analyze for nonzero s. Hence, we move to a "dual" description of a single sector and consider the dynamics of the two sector boundaries at angular positions $\phi_1(t)$ and $\phi_2(t)$ at time t. Without loss of generality, let $\phi_1(t) > \phi_2(t)$ so that $\phi_1(t) - \phi_2(t)$ is the angular sector size.

Suppose that the dynamics of each boundary is independent, and that during each generation we either gain or lose a single cell at each boundary due to local competition with adjacent red cells. Since the cells will generally be staggered along the population frontier, there will be slight variation in the change in the sector size when the sector loses or gains a cell. Hence, we define a parameter $\tilde{a}_r$ equal to the average change in sector size when we add a single cell at the sector boundary. The parameter $\tilde{a}_r$ should be close to the effective lattice spacing a_r. The master equation for the probability $P(\phi_1, t)$ of observing a boundary at ϕ_1 at time t reads

$$\partial_t P(\phi_1, t) = \frac{p_G}{\tau_g} P\left(\phi_1 - \frac{\tilde{a}_r}{R(t)}, t\right) - \left(\frac{1 - p_G}{\tau_g}\right) P\left(\phi_1 + \frac{\tilde{a}_r}{R(t)}, t\right), \tag{B.8}$$

where $\tau_g = 1$ is the generation time, and we approximate $[P(\phi, t + \tau_g) - P(\phi, t)]/\tau_g$ by

$\partial_t P(\phi, t)$. The probability of a green cell out-competing a red cell is given by p_G.

Recall that our Bennett model lattice is constructed so that the circular population radius advances by approximately one cell diameter per generation. Thus, the angular position of the sector boundary at the frontier is determined by the color of approximately one cell. If a single green cell competes with adjacent red cells, the Domany-Kinzel rules (Eq. (3.1)) imply that for $s \ll 1$,

$$p_G \approx \frac{1}{2} - \frac{\gamma s}{4}, \tag{B.9}$$

where $\gamma = 1$ for simple competition between a pair of green and red cells. The disorder in the lattice, however, may change γ slightly as there can be a variable number of total parents n_T that determine p_G (see discussion of Eq. (3.1) in the main text). The continuum limit of the master equation Eq. (B.8) yields the Fokker-Planck equation, namely

$$\partial_t P(\phi_1, t) = -\frac{\gamma s \tilde{a}_r}{2\tau_g R(t)}\,\partial_{\phi_1} P(\phi_1, t) + \frac{\tilde{a}_r^2}{2\tau_g [R(t)]^2}\,\partial_{\phi_1}^2 P(\phi_1, t). \tag{B.10}$$

Note that the other boundary position $\phi_2(t)$ also obeys Eq. (B.10). The Fokker-Planck equations can be converted to stochastic differential equations for $\phi_1(t)$ and $\phi_2(t)$ using standard techniques [15]. Subtracting the stochastic differential equation for $\phi_2(t)$ from the one for $\phi_1(t)$ yields

$$\frac{d\phi}{dt} = \frac{\gamma \tilde{a}_r s}{\tau_g R(t)}\,\phi(t) + \sqrt{\frac{2\tilde{a}_r^2}{\tau_g [R(t)]^2}}\,\eta(t), \tag{B.11}$$

where $\phi(t) = \phi_1(t) - \phi_2(t)$ is the angular sector size and $\eta(t)$ is a Gaussian white noise (with $\langle \eta(t)\eta(t')\rangle = \delta(t - t')$). Upon defining the dimensionless conformal time

$\tau \equiv t_c(t)/t^* = vt/(R_0 + vt) = vt/R(t)$, we find that the Fokker-Planck equation associated with Eq. (B.11) (after changing variables from t to τ) must be the one given by Eq. (3.31) in the main text, with

$$w = \frac{\gamma \tilde{a}_r s}{v \tau_g} \quad \text{and} \quad \Delta = \frac{\tilde{a}_r^2}{R_0 v \tau_g}. \tag{B.12}$$

This simple model is not quite faithful to the simulations on an amorphous Bennett model lattice because the disorder in the lattice will introduce variability in the constant $\tilde{a}_r$ and the exact number of cells at the boundary of the sector per generation. Some of these effects are modelled by γ, but we expect a slight s dependence in Δ, as well. Also, the cell density of the Bennett model lattice decreases a little with increasing population radius. However, by fitting the parameters $\tilde{a}_r$ and γ to simulation results, we partially compensate for these complications, as confirmed by checks of the analytic results against our computer simulations.

When $s = 0$, Eq. (3.31) in the main text reduces to a simple diffusion equation with diffusivity $\Delta = \tilde{a}_r^2/R_0 v \tau_g$, which can be solved exactly. This solution is fitted to simulation results to find $\tilde{a}_r$. We find $\tilde{a}_r \approx 0.91$ for $R_0 = 300$ (recall that $v = 1$ in the simulations). Next, we pick a large initial sector angle $\phi_0 \approx \pi/2$ and measure the average angular size $\langle \phi(t) \rangle$ of the green sector as a function of time t. Since the sector never goes extinct in this case, the dynamics of $\langle \phi(t) \rangle$ is controlled by the selection bias. The green sector then forms, on average, the logarithmic spiral shape described by Eq. (3.35) (see also Ref. [92]). The shape of the spiral in Eq. (3.35) implies that $w = (\langle \phi(t) \rangle - \phi_0)/(\ln R(t) - \ln R_0)$ is a time independent constant. The slope of the plot of w versus s yields an estimate for $\gamma \tilde{a}_r / v$. As expected from our simple model, we find

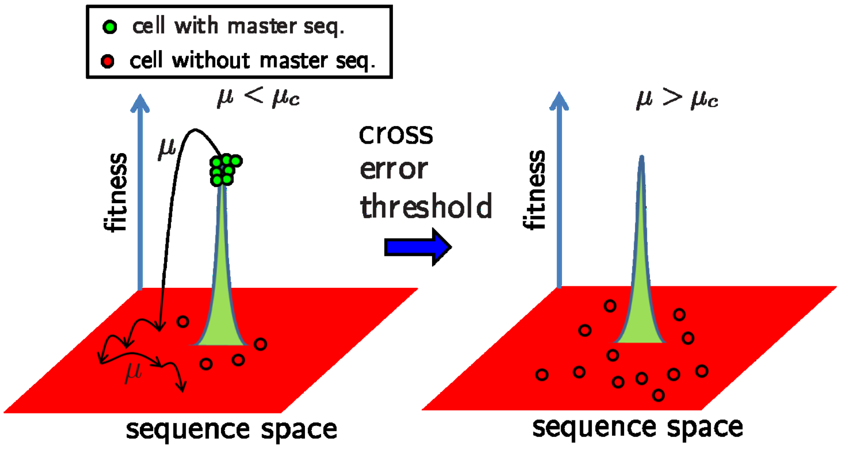


Figure B.1.1: (Color online) A schematic of the quasi-species model in which each cell in a population has a sequence distributed in a fitness landscape with a single sharp peak in green (light gray) surrounded by a flat landscape in red (dark gray). As cells divide, each individual reproduces with a mutational error rate μ that moves it around in sequence space (black, thin arrows). Below a critical “error threshold” ($\mu < \mu_c$), a finite fraction of the population has the master sequence (i.e., is a green (light gray) cell). Cells deviating from this ideal sequence are red (dark gray). When $\mu > \mu_c$, the population fitness distribution becomes delocalized, and the fitness peak is depopulated.

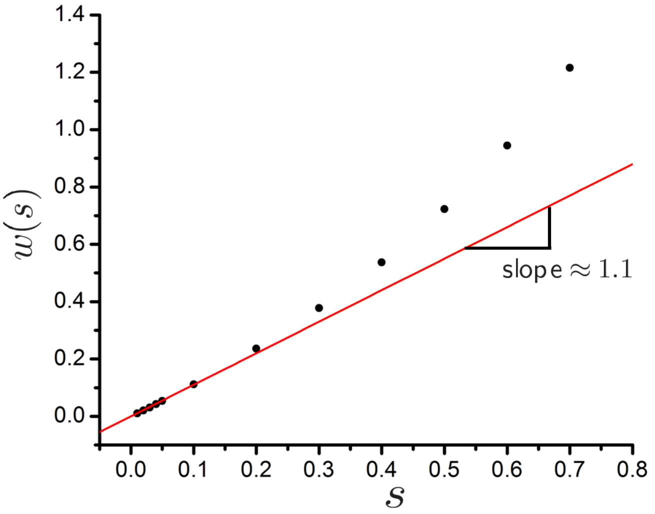


Figure B.3.1: (Color online) The speed $w(s)$ of a green sector boundary at the frontier with various selection advantages s, invading an otherwise red population at the frontier. The solid red line is the best fit of the speed to a linear function for $s \leq 0.1$ ($w(s) \approx 1.1s$). The initial angular sector size is $\phi_0 \approx \pi/2$, and the initial homeland radius $R_0 = 300$. $w(s)$ is calculated by averaging $(\langle\phi(t)\rangle - \phi_0)/\ln[R(t)/R_0]$ over approximately 10^3 generations, where $\langle\phi(t)\rangle$ is the angular size of the green sector at time t, averaged over about 3×10^4 simulation runs.

in Fig. B.3.1 a linear relationship between w and s for $s \lesssim 1$ that yields $\gamma \approx 1.2$

Appendix C

Chapter 4 Appendices

C.1 Response Functional Formalism

The response functional formalism allows us to calculate moments of the coarse-grained green cell density field $f(\mathbf{x}, t)$ (with $\mathbf{x}$ some position along the front at time t), such as $\langle f(\mathbf{x}, t)\rangle$, $\langle f(\mathbf{x}, t) f(\mathbf{x}', t')\rangle$, etc., where the averages are over all possible realizations of the noise $\xi(\mathbf{x}, t)$ in the Langevin equation (4.16) obeyed by $f(\mathbf{x}, t)$. Schematically, then, for any functional $\mathcal{G}[f]$ of $f \equiv f(\mathbf{x}, t)$, we may take an average $\langle \mathcal{G}[f]\rangle = \int \mathcal{D}\xi \, \mathcal{G}[f] P[\xi]$, where we integrate over all possible noise instances $\xi(\mathbf{x}, t)$, each of which has a probability weight

$$P[\xi] \propto \exp\left[-\int \mathrm{d}\mathbf{x} \, \mathrm{d}t \, \frac{\xi(\mathbf{x}, t)^2}{2}\right]. \tag{C.1}$$

We perform the integration over ξ by introducing a functional Dirac delta function. This allows us to change variables from $\xi(\mathbf{x}, t)$ to $f(\mathbf{x}, t)$, ensuring that $f(\mathbf{x}, t)$ obeys

Eq. (4.16), which we can write generally as

$$\partial_t f(\mathbf{x},t) = -R[f(\mathbf{x},t)] + \sqrt{V[f(\mathbf{x},t)]}\xi(\mathbf{x},t), \tag{C.2}$$

where $R[\ldots]$ is the "reaction" term (see Chapter 1) and $\sqrt{V[\ldots]}$ is the noise strength. Both of these are easily identifiable in Eq. (4.16). The integration over ξ then becomes a Gaussian integral as follows:

$$\begin{aligned}
\langle \mathcal{G}[f(\mathbf{x},t)]\rangle &= \int \mathcal{D}\xi\, \mathcal{D}f\, \mathcal{G}[f]\, \delta\left(\partial_t f + R[f] - V[f]\xi\right)\, P[\xi] \\
&= \int \mathcal{D}\xi\, \mathcal{D}f\, \mathcal{D}[i\tilde{f}]\, \mathcal{G}[f] \\
&\qquad \times \exp\left[-\int \mathrm{d}\mathbf{x}\, \mathrm{d}t\, \tilde{f}\left(\partial_t f + R[f] - \sqrt{V[f]}\,\xi\right) - \int \mathrm{d}\mathbf{x}\, \mathrm{d}t\, \frac{\xi(\mathbf{x},t)^2}{2}\right] \\
&= \int \mathcal{D}f\, \mathcal{D}[i\tilde{f}]\, \mathcal{G}[f] \exp\left[-\int \mathrm{d}\mathbf{x}\, \mathrm{d}t\, \tilde{f}\left(\partial_t f + R[f] - \frac{\tilde{f}V[f]}{2}\right)\right] &\text{(C.3)} \\
&\equiv \int \mathcal{D}f\, \mathcal{D}[i\tilde{f}]\, \mathcal{G}[f]\, e^{-\mathcal{J}[f,\tilde{f}]}, &\text{(C.4)}
\end{aligned}$$

where we have rewritten the functional Dirac delta function as a functional Fourier transform by introducing a function $\tilde{f} = \tilde{f}(\mathbf{x},t)$ (the analogue of the wave-number $\mathbf{k}$ for the usual Fourier transform) that we integrate over the imaginary axis for every $\mathbf{x}$ and t. The function $\tilde{f}(\mathbf{x},t)$ is often called a response field. (Note that we have neglected the Jacobian associated with changing variables in the delta function and we have absorbed any normalization constants in the functional measures $\mathcal{D}f$, $\mathcal{D}[i\tilde{f}]$. For more details and subtleties, we refer to the excellent book by Taüber [168].) We see that the force and noise terms in the Langevin equation, when interpreted according to Eq. (C.3), are now on equal footing. Indeed, we can now interpret any average $\langle \ldots \rangle$ over instances of the noise $\xi(\mathbf{x},t)$ as averages weighted by a Boltzmann-like factor $e^{-\mathcal{J}[\tilde{f},f]}$, where $\mathcal{J}[\tilde{f},f]$ is given by Eq. (4.17) and we integrate over all possible $f(\mathbf{x},t)$ and $\tilde{f}(\mathbf{x},t)$.

We now want to use the response function formalism to determine the survival probability. We follow Janssen's analysis in Ref. [140] and consider first the probability of forming an N-cell sector of mutants from an N_0-particle initial one in some time $T \gg \tau_g$. We can calculate the probability by counting up the fraction of noise histories $\xi(\mathbf{x},t)$ that yield a cluster of N cells. Hence, using the functional integration technique in Eq. (C.4), we find

$$P_N(T) \approx \int \mathcal{D}f\,\mathcal{D}\tilde{f}\,\delta\left[N - \int \mathrm{d}^2\mathbf{x}\,\mathrm{d}t\,\theta(T-t)f(\mathbf{x},t)\right] e^{-\mathcal{J}[\tilde{f},f]+N_0\tilde{f}(\mathbf{0},0)}, \tag{C.5}$$

where $\theta(x)$ is the Heaviside theta function and we have explicitly included the initial condition term in the exponential (see Eq. (4.18) and surrounding discussion in the main text). The Dirac delta function in Eq. (C.5) ensures that we only count the noise histories that yield a cluster of N cells. These probabilities $P_N(T)$ can be used to find the survival probability via a generating function or discrete Laplace transform

$$\begin{aligned} G(\zeta,T) &\equiv \sum_{N=0}^{\infty} P_N(T)e^{-\zeta N} && \text{(C.6)}\\ &\approx \int \mathrm{d}N \int \mathcal{D}f\,\mathcal{D}\tilde{f}\,\delta\left[N - \int \mathrm{d}^2\mathbf{x}\,\mathrm{d}t\,\theta(t-T)f(\mathbf{x},t)\right] e^{-\mathcal{J}[\tilde{f},f]+N_0\tilde{f}(\mathbf{0},0)-\zeta N} && \text{(C.7)}\\ &= \int \mathcal{D}f\,\mathcal{D}\tilde{f}\,e^{-\mathcal{J}[\tilde{f},f]+N_0\tilde{f}(\mathbf{0},0)-\zeta\int \mathrm{d}^2\mathbf{x}\,\mathrm{d}t\,\theta(T-t)f(\mathbf{x},t)}. && \text{(C.8)} \end{aligned}$$

This generating function is related to the long-time survival probability P_∞ of the mutant cluster, which turns out to be given by [140]

$$P_\infty = 1 - \lim_{\zeta\to 0^+}\lim_{T\to\infty} G(\zeta) \tag{C.9}$$

$$\approx 1 - \lim_{\zeta\to 0^+}\lim_{T\to\infty} \int \mathcal{D}f\,\mathcal{D}\tilde{f}\; e^{-\mathcal{J}[\tilde{f},f]+N_0\tilde{f}(\mathbf{0},0)-\zeta\int \mathrm{d}^2\mathbf{x}\,\mathrm{d}t\,\theta(T-t)f(\mathbf{x},t)} \tag{C.10}$$

$$P_\infty \approx 1 - \lim_{\zeta\to 0^+}\lim_{T\to\infty} \left\langle e^{N_0\tilde{f}(\mathbf{0},0)}\right\rangle_{\mathcal{J}[\tilde{f},f]+\zeta\int \mathrm{d}^2\mathbf{x}\,\mathrm{d}t\,\theta(T-t)f(\mathbf{x},t)}, \tag{C.11}$$

where the average is performed using the exponentiated field theoretic action in Eq. (4.17), shifted by the Heaviside theta function term: $\mathcal{J}[\tilde{f},f] \to \mathcal{J}[\tilde{f},f]+\zeta\int \mathrm{d}^2\mathbf{x}\,\mathrm{d}t\,\theta(T-t)f(\mathbf{x},t)$.

For a mutant survival probability on an infinite, flat front [$\Theta = 0$ in Eq. (4.16)], it is known that a mean-field approximation (up to logarithmic corrections) works around the special "voter model" case of $s = 0$ with one time dimension and two spatial dimensions. Here, we assume mean-field theory also works for $0 < s \lesssim 1$ and in the presence of inflation [$\Theta > 0$ in Eq. (4.16)]. We will check our mean-field approximation by comparing the results to simulations and find that it works remarkably well. Within the mean-field approximation, the average in Eq. (C.11) can be evaluated by substituting in the spatially uniform, saddle-point approximation for $\tilde{f}(t = 0) \equiv \tilde{f}(\mathbf{0}, 0)$. The saddle point equations, obtained by varying the shifted action $\mathcal{J}[f,\tilde{f}]$ with respect to $\tilde{f}$ and f are, respectively, $f = 0$ and Eq. (4.21) in the main text. We can rewrite the latter equation by introducing a new time variable $\bar{t} \equiv t/T$ so that $\bar{t} \in [0, 1]$. Then, incorporating the step function in Eq. (4.21) into an appropriate boundary condition, we find:

$$\partial_{\bar{t}}\tilde{f}(\bar{t}) = -s\tau_g^{-1}T\tilde{f}(\bar{t}) - \frac{\Delta T[\tilde{f}(\bar{t})]^2}{\tau_g(1+\bar{t}T/t^*)^{2\Theta}} + \zeta T \quad \text{with} \quad \tilde{f}(1) = 0. \tag{C.12}$$

This equation can be solved for $\tilde{f}(t = 0)$ in the $T \to \infty$ and $\zeta \to 0^+$ limit for various growth exponents Θ. We will do this in Appendix C.2.

C.2 Survival Probability Calculation in Three Dimensions

We want to solve Eq. (C.12) for various Θ. To emphasize that we are looking for solutions at arbitrary Θ, let us introduce a subscript to our response function: $\tilde{f}_\Theta(\bar{t})$. Recall that Eq. (C.12) describes a time history of $\tilde{f}_\Theta(\bar{t})$ in the range $0 \leq \bar{t} \leq 1$. Equation (C.12) is a Ricatti equation [169] and we can solve it by introducing a function $u_\Theta(\bar{t})$ such that $\tilde{f}_\Theta(\bar{t}) = (1+\bar{t}T/t^*)^{2\Theta}\partial_{\bar{t}} u_\Theta(\bar{t})/[u_\Theta(\bar{t})\Delta T]$. The Ricatti equation (C.12) is then transformed into a linear second-order differential equation that reads

$$\frac{d^2 u_\Theta}{d\bar{t}^2} + \left(sT + \frac{2\Theta T/t^*}{1+\bar{t}T/t^*}\right)\frac{du_\Theta}{d\bar{t}} - \frac{\Delta\zeta T^2}{(1+\bar{t}T/t^*)^{2\Theta}}\,u_\Theta = 0. \tag{C.13}$$

At this point, we specialize to linear inflation and set $\Theta = 1$. Then, upon changing variables to $z \equiv s(t^* + T\bar{t})$ and introducing a new function $w_\Theta(z) \equiv e^z z^{(1-\nu)/2} u_\Theta(z)$ (where $\nu^2 \equiv 1 + 4(t^*)^2\Delta\zeta$), we find the confluent hypergeometric differential equation

$$z\,\frac{d^2 w_{\Theta=1}}{dz^2} + (1+\nu-z)\frac{dw_{\Theta=1}}{dz} - \frac{3+\nu}{2}\,w_{\Theta=1}(z) = 0. \tag{C.14}$$

The general solution for $w_{\Theta=1}(z)$ is given by a linear combination of special functions [4],

$$w_{\Theta=1}(z) = A_1\,{}_1F_1\left[(3+\nu)/2; 1+\nu; z\right] + A_2\,U\left[(3+\nu)/2, 1+\nu, z\right], \tag{C.15}$$

where A_1 and A_2 are arbitrary constants, and ${}_1F_1[\alpha;\beta;z]$ and $U[\alpha,\beta,z]$ are the confluent hypergeometric functions of the first and second kind, respectively. Upon substituting this general solution into the expression for $\tilde{f}_\Theta(z)$, we find

$$\begin{aligned}\tilde{f}_{\Theta=1}(z) = \frac{sz^2}{\Delta(st^*)^2}\Bigg\{&-1+\frac{\nu-1}{2z}\\ &+\left[\frac{3+\nu}{2}\right]\frac{B\,{}_1F_1\left[(5+\nu)/2;2+\nu;z\right]/(1+\nu)-U\left[(5+\nu)/2,2+\nu,z\right]}{U\left[(3+\nu)/2,1+\nu,z\right]+B\,{}_1F_1\left[(3+\nu)/2;1+\nu;z\right]}\Bigg\},\end{aligned} \tag{C.16}$$

where we will now fix $B \equiv A_1/A_2$ by applying the boundary condition $\tilde{f}_{\Theta=1}[z = s(t^*+T)] = 1$. In the $T \gg t^*$ limit, we have $s(t^*+T) \approx sT$. Our boundary condition, after taking the $\zeta \to 0^+$ limit is

$$\begin{aligned}\tilde{f}_{\Theta=1}(sT) &= -\frac{sT^2}{\Delta(t^*)^2}\left[1-\frac{B\,{}_1F_1\left[3;3;sT\right]-2U\left[3,3,sT\right]}{U\left[2,2,sT\right]+B\,{}_1F_1\left[2;2;sT\right]}\right]\\ &= -\frac{sT^2}{\Delta(t^*)^2}\left[1+\frac{2U\left[3,3,sT\right]-B}{U\left[2;2;sT\right]+B}\right]=1.\end{aligned} \tag{C.17}$$

The solution for B is

$$B = -\frac{2sT^2}{(t^*)^2\Delta}U\left[3,3,sT\right]-\left[1+\frac{sT^2}{(t^*)^2\Delta}\right]U\left[2;2;sT\right]. \tag{C.18}$$

In the long time $T \to \infty$ limit, we will have $B \to 0$ if $s > 0$ and $|B| \to \infty$ if $s < 0$. Hence, we treat these cases separately. We will find a jump discontinuity in the limit at $s = 0$. To find the survival probability, we must evaluate $\tilde{f}_{\Theta=1}(\bar{t})$ at $\bar{t} = 0$ (or $z = st^*$).

The calculation for $s > 0$ in the $T \to \infty$ limit leads to

$$\begin{aligned}\tilde{f}_{\Theta=1}(st^*) &= -\frac{s}{\Delta}\left\{1 - \frac{B - 2U\left[3,3,st^*\right]}{B + U\left[2,2,st^*\right]}\right\} \\ &= -\frac{s}{\Delta}\left[1 + \frac{2U\left[3,3,st^*\right]}{U\left[2,2,st^*\right]}\right] = -\frac{s}{\Delta}\left\{\frac{1}{st^*[1 + st^* e^{st^*}\mathrm{Ei}(-st^*)]}\right\},\end{aligned} \tag{C.19}$$

where $\mathrm{Ei}(x)$ is the exponential integral function [4]. For $s < 0$, we have $\tilde{f}_{\Theta=1}(st^*) = 0$. The solution method is analogous in the marginal $\Theta = 1/2$ case. Indeed, Eq. (C.13) may be transformed into another confluent hypergeometric differential equation. Solving this equation and applying the same limits as in the $\Theta = 1$ case, we find that for $s > 0$

$$\tilde{f}_{\Theta=1/2}(st^*) = -\frac{s}{\Delta}\left[1 + \frac{U\left[2,2,st^*\right]}{U\left[1,1,st^*\right]}\right] = \frac{1}{\Delta t^* e^{st^*}\mathrm{Ei}(-st^*)} \tag{C.20}$$

and $\tilde{f}_{\Theta=1/2}(st^*) = 0$ when $s < 0$. Upon substituting this solution into Eq. (4.20), we find Eq. (4.33) in the main text.

C.3 Survival Probability for $s < 0$

Our theory predicts that a mutation with a selective disadvantage $s < 0$ occurring at the frontier of a circular or spherical population only survives if genetic drift allows it to wrap all the way around the population. Otherwise the deterministic dynamics will collapse the mutant sector and $P(t \to \infty, s < 0) \equiv P_\infty(s < 0) = 0$. However, the convergence to this limiting survival probability is very slow, due to the deterministic dynamics (logarithmic spiral shapes, etc.) described in the main text. To show this, we plot the survival probability $P(t, s)$ at various times t for circular and spherical frontiers in Fig. C.3.1(A,B). Note that for $s > 0$, the probabilities converge rapidly to their steady-state value. Conversely, for $s < 0$, the data points at different t do not overlap in Fig. C.3.1(A,B), indicating that the probabilities have not yet converged. We check that the survival probability has not yet approached a steady-state for small negative s in Fig. C.3.1(C,D) for both circular and spherical frontiers.

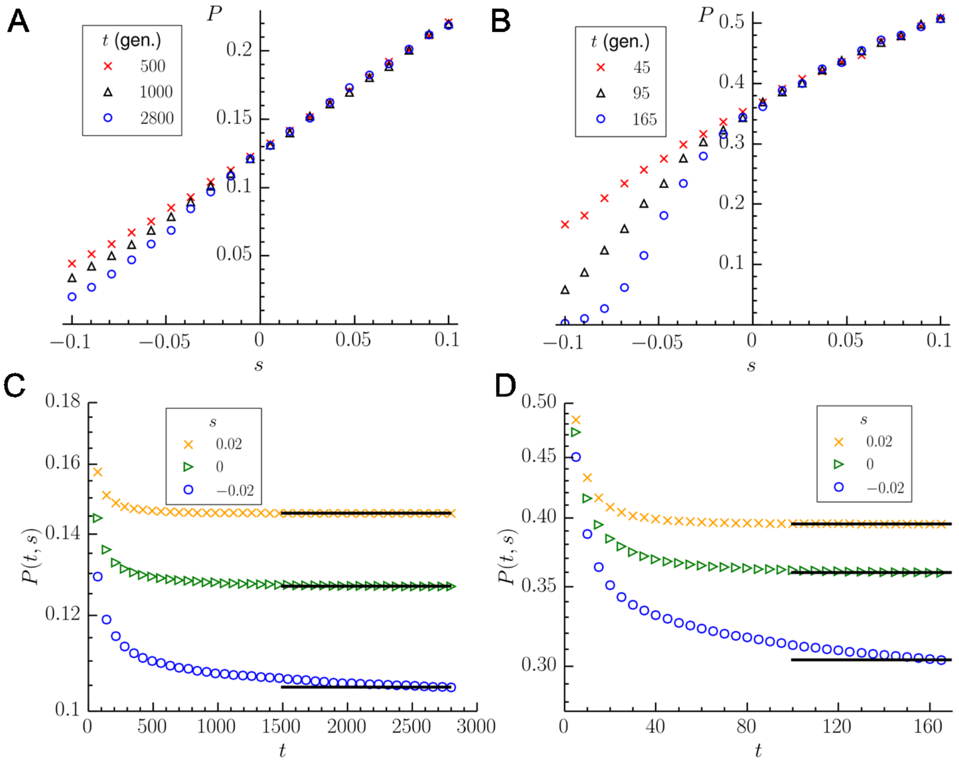


Figure C.3.1: The survival probability $P \equiv P(t, s)$ at time t generations and selection parameter s of a cluster formed from a single neutral green cell at (A,C) a circular frontier with initial radius $R_0 = 20$ cell diameters and (B, D) a spherical frontier with $R_0 = 5$ cell diameters. We plot P as a function of s for various t in (A,B) and as a function of t for a few s values in (C,D) (with a logarithmic vertical scale). When $s < 0$ in (A,B), the survival probability has not yet converged and continues to decrease with increasing t. Note that in the $s = -0.02$ case in (C,D), the probability continues to decrease as we increase t, indicating that it has not yet converged to a long-time value. The horizontal black lines denote the survival probability values at the longest time simulated [2800 generations in (C) and 170 generations in (D)]. Our theory predicts the probability will eventually decay to a vanishingly small (compared to, say, the limiting probability for $s = 0^{+}$) residual probability due to the very small chance that sectors wrap around the entire population.

Appendix D

Chapter 5 Appendices

D.1 Flat Front Models and Phase Boundaries

The update rules for the flat front models in two and three dimensions are given by Eq. (2) in the main text. As discussed previously, these rules are implemented on a square lattice in two dimensions, and on a close-packed hexagonal array in three dimensions. Figure D.1.1 summarizes the update rules for the two- and three-dimensional models, where one dimension represents time.

In two dimensions we find a demixing regime for parameter values $\alpha = \beta < \alpha_c \approx 0.1242$ as discussed in the main text. In this regime, we can model the domain walls as pair-annihilating random walkers. During each generation, domain walls can hop to the left or right by one lattice site through updates such as $\circ \circ \circ | \bullet \bullet\bullet \to \circ \circ | \bullet \bullet \bullet \bullet$ (hop to left). For $\alpha = \beta$, the hops to the left and right occur with equal probability p. If we ignore possible domain branching events in each generation (such as $\circ \circ \circ | \bullet \bullet\bullet \to \circ \circ | \bullet | \circ | \bullet \bullet$),

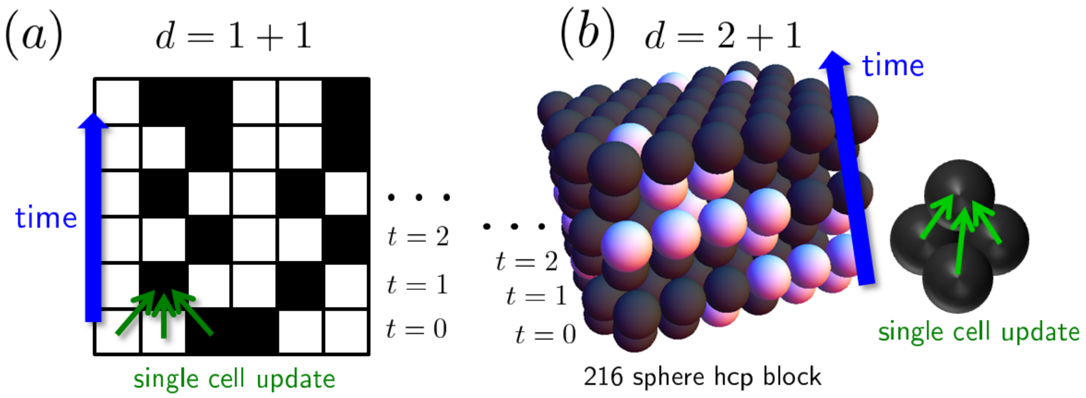


Figure D.1.1: (Color online) (a) A simulation of a small two-dimensional range expansion with a flat front extending along the horizontal direction. A single cell update is shown by the three green, thin arrows, with update rules given by Eq. (2) in the main text. Each horizontal row of squares in the lattice represents the state of the population front after t generations. (b) A three-dimensional flat front simulation. The generations are triangular lattices stacked on top of each other to form a hexagonal close-packed (hcp) array. A single cell update in the array is again indicated by the green arrows [probabilities given by Eq. (2) in main text].

then $p \approx 1/3 + \alpha$. Hence, we can model the domain walls as random walkers with an effective diffusivity $D_{\text{eff}} \approx p\ell^2/\tau_g \approx 1/3 + \alpha$, where ℓ is the lattice spacing, τ_g the generation time, and we set $\ell = \tau_g = 1$ for convenience. We can also perform a scaling analysis of the average interface density $\langle \rho(t) \rangle$ for various offsets $\tau = \alpha - \alpha_c$. We find the scaling function for $d = 1 + 1$ shown in Fig. D.1.2. The bottom branch of the scaling function is consistent with the pair-annihilation dynamics in the genetic demixing regime that yields $\langle \rho(t) \rangle \sim t^{-1/2}$. The data collapse is also consistent with the expected DP2 scaling results for the interface density decay exponent $\theta_{\text{DP2}} \approx 0.285(5)$ and transverse correlation length exponent $\nu_{\parallel} \approx 3.22(6)$ [103].

The dynamics of three dimensional range expansions resembles an overdamped Ising model below T_c with a nonconserved order parameter for $\alpha = \beta < 0$. The coarsening in the system is driven by a dynamically generated surface tension, calculated in Fig. 3

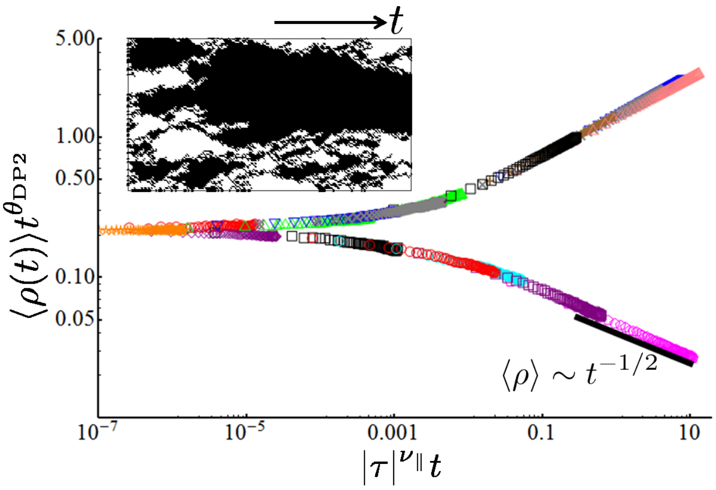


Figure D.1.2: (Color online) The scaling function for the interface density (averaged over at least 600 runs) near the DP2 transition ($d = 1+1$) along the line $\alpha = \beta$ for flat front sizes $L \gtrsim 2000$. Different sets of colored points are different offsets: $\tau = \alpha - \alpha_c = \pm 0.001, \pm 0.002, \ldots$. The data collapse is consistent with expected DP2 results $\theta_{\mathrm{DP2}} \approx 0.285(5)$ and $\nu_\parallel \approx 3.22(6)$ [103]. The lower branch describes the genetic demixing regime, characterized by the pair-annihilation dynamics of diffusing domain walls. The inset shows the population for $\alpha = \beta = \alpha_c \approx 0.1242$ at $t = 200$ generations and an $L = 100$ system size.

in the main text using an initially straight, long interface between a black and white cell domain. Another interesting initial condition is a black cell "droplet" in a sea of white cells [152]. Example evolutions for various α and β are shown in Fig. D.1.3. The dynamical surface tension shrinks the droplet, similar to non-conservative dynamics in an Ising model quenched below the critical temperature. However, the droplet dissolves away as $\alpha = \beta$ approaches 0 from negative values and becomes positive. Interestingly, a related model studied in Ref. [159] [based on the discretization of Eq. (1) in the main text] exhibits a transition from an Ising model to a voter model coarsening in the presence of strong noise for $\alpha = \beta < 0$. We did not observe such a transition in our model and the voter model coarsening seems to occur only for $\alpha = \beta = 0$.

To find the shape of the mutualistic "wedge" (see Fig. 1 in main text) near the DP2

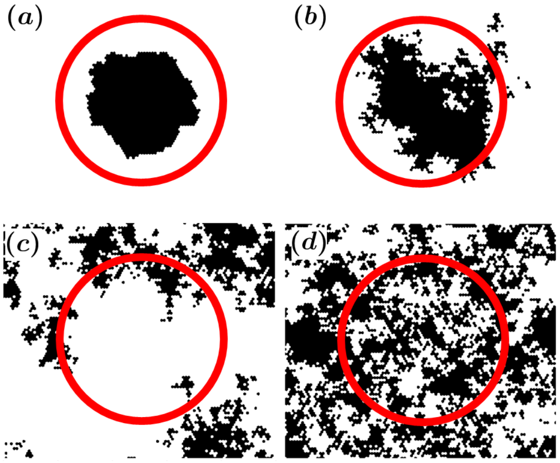


Figure D.1.3: (Color online) A black droplet with initial radius $R = 30$ cells (indicated by thick red circle) is placed in a white cell population. The droplet is then evolved to $t = 5000$ generations using the following parameters: (a) $\alpha = \beta = -0.2$ (b) $\alpha = \beta = -0.02$ (c) $\alpha = \beta = 0$ (d) $\alpha = \beta = 0.02$.

point for $d = 1 + 1$ and near the voter model point for $d = 2 + 1$, we find the points (α_i, β_i) $(i = 1, 2, \ldots, N)$ in the (α, β)-plane (near the DP2 and voter model points) for which the interface density $\langle \rho(t) \rangle$ decays at long times t with the expected directed percolation exponent $\langle \rho(t) \rangle \sim t^{-\theta_{\mathrm{DP}}}$, with $\theta_{\mathrm{DP}} \approx 0.159464$ for $d = 1 + 1$ and $\theta_{\mathrm{DP}} \approx 0.452$ for $d = 2 + 1$ [95]. In Fig. D.1.4, we show these $N \approx 10$ points in the (r, s)-plane (by transforming to $r_i = \alpha_i - \beta_i$ and $s_i = \alpha_i + \beta_i$) and fit them to the expected functional forms discussed in the main text. We find the crossover exponent $\phi \approx 1.9(1)$ by fitting. We estimate the error in ϕ by monitoring the variation in the effective exponent $\phi_e(r_i)$ (see, e.g., Ref. [158]) for each point (α_i, β_i) as we approach the DP2 transition. The exponent is given by

$$\phi_e(r_i) \equiv \frac{\ln[s_c(r_i) - s_c(0)] - \ln[s_c(r_{i-1}) - s_c(0)]}{\ln(r_i) - \ln(r_{i-1})}. \tag{D.1}$$

The same analysis can be applied to the rough front model, as shown in Fig. D.1.4.

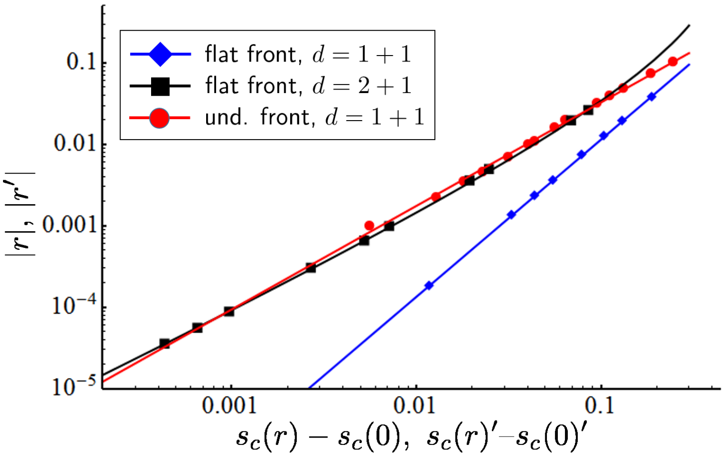


Figure D.1.4: (Color online) We track the shape of the mutualistic phase boundary by finding the points in the (r, s)-plane ($s = \alpha + \beta$ and $r = \alpha - \beta$) for which the average interface density $\langle\rho(t)\rangle$ decays at long times t with a characteristic DP power law $\langle\rho(t)\rangle \sim t^{-\theta_{\mathrm{DP}}}$ (primes for undulating fronts). We do this for the flat front models for both $d = 1+1$ near the special DP2 point (blue diamonds) and for $d = 2+1$ dimensions near the voter model point (black squares). The lines through the points are fits to the functions expected from cross-over scaling: $|r| = A_{2\mathrm{D}}[s_c(r) - s_c(0)]^{\phi}$ for $d = 1+1$ and $|r| = A_{3\mathrm{D}}\, s_c(r)/|\ln[s_c(r)/s_0]|$ for $d = 2+1$. The fitting parameters are $A_{2\mathrm{D}} \approx 0.970$, $A_{3\mathrm{D}} \approx 0.582$, $s_0 \approx 0.551$, and $\phi \approx 1.93$. $A_{2\mathrm{D}}$, $A_{3\mathrm{D}}$, and s_0 are non-universal constants and ϕ is a cross-over exponent. Similarly, the red circles are points in the (r', s')-plane (for the undulating front model in $d = 1+1$) near the "rough DP2" transition for which $\langle\rho(t)\rangle \sim t^{-\theta'_{\mathrm{DP}}}$, where $\theta'_{\mathrm{DP}} \approx 0.3125$ is the characteristic "rough DP" exponent found in Ref. [94]. The red line through the red circles is a fit to $|r'| = A'_{2\mathrm{D}}[s'_c(r') - s'_c(0)]^{\phi'}$. We find the non-universal amplitude $A'_{2\mathrm{D}} \approx 0.607$ and the crossover exponent $\phi' \approx 1.27$ by fitting.

Using the effective exponent technique, we find $\phi' \approx 1.27(15)$.

D.2 Undulating Front Models

It was convenient to implement the undulating front models on a triangular lattice and on a hexagonal close-packed lattice in two and three dimensions, respectively. Figure D.2.1 shows how cells are updated in the model. Note that the interaction strengths α' and β' can be negative, but they are bounded from below to ensure that each cell has a positive

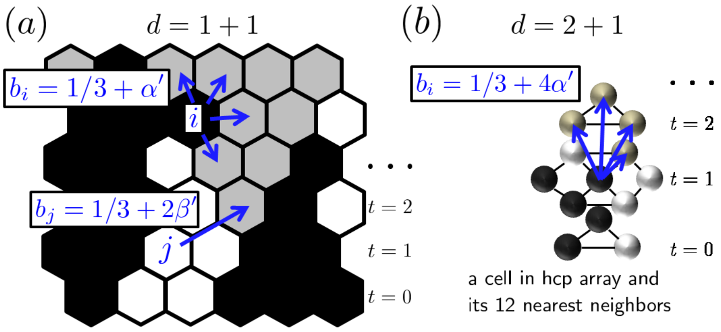


Figure D.2.1: (Color online) (*a*) A two-dimensional, undulating front simulation is implemented on a triangular lattice of potential cell positions. The front grows upward, with unoccupied space available to the black and white organisms indicated in grey. Two examples of the cell growth steps with rates b_i [calculated using Eq. (5) in the main text] are shown. Upon selecting a cell to divide, we choose a random empty (grey) nearest neighbor site to place the daughter cell (blue arrows from cell i). If one is unavailable, a random empty next nearest neighbor site is selected (blue arrow from cell j). As discussed in the main text, the time t (measured in generations) is assigned by taking horizontal cuts through the lattice behind the undulating front. (*b*) A three-dimensional, undulating front model is implemented on the hcp lattice. We illustrate an update step using a single cell's 12 nearest neighbor sites. The cells have a birthrate b_i [Eq. (5) in the main text] calculated using the occupied nearest neighbor sites (black and white cells) and daughter cells are placed into empty nearest neighbor sites (four blue arrows into grey spheres). The background birth rates (without mutualism) are $1/3$ for all cells.

birthrate [see Eq. (5) in main text]. We show some sample two dimensional range expansion simulations in Fig. D.2.2 for various interaction strengths α' and β'. Note the enhanced roughness near the DP transition line and for antisymmetric mutualism ($\alpha' \neq \beta'$ in the mixed phase) in Fig. D.2.2(*c*) and (*d*).

As discussed in the main text, the dynamics at $\alpha' = \beta' = 0$ for $d = 1 + 1$ is easier to understand as the front roughness is insensitive to the evolutionary dynamics at this point. The domain interface dynamics inherit the fluctuations of the rough population front and perform super-diffusive random walks. Figure D.2.4(*a*) shows an example of the population evolution. Interestingly, we see in Fig. D.2.4(*b*) that the scaling function associated with the heterozygosity can be fit to the same error function form as in the

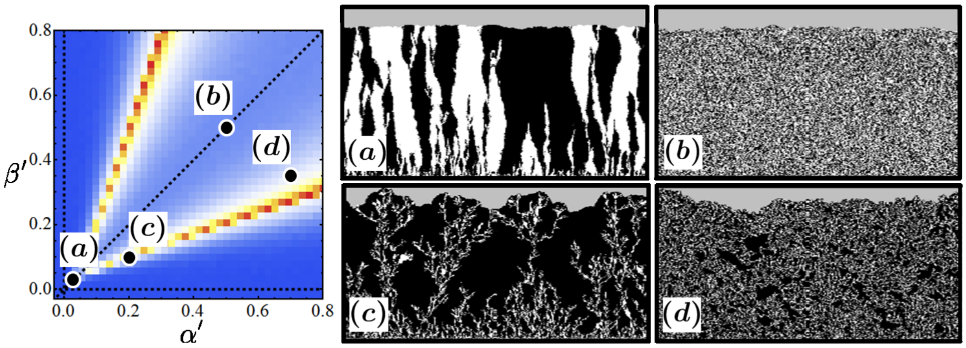


Figure D.2.2: (Color online) A map of the front roughness at time $t = 2500$ generations for an undulating front in $d = 1 + 1$ dimensions with a system size $L = 1000$ is shown on the left [see Fig. 4 in main text for the color bar scale]. In (a)-(d) we show sample evolutions of the system after about 500 generations at the indicated regions of the map. The simulation parameters are: (a) The DP2 critical point $\alpha' = \beta' = \alpha'_c = 0.0277$. ($b$) Symmetric mutualistic phase at $\alpha' = \beta' = 0.5$ (c) Near a rough DP transition at $\alpha' = 0.2$ and $\beta' = 0.1$. An all black cell front is the absorbing state. (d) An asymmetric mutualistic regime with $\alpha' = 0.7$ and $\beta' = 0.35$ (black is somewhat favored).

flat front case in the demixing regime [Eq. (3) in main text] if we replace the scaling variable $x/(8D_{\text{eff}}t)^{1/2}$ by $x/(Dt)^{2/3}$, where D is an effective "super-diffusion" coefficient.

The scaling function for the average interface density $\langle\rho(t)\rangle$ along $\alpha' = \beta'$, shown in Fig. D.2.3, is consistent with an exponent $\theta'_{\text{DP2}} \approx 0.61(1)$ describing the decay of $\langle\rho(t)\rangle$ at the rough DP2 point and a transverse correlation exponent $\nu'_{\|} \approx 2.8(2)$. Note that the flat front results $\theta_{\text{DP2}} \approx 0.285(5)$ and $\nu_{\|} \approx 3.22(6)$ are markedly different (see Fig. D.1.2 and Ref. [103]). Also, the bottom branch of the scaling function in Fig. D.2.3 confirms the rough voter model dynamics, $\langle\rho(t)\rangle \sim t^{-2/3}$, in the genetic demixing regime $\alpha' = \beta' < \alpha'_c$.

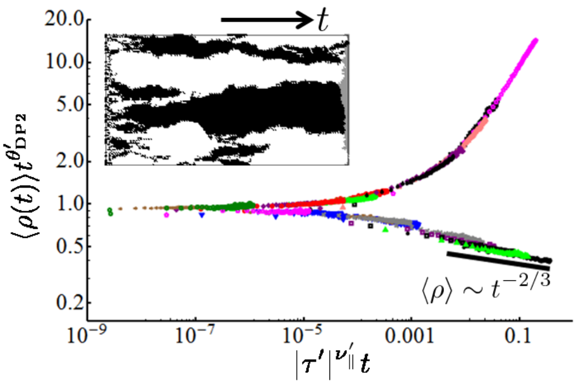


Figure D.2.3: (Color online) The scaling function for the interface density (averaged over at least 150 runs) near the rough DP2 transition ($d = 1 + 1$) along the line $\alpha' = \beta'$ for system sizes $L \gtrsim 2000$. Different sets of colored points are various offsets: $\tau' = \alpha' - \alpha'_c = \pm 0.001, \pm 0.002, \ldots$. The data collapse is consistent with the interface density decay exponent $\theta'_{\mathrm{DP2}} \approx 0.61(1)$ and correlation length exponent $\nu'_{\parallel} \approx 2.8(2)$. The upper branch of the function describes the mutualistic regime. The lower branch describes the genetic demixing regime. The inset shows the population for $\alpha' = \beta' = \alpha'_c$ at $t = 200$ generations and an $L = 100$ system size.

D.3 The Rough DP2 Transition in $d = 1 + 1$ Dimensions

The critical exponents along the two rough DP transition lines in $d = 1 + 1$ dimensions are consistent with those calculated by Frey et al. for a different model exhibiting a rough DP transition [94]. The exceptional DP2 transition point where the two lines meet has markedly different scaling properties, however. Unlike the rough DP transition, the characteristic widths $\xi_{\perp}$ and lengths $\xi_{\parallel}$ of genetic domains (in the space-like and time-like directions, respectively) diverge with different exponents $\xi_{\perp} \sim (\tau')^{-\nu'_{\perp}}$ and $\xi_{\parallel} \sim (\tau')^{-\nu'_{\parallel}}$ as we approach the DP2 point and the offset τ' approaches zero: $\tau' \equiv \alpha' - \alpha'_c \to 0$ (with $\tau' > 0$). The exponents $\nu'_{\perp,\parallel}$ are calculated by looking at all compact, black and white clusters with finite sizes for various points inside the mutualistic phase, i.e., at various offsets $\tau' \equiv \alpha' - \alpha'_c > 0$ from the DP2 point with $\beta' = \alpha'$. Following Ref. [170], we

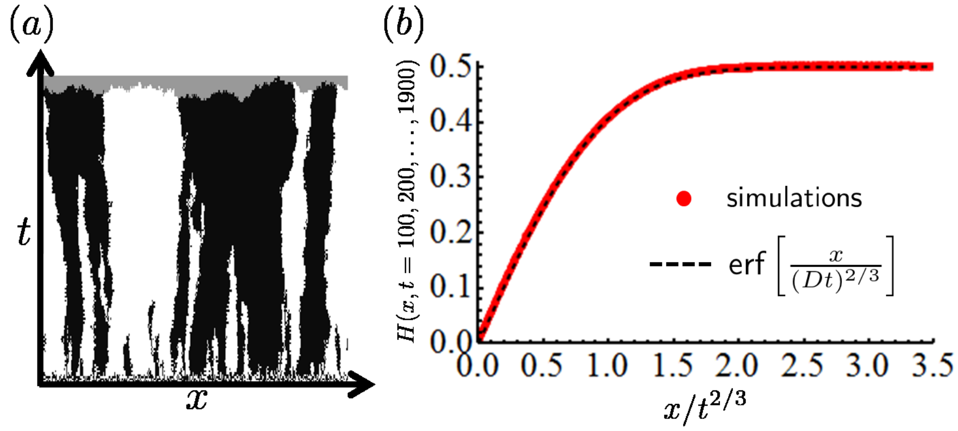


Figure D.2.4: (Color online) (a) Undulating front model simulation in $d = 1 + 1$ dimensions at $\alpha' = \beta' = 0$ (system size $L = 300$) evolved for about 150 generations. The space-like and time-like axes (x and t, respectively) are also shown. (b) We collapse the heterozygosity $H(x, t)$ for $\alpha' = \beta' = 0$ at different times $t = 100, 200, \ldots, 1900$ generations and positions x onto a single universal function using the scaling variable $x/t^{2/3}$, where $2/3 = 1/\tilde{z}$ is related to the dynamic exponent $\tilde{z}$ of the KPZ model [93, 160]. The conjectured analytic form of the scaling function is shown by the black dashed line, where we find $D \approx 1.1$ by fitting. The simulation uses an $L = 2000$ system size and $H(x, t)$ is averaged over 1600 runs.

calculate

$$\xi_{\parallel} = \frac{\sum_i M_i \ell_{\parallel,i}}{\sum_i M_i} \quad \text{and} \quad \xi_{\perp} = \frac{\sum_i M_i \ell_{\perp,i}}{\sum_i M_i}, \tag{D.2}$$

where we sum over all compact clusters i. Each such cluster i contains M_i cells (and hence has "mass" M_i), and has maximum extensions in the space-like and time-like directions given by $\ell_{\perp,i}$ and $\ell_{\parallel,i}$, respectively. All lengths are measured in units of the cell diameter.

We identify compact clusters in the simulations by using a sequential, two-pass image segmentation algorithm [171]. In the first pass, we scan and label the cells along the space-like direction [left to right in Fig. D.2.4(a)] and advance one row at a time along the time-like direction [from bottom to top in Fig. D.2.4(a)]. As we scan, connected regions of cells are partially identified by assigning cells in each region the same label.

We modify the usual segmentation algorithm by testing each cell for connectivity only if it is the same color as the two cells labelled in the previous row and the two cells to adjacent to it in its own row (taking into account the periodic boundary conditions along the space-like direction). In the second pass, we complete the segmentation by assigning the same label to all cells that end up part of a connected region after the first pass (see Ref. [171] for more details). We then treat all labelled clusters with masses $M_i \leq 3$ or $\ell_{\parallel,i} = 1$ as the "mutualistic phase" and discard them when performing the summation in Eq. (D.2). The leftover clusters are shown in different colors in the inset of Fig. D.3.1. With a few pathological cases at contorted domain walls, the mutualistic phase will be all the cells which neighbor at least three cells of the opposite type.

After calculating $\xi_{\perp,\parallel}$ for various front sizes L and offsets τ', we perform a data collapse in Fig. D.3.1 and find $\nu'_\perp = 1.8(2)$ and $\nu'_\parallel = 2.6(2)$. Note that within the error margin, $\nu'_\perp$ is the same as the flat front result, $\nu_\perp \approx 1.83(3)$ [95]. However, the front undulations suppress $\nu'_\parallel$ relative to flat fronts (for which $\nu_\parallel \approx 3.22(6)$ [95]). The errors are estimated by varying the exponents and estimating the range of values for which we still find a data collapse. In the rough DP case studied in Ref. [94], $\nu'_\parallel = \nu'_\perp \approx 1.6(1)$. Thus, the clusters have a markedly different, more asymmetric shape near the rough DP2 transition. This feature makes sense as the more strongly enhanced front roughness along the DP lines [see Fig. D.2.2(c)] results in a stronger coupling between the dynamics in the time-like and space-like directions. At the DP2 point, this coupling is weaker but still significant: the exponents $\nu'_\perp$ and $\nu'_\parallel$ are closer to each other than their flat front counterparts.

It is also possible to study the scaling of the front roughness. As discussed in Ref. [94], for a fixed population front size L and time t, we expect the average interface

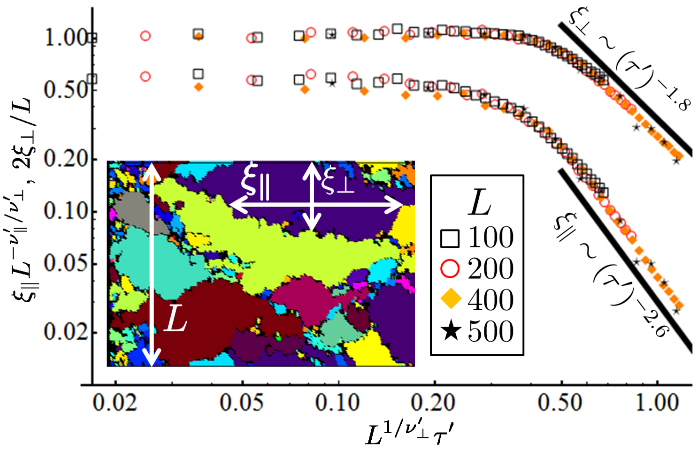


Figure D.3.1: (Color online) We measure the correlation lengths describing compact domains ($\xi_\parallel$ and $\xi_\perp$ in inset) using Eq. (D.2) for various front sizes L along the time-like and space-like directions as we approach the rough DP2 transition at $\alpha' = \beta' = \alpha'_c \approx 0.0277$ for various offsets $\tau' \equiv \alpha' - \alpha'_c > 0$ along the line $\alpha' = \beta'$. We average over at least 150 runs. The inset shows individual compact domains that we sum over in Eq. (D.2) in different colors. The time-like direction is horizontal. Our algorithm for determining compact domains of black and white cells excises slivers of mutualistic mixing at black/white boundaries. This is done using a modified image segmentation algorithm as described in this supplement. The data for various L are collapsed using finite size scaling [172] and we find power law divergences in the lengths $\xi_{\parallel,\perp}$ for large L.

height fluctuation $\langle \Delta h(t,L)\rangle$ to obey

$$\langle \Delta h(t,L)\rangle \sim \begin{cases} t^{\gamma} & t \ll t_\times \\ L^{\tilde{\alpha}} & t \gg t_\times \end{cases}, \tag{D.3}$$

where γ is the growth exponent, $\tilde{\alpha}$ is the roughness exponent, and the crossover time $t_\times$ satisfies $t_\times \sim L^{\tilde{z}}$ with dynamic exponent $\tilde{z} = \tilde{\alpha}/\gamma$. When $\alpha' = \beta' = 0$, we expect the dynamics to fall into the KPZ universality class, for which $\gamma = 1/3$ and $\tilde{\alpha} = 1/2$ (so $\tilde{z} = 3/2$). We find consistent results $\gamma \approx 0.31(2)$ and $\tilde{\alpha} \approx 0.48(5)$ for $\alpha' = \beta' = 0$. We extrapolated the value and error in γ at long times by using the effective exponent technique mentioned above. The exponent $\tilde{\alpha}$ was calculated by varying L and measuring the final, saturated roughness at long times. At the DP2 point, we find similar results $\gamma_{\mathrm{DP2}} \approx 0.31(1)$ and $\tilde{a}_{\mathrm{DP2}} \approx 0.48(4)$. Hence, the dynamics also appear consistent with the KPZ universality class. However, more extensive simulations are necessary to verify that the coupling of the DP2 evolutionary dynamics to the front undulations do not change the KPZ scaling. Note that the front undulation scaling along the DP phase transition lines is dramatically different from KPZ scaling, as discovered in Ref. [94]. It would also be interesting to check if the interfaces between black and white domains inherit the front fluctuations in the same way as the neutral case at $\alpha' = \beta' = 0$, so that the relation $\theta'_{\mathrm{DP2}} = 1/\tilde{z}_{\mathrm{DP2}}$ holds. As discussed in the main text, we found $\theta'_{\mathrm{DP2}} \approx 0.61(1)$. Hence, if the domain walls also inherit the front undulations, it is possible that the front roughness actually has a slightly larger dynamic exponent than in the KPZ class, with $\tilde{z}_{\mathrm{DP2}} \approx 1.64$ (instead of $\tilde{z} = 3/2$).

Bibliography

[1] H. Vlamakis, Y. Chai, P. Beauregard, R. Losick, and R. Kolter, Nature Rev. Microbiol. **11**, 157 (2013).

[2] O. Hallatschek, P. Hersen, S. Ramanathan, and D. R. Nelson, Proc. Nat. Acad. Sci. **104**, 19926 (2007).

[3] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, 2001).

[4] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, Oxford, 2007), 7th ed.

[5] P. Yu, M. Mustata, J. J. Turek, P. M. W. French, M. R. Melloch, and D. D. Nolte, Appl. Phys. Lett. **83**, 575 (2003).

[6] M. V. Smoluchowski, Z. Phys. **17**, 557 (1916).

[7] G. H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam, 1994).

[8] J. Folkman and M. Hochberg, J. Exp. Med. **138**, 745 (1973).

[9] O. Hallatschek and D. R. Nelson, Evolution **64**, 193 (2010).

[10] P. A. P. Moran, Math. Proc. of the Camb. Phil. Soc. **54**, 60 (1958).

[11] J. F. Crow and M. Kimura, *An Introduction To Population Genetics Theory* (Harper & Row, New York, 1970).

[12] T. Duty, Ph.D. thesis, Virginia Tech (2000).

[13] K. S. Korolev, M. Avlund, O. Hallatschek, and D. R. Nelson, Rev. Mod. Phys. **82**, 1691 (2010).

[14] N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007), 3rd ed.

[15] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985), 2nd ed.

[16] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge University Press, Cambridge, 2010).

[17] O. Hallatschek and K. S. Korolev, Phys. Rev. Lett. **103**, 108103 (2009).

[18] R. W. Taylor, M. J. Barron, G. M. Borthwick, A. Gospel, P. F. Chinnery, D. C. Samuels, G. A. Taylor, S. M. Plusa, S. J. Needham, L. C. Greaves, et al., J. Clin. Invest. **112**, 1351 (2003).

[19] N. Barker, J. H. van Es, J. Kuipers, P. Kujala, M. van den Born, M. Cozijnsen, A. Haegebarth, J. Korving, H. Begthel, P. J. Peters, et al., Nature **449**, 1003 (2007).

[20] H. J. Snippert, L. G. van der Flier, T. Sato, J. H. van Es, M. van den Born, C. Kroon-Veenboer, N. Barker, A. M. Klein, J. van Rheenen, B. D. Simons, et al., Cell **143**, 134 (2010).

[21] C. Lopez-Garcia, A. M. Klein, B. D. Simons, and D. J. Winton, Science **330**, 822 (2010).

[22] M. J. I. Müller, B. I. Neugeboren, D. R. Nelson, and A. W. Murray, *Genetic drift opposes mutualism during spatial population expansion* (2014).

[23] B. Momeni, K. A. Brileya, M. W. Fields, and W. Shou, eLife **2**, e00230 (2013).

[24] K. S. Korolev and D. R. Nelson, Phys. Rev. Lett. **107**, 088103 (2011).

[25] M. Johnston, Trends Genet. **15**, 29 (1999).

[26] E. Reigenberger, E. Boles, and M. Ciriacy, Eur. J. Biochem. **245**, 324 (1997).

[27] M. M. C. Meijer, J. Boonstra, A. J. Verkleij, and C. T. Verrips, Biochim. Biophys. Acta **1277**, 209 (1996).

[28] F. Kreuzer, Respir. Physiol. **9**, 1 (1970).

[29] F. Bronner, D. Pansu, and W. D. Stein, Am. J. Physiol. Gastrointest. Liver Physiol. **250**, G561 (1986).

[30] A. Carruthers, Physiol. Rev. **70**, 1135 (1990).

[31] M. J. Grimson and G. C. Barker, Phys. Rev. E **49**, 1680 (1994).

[32] R. Willaert and G. Baron, Appl. Microbiol. Biotechnol. **39**, 347 (1993).

[33] M. B. Cassidy, H. Lee, and J. T. Trevors, J. Ind. Microbiol. **16**, 79 (1996).

[34] K. S. K. I. G. Najafpour, H. Younesi, Bioresource Technol. **92**, 251 (2004).

[35] H. C. Berg and E. M. Purcell, Biophys. J. **20**, 193 (1977).

[36] H. C. Berg, *Random Walks in Biology* (Princeton University Press, New Jersey, 1993).

[37] R. Zwanzig, Proc. Natl. Acad. Sci. USA **87**, 5856 (1990).

[38] R. Zwanzig and A. Szabo, Biophys. J. **60**, 671 (1991).

[39] R. I. Cukier, J. Stat. Phys. **42**, 69 (1986).

[40] D. F. Calef and J. M. Deutch, Ann. Rev. Phys. Chem. **34**, 493 (1983).

[41] S. Torquato, *Random Heterogeneous Materials: Microstructure and macroscopic properties* (Springer Science & Business Media, New York, 2001).

[42] A. Natarajan and F. Srienc, Metab. Eng. 1 **1**, 320 (1999).

[43] B. Antuna and M. A. Martinez-Anaya, Inter. J. of Food Microbiol. **18**, 191 (1993).

[44] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer-Verlag, Berlin, 1989).

[45] S. V. G. Menon and D. C. Sahni, Phys. Rev. A **32**, 3832 (1985).

[46] K. Asano, *Mass Transfer: From fundamentals to modern industrial applications* (Wiley-VCH, Weinheim, Germany, 2006).

[47] C. Guthriw and G. R. Fink, *Guide to Yeast Genetics and Molecular and Cell Biology*, vol. 350 (Gulf Professional Publishing, 2002).

[48] L. G. Longsworth, Journal of the American Chemical Society **75**, 5705 (1953).

[49] H. V. Urk, P. R. Mark, W. A. Scheffers, and J. P. V. Dijken, Yeast **4**, 283 (1988).

[50] D. E. Nelson and K. I. Young, J. Bacteriol. **182**, 1714 (2000).

[51] D.-S. Kim and H. S. Fogler, Appl. Microbiol. Biotechnol. **52**, 839 (1999).

[52] P. M. Richards, J. Phys. Chem. **85**, 3520 (1986).

[53] R. F. Kayser and J. B. Hubbard, Phys. Rev. Lett. **51**, 79 (1983).

[54] M. Tokuyama and R. I. Cukier, J. Chem. Phys. **76**, 6202 (1982).

[55] J. Krüger, Physica A **169**, 393 (1990).

[56] I. D. C. Forney, Master's thesis, MIT, Cambridge, MA (2007).

[57] C. A. Miller and S. Torquato, Phys. Rev. B **40**, 7101 (1989).

[58] L. Zhing and Y. C. Chiew, J. Chem. Phys. **93**, 2658 (1990).

[59] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, Adv. in Comp. Math. **5**, 329 (1996).

[60] B. J. M. Hannoun and G. Stephanopoulos, Biotechnol. Bioeng. **28**, 829 (1986).

[61] S. Torquato, T. M. Truskett, and P. G. Debenedetti, Phys. Rev. Lett. **84**, 2064 (2000).

[62] M. Muthukumar and R. I. Cukier, J. Stat. Phys. **26**, 453 (1981).

[63] J. R. Lebenhaft and R. Kapral, J. Stat. Phys. **20**, 25 (1979).

[64] R. I. Cukier and K. F. Freed, J. Chem. Phys. **78**, 2573 (1983).

[65] R. I. Cukier, J. Phys. Chem. **87**, 582 (1983).

[66] J. R. Kalnin, E. A. Kotomin, and J. Maier, J. Phys. Chem. Solids **63**, 449 (2002).

[67] B. Jönsson, H. Wennerström, P. G. Nilsson, and P. Linse, Colloid & Polymer Sci. **264**, 77 (1986).

[68] (2008), COMSOL Multiphysics User's Guide, Version 3.5a.

[69] C. H. Bennett, J. Appl. Phys. **43**, 2727 (1972).

[70] J. H. Koschwanez, K. R. Foster, and A. W. Murray, PloS Biol **9**, e1001122 (2011).

[71] M. J. Brauer, C. Huttenhower, E. M. Airoldi, R. Rosenstein, J. C. Matese, D. Gresham, V. M. Boer, O. G. Troyanskaya, and D. Botstein, Mol Biol Cell **19**, 352 (2008).

[72] J. A. Barnett and K.-D. Entian, Yeast **22**, 835 (2005).

[73] C. Mateus and S. V. Avery, Yeast **16**, 1313 (2000).

[74] B. Nguyen, A. Upadhyaya, A. van Oudenaarden, and M. P. Brenner, Biophys. J. **86**, 2740 (2004).

[75] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., Nat Meth **9**, 676 (2012).

[76] S. Preibisch, S. Saalfeld, and P. Tomancak, Bioinformatics **25**, 1463 (2009).

[77] C. H. Li and P. K. S. Tam, Pattern Recogn. Lett. **19**, 771 (1998), URL `http://dx.doi.org/10.1016/S0167-8655(98)00057-9`.

[78] K. S. Korolev, J. B. Xavier, D. R. Nelson, and K. R. Foster, Amer. Nat. **178**, 538 (2011).

[79] L. M. F. Merlo, J. W. Pepper, B. J. Reid, and C. C. Maley, Nat. Rev. Cancer **6**, 924 (2006).

[80] J. A. Cuesta, J. Aguirre, J. A. Capitán, and S. C. Manrubia, Phys. Rev. Lett. **106**, 028104 (2011).

[81] J. A. Capitán, J. A. Cuesta, S. C. Manrubia, and J. Aguirre, PLoS ONE **6**, e23358 (2011).

[82] G. Hewitt, Nature **405**, 907 (2000).

[83] A. Templeton, Nature **416**, 45 (2002).

[84] C. Parmesan, N. Ryrholm, C. Stefanescu, J. K. Hill, C. D. Thomas, H. Descimon, B. Huntley, L. Kaila, J. Kullberg, T. Tammaru, et al., Nature **399**, 579 (1999).

[85] A. T. Hitch and P. L. Leberg, Conservation Biology **21**, 534 (2007).

[86] B. J. Phillips, G. P. Brown, J. K. Webb, and R. Shine, Nature **439**, 803 (2006).

[87] M. E. Gray, T. W. Sappington, N. J. Miller, J. Moeser, and M. O. Bohn, Annu. Rev. Entomol. **54**, 303 (2009).

[88] M. Kimura and G. Weiss, Genetics **49**, 561 (1964).

[89] W. J. Ewens, *Mathematical Population Genetics*, vol. I (Springer, New York, 2004), 2nd ed.

[90] J. D. Murray, *Mathematical Biology*, vol. I - An Introduction (Springer-Verlag, Berlin, 2001).

[91] R. A. Fisher, Annals of Eugenics **7**, 355 (1937).

[92] K. S. Korolev, M. J. I. Müller, N. Karohan, A. W. Murray, O. Hallatschek, and D. R. Nelson, Phys. Biol. **9**, 026008 (2012).

[93] Y. Saito and H. Müller-Krumbhaar, Phys. Rev. Lett. **74**, 4325 (1995).

[94] J.-T. Kuhr, M. Leisner, and E. Frey, New J. Phys. **13**, 113013 (2011).

[95] H. Hinrichsen, Adv. in Phys. **49**, 815 (2000).

[96] M. Lynch, R. Bürger, D. Butcher, and W. Gabriel, Journal of Heredity **84**, 339 (1993).

[97] Y. Timofeeva and S. Coombes, Phys. Rev. E **70**, 062901 (2004).

[98] O. Hallatschek, P. Hersen, S. Ramanathan, and D. R. Nelson, PNAS **104**, 19926 (2007).

[99] S. Lübeck, Int. J. Mod. Phys. B **18**, 3977 (2004).

[100] E. Domany and W. Kinzel, Phys. Rev. Lett. **53**, 311 (1984).

[101] T. M. Liggett, *Interacting Particle Systems* (Springer-Verlag, New York, 1985).

[102] S. Lübeck, J. Stat. Mech. p. P09009 (2006).

[103] M. Henkel, H. Hinrichsen, and S. Lübeck, *Non-Equilibrium Phase Transitions*, vol. I - Absorbing Phase Transitions (Springer Science, The Netherlands, 2008).

[104] M. Rubinstein and D. R. Nelson, Phys. Rev. B **26**, 6254 (1982).

[105] R. J. Glauber, J. Math. Phys. **4** (1963).

[106] S. A. Palumbo, M. G. Johnson, V. T. Rieck, and L. D. Witter, J. Gen. Microbiol. **66**, 137 (1971).

[107] B. F. Gray and N. A. Kirwan, Biophys. Chem. **1**, 204 (1974).

[108] B. O. Neill, *Elementary Differential Geometry* (Academic Press, San Diego, 1997), 2nd ed.

[109] A. Ali and S. Grosskinsky, Adv. Complex Syst. **13**, 249 (2010).

[110] E. B. Saff and A. D. Snider, *Fundamentals of Complex Analysis with Applications to Engineering and Science* (Pearson Education, New Jersey, 2003), 3rd ed.

[111] M. Eigen, J. McCaskill, and P. Schuster, J. Phys. Chem. **92**, 6881 (1988).

[112] K. Jain and J. Krug, in *Structural Approaches to Sequence Evolution*, edited by U. Bastolla, M. Porto, H. E. Roman, and M. Vendruscolo (Springer Berlin Heidelberg, 2007), Biological and Medical Physics, Biomedical Engineering, pp. 299–339.

[113] J. J. Bull, L. A. Meyers, and M. Lachmann, PLoS Comput Biol **1**, e61 (2005).

[114] M. J. Kearny and S. N. Majumdar, J. Phys. A: Math. Gen. **38**, 4097 (2005).

[115] B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. D. Jr., and K. W. Kinzler, Science **339**, 1546 (2013).

[116] L. A. Kunz-Schughart, Cell Biol. Int. **23**, 157 (1999).

[117] M. T. Santini and G. Rainaldi, Pathobiology **67**, 148 (1999).

[118] F. Hirschhaeuser, H. Menne, C. Dittfeld, J. West, W. Mueller-Klieser, and L. A. Kunz-Schughart, J. Biotechnol. **148**, 3 (2010).

[119] S. M. Weis and D. A. Cheresh, Nat. Med. **17**, 1359 (2011).

[120] A. Shirinifard, J. S. Gens, B. L. Zaitlen, N. J. Poplawski, M. Swat, and J. A. Glazier, PLoS ONE **4**, e7190 (2009).

[121] T. Alarcón, H. M. Byrne, and P. K. Maini, Multiscale Model. Sim. **3**, 440 (2005).

[122] G. Cheng, J. Tse, R. K. Jain, and L. L. Munn, PLoS ONE **4**, e4632 (2009).

[123] F. Montel, M. Delarue, J. Elgeti, L. Malaquin, M. Basan, T. Risler, B. Cabane, D. Vignjevic, J. Prost, G. Cappello, et al., Phys. Rev. Lett. **107**, 188102 (2011).

[124] F. Montel, M. Delarue, J. Elgeti, D. Vignjevic, G. Cappello, and J. Prost, New J. Phys. **14**, 055008 (2012).

[125] A. Brú, S. Albertos, J. L. Subiza, J. L. García-Asenjo, and I. Brú, Biophys. J. **85**, 2948 (2003).

[126] E. L. Stott, N. F. Britton, J. A. Glazier, and M. Zajac, Math. Comput. Model. **30**, 183 (1999).

[127] T. Antal, P. L. Krapivsky, and M. A. Nowak (2013), `arXiv:1308.1564`.

[128] J. Arvo, in *Graphics Gems III*, edited by D. Kirk (Academic Press Inc., London, 1992).

[129] C. R. Doering, C. Mueller, and P. Smereka, Physica A **325**, 243 (2003).

[130] E. A. Martens, R. Kostadinov, C. C. Maley, and O. Hallatschek, New J. Phys. **13**, 115014 (2011).

[131] E. Moro, Phys. Rev. Lett. **87**, 238303 (2001).

[132] S. Pigolotti, R. Benzi, P. Perlekar, M. H. Jensen, F. Toschi, and D. R. Nelson, Theor. Popul. Biol. **84**, 72 (2013).

[133] A. Ali, R. C. Ball, S. Grosskinsky, and E. Somfai, Phys. Rev. E **87**, 020102(R) (2013).

[134] A. Ali, R. C. Ball, S. Grosskinsky, and E. Somfai, J. Stat. Mech. **2013**, P09006 (2013).

[135] C. Escudero, J. Phys. A: Math. Theor. **46**, 355403 (2013).

[136] H. P. de Vladar and N. Barton, Trends in Ecology and Evolution **26**, 424 (2011).

[137] P. C. Martin, E. D. Siggia, and H. A. Rose, Phys. Rev. A **8**, 423 (1973).

[138] H. K. Janssen, Z. Phys. B Con. Mat. **23**, 377 (1976).

[139] C. de Dominicis, J. Phys. Colloques **37**, C1 (1976).

[140] H. K. Janssen, J. Phys.: Condens. Matter **17**, S1973 (2005).

[141] J. T. Cox and D. Griffeath, Ann. Prob. **14**, 347 (1986).

[142] G. Ódor, Rev. Mod. Phys. **76**, 663 (2004).

[143] A. Anderson and O. Hallatschek (2014), unpublished work.

[144] E. H. Wintermute and P. A. Silver, Mol. Syst. Biol. **6** (2010).

[145] J. D. Styrsky and M. D. Eubanks, Proc. R. Soc. B **274**, 151 (2007).

[146] C. R. Dickman, Trends Ecol. Evol. **7**, 194–197 (1992).

[147] R. A. Long and F. Azam, Appl. Environ. Microbiol. **67**, 4975 (2001).

[148] L. Dall'Asta, F. Caccioli, and D. Beghé, EPL **101**, 18003 (2013).

[149] L. Canet, H. Chaté, B. Delamotte, I. Dornic, and M. A. Muñoz, Phys. Rev. Lett. **95**, 100601 (2005).

[150] J. L. Cardy and U. C. Täuber, J. Stat. Phys. **90**, 1 (1998).

[151] O. A. Hammal, H. Chaté, I. Dornic, and M. A. Muñoz, Phys. Rev. Lett. **94**, 230601 (2005).

[152] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen, Phys. Rev. Lett. **87**, 045701 (2001).

[153] T. Roose, S. J. Chapman, and P. K. Maini, SIAM Rev. **49**, 179 (2007).

[154] S. Torquato, Phys. Biol. **8**, 015017 (2011).

[155] P. Grassberger, F. Krause, and T. von der Twer, J. Phys. A **17**, L105 (1984).

[156] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (World Scientific, New Jersey, 2006), 3rd ed.

[157] M. E. Fisher and D. R. Nelson, Phys. Rev. Lett. **32**, 1350 (1974).

[158] G. Ódor and N. Menyhárd, Phys. Rev. E **78**, 041112 (2008).

[159] D. I. Russell and R. A. Blythe, Phys. Rev. Lett. **106**, 165702 (2011).

[160] M. Kardar, G. Parisi, and Y.-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).

[161] D. Forster, D. R. Nelson, and M. J. Stephen, Phys. Rev. A **16**, 732 (1977).

[162] C. Vulin, J.-M. D. Meglio, A. B. Lindner, A. Daerr, A. Murray, and P. Hersen, Biophys. J. **106**, 2214 (2014).

[163] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, New Jersey, 1999), 3rd ed.

[164] S. K. Lucas, J. Comput. Appl. Math. **64**, 269 (1995).

[165] M. Muthukumar, J. Chem. Phys. **76**, 2667 (1982).

[166] C. O. Wilke, BMC Evol. Bio. **5** (2005).

[167] F. Vazquez and C. López, Phys. Rev. E **78** (2008).

[168] U. C. Täuber, *Critical Dynamics* (Cambridge University Press, Cambridge, 2014).

[169] V. F. Zaitsev, *Handbook of exact solutions for ordinary differential equations* (CRC Press, Boca Raton, Florida, 2003).

[170] A. Gelimson, J. Cremer, and E. Frey, Phys. Rev. E **87**, 042711 (2013).

[171] L. Shapiro and G. Stockman, *Computer Vision* (Prentice Hall, New Jersey, 2001).

[172] M. E. Fisher and M. N. Barber, Phys. Rev. Lett. **28**, 1516 (1972).