UCLA
UCLA Electronic Theses and Dissertations

Title

Population Genetics in a Single Organism: Models of Neurospora crassa Nuclear Dynamics

Permalink

https://escholarship.org/uc/item/1j6212xp

Author

Wang, Teng

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Population Genetics in a Single Organism:
Models of Neurospora crassa Nuclear Dynamics

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mathematics

by

Teng Wang

2015

ABSTRACT OF THE DISSERTATION

Population Genetics in a Single Organism:
Models of Neurospora crassa Nuclear Dynamics

by

Teng Wang

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2015

Professor Marcus Leigh Roper, Chair

In this thesis I will analyze the dynamics of two-populations of nuclei within a filamentous fungus. The models developed here are designed to directly represent recent experiments studying *Neurospora crassa* chimera created from two genetically different nuclear populations.

The focus will be on analyzing how dynamics within the network are affected by the geometry of the fungus network by building and validating spatial population genetic models. We start from simulating some intuitive models for nuclear population dynamics during fungal growth, then provide asymptotic and analytic insights on the dynamics involved, and the relationship between population dynamics and the branching structure of the network. Comparisons among the simulation, the simplified model, and existing experiment data will also be discussed.

The dissertation of Teng Wang is approved.

Janet S. Sinsheimer,

Luminita Aura Vese

Joseph M. Teran

Marcus Leigh Roper, Committee Chair

University of California, Los Angeles

2015

To my parents

Table of Contents

LIST OF FIGURES

List of Tables

I would like to thank my Chinese friends. There are quite a few, but I will list these names in particular: Feng Guan, Wenyun Yang, Changyong Yin, Wenbo Sheng. With them I share some of my happiest moments in recent years. And without them I cannot overcome a lot of difficulties during the same period.

Finally, I would like to thank my parents. Their love and support weighs more than everything else in the world for me. This thesis is dedicated to them.

VITA

2010	B.S. (Mathematics), Fudan University, Shanghai, China.
2012	M.A. (Mathematics), UCLA, Los Angeles, California.
2012-2015	Teaching Assistant, Department of Mathematics, UCLA.
2010-2015	Research Assistant, Department of Mathematics, UCLA.
2014	Summer intern in risk analytics, Western Asset Management, Pasadena, California.

PRESENTATIONS

"Population Genetics in a Single Organism: Models of *Neurospora crassa* Nuclear Dynamics", Poster session in Mathematical Modeling in Biomedicine: The Carol Newton Legacy Symposium, March 9, 2015.

CHAPTER 1

Introduction

The evolution of new lineages, species and groups is strongly affected by the spatial distributions of organisms. Organisms can only compete or cooperate with each other if they are close enough to interact.

Accordingly geographic isolation is known to be a motor for the evolution of new species ("callopalmy") [Gou02]. This is isolation, need not be absolute; organisms that fund themselves at the periphery of a population during a range expansion are more likely to move into the new territory and to populate it with offspring than organisms far from the periphery. This effect, which leads to newly founded territories not reflecting the diversity of the original population has been observed in model microbial systems [HN08] [KAH10] and forms the basis of efforts to reconstruct historical patterns of dispersal and emigration. [HKR98] [GFV00]

Realistic representation of spatial structure in population genetic models for evolution remains an unmet challenge. Although these models have had some success, most of them do not represent the spatial structure of the populations that they are intended to model.

Indeed for most wild populations there is both inadequate resolution of the current spacial distribution of its members (that is where the organisms are currently) and too much uncertainty about historical distributions (where organisms, or the lineages they are derived from) were in the past. Thus although toy models that incorporate spatial structure have been shown to exhibit surprising dynamics [OHL06], it is not known whether these dynamics occur in any real system. Recent efforts have focused on creating model microbial systems in which both the dispersal of organisms, which controls their spatial distributions, and the set

of neighbors that they interact with can be mapped and controlled. For example Gore et al. [GGR13] cultured yeast ecosystems in microwells, and simulated dispersal by transferring fixed numbers of cells between micro-wells. Hallatschek et al. grew E Coli bacteria on solid agar mapped cell lineages colonies to grow [KAH10].

However both sets of experiments are at several removes from dispersal in nature; since the natural abilities of organisms to disperse are supposed (in the experiments by growing cells in physically sperate micro-wells). In the Hallatschek experiments by removing the machinery such as flagella [BB91] [DTM05] [Ber08] or surfactant expression [ARK09] that would allow wild type cells to explore new environments.

Remarkably much of the recent work on model microbial systems was anticipated around 60 years ago by the pioneering microbiologist Guido Pontecorvo [PG44] [Pon59] [Pon75] who proposed studying these dynamics using chimeric fungi for an experimental model. Pontecorvo's experiments used two-component *Penicillium notatum* fungi, that harbored two different nuclear genotypes. He investigated how these genotypes tended to segregate(or sector) as the chimeric cells grew. Although he could measure differences in fitness(that is, rates of division) from the geometry of the sectors, he did not introduce models for these processes, likely because his method for distinguishing the two genotypes(based on the pigmentation of the spores that they produced) could not give a quantitative read on their relative abundances, and because at the time little was known about the motions of nuclei within living fungal cells.

More recently nuclear labeling using fluorescent proteins has made it possible to map nuclear movements within living cells.

Nevertheless this nuclear dynamics within chimera are an attractive system for building and validating spatial population genetic models, partly because no special treatments is needed to control dispersal of nuclei across cells(this is similar of nuclear movement occur in the lab as in nature). Nuclear dispersal patterns can also be manipulated genetically - for example in *Neurospora crassa*, which is the main model system for which models will be built here, hundreds of wild-type variants with different cell morphologies have been isolated

[Gla04] [GD06], while several single gene mutations are also known to alter morphology [FSJ05] [GJS00] [SRY10].

Moreover, single fungal cells often harbor genetically diverse nuclei. Internal genetic diversity created by mutation when nuclei divide [MR63] or more rarely when two different cells may fuse and exchange nuclei. [Bus14] [RET11]

Filamentous fungi are able to harbor genetically diverse nuclei within large multinucleate compartments. The ability to support multiple genomes within a single cell is thought to a key part of fungal life histories, since it provides a pathway to adapt to new hosts, changing environments and substrates, as well as to cope with mutations introduced during division of individual nuclei. However very little is known about whether or how this internal genetic diversity is maintained during fungal growth.

In particular it is explained in Chapter 2 that in any finite, initially diverse population, genetic drift(the random fluctuations in the proportions of different lineages) will eventually fix a single lineage, i.e, one subpopulation will take over even if the two populations have matched fitness, and do not interact competitively.

Different parts of the fungal cell can fix to different genotypes to produce the sectors observed by Pontecorvo and Gemmell in *Penicillium notatum* and more recently in other model microbial systems [HN08]. However this sectoring was not seen in the filamentous fungus *Neurospora crassa*, which is the main model for this thesis [RSH13]. It has previously been argued that the lack of sectoring in *Neurospora crassa* is a result of its relatively fast nuclear dispersal [RSH13] [Lew05].

In this thesis I will analyze the dynamics of two-populations of nuclei within a filamentous fungus. The models developed here are designed to directly represent recent experiments studying *Neurospora crassa* chimera created from two genetically different nuclear populations in which the two types of nuclei can be distinguished because they expose differently fluorescently labeled versions of the the same histone.[MSC15] Since the nuclei are genetically identical except for the differently colored version of the same protein they divide at the

same rate. Additionally although genetically different nuclei have previously been shown to have complex ecological-like dynamics(see [Mah05] for a review of experiments), the nuclei in these synthetic chimera are too similar for interactions between different nuclei to differ from interactions between genetically identical nuclei.

Instead the focus of this thesis will be on how dynamics within the network are affected by the geometry of the fungus network. Filamentous fungal cells form a branched and often loopy network. Previous work has shown that nuclei usually follow complex paths through this network as they move from the sites of production(typically in the middle of the fungus) to sites of growth that are typically distributed along the edges of the fungus. Roper et al. [RSH13] argued that these complex paths were an adaptive feature of the network, creating physical mixing of the genetically diverse nuclei present within the fungus, just as multi-directional flows can be used to mix two different fluids.[SOW06] This idea remains controversial because it is generally assumed that biological transport networks such as fungal cells are optimized to minimize the cost of transporting matter.

This is typically assumed for all biological networks including plant, vasculature, cellular tubes of slime molds, and animal blood networks [HC13]. Yet there are thought to be 1.5 million species of filamentous fungus [Bla11], and among the known species all networks are visibly wired differently. Understanding the physical principles according to which different biological networks are wired will help to explain the morphological diversity of biological networks. At the same time understand how biological networks mix effectively may aid in the design of human transport networks, e.g, for cars or data, which must accomplish similar objectives and have similar tradeoffs in terms of the cost of building the network.

We start from simulating some intuitive models for nuclear population dynamics during fungal growth, then provide asymptotic and analytic insights on the dynamics involved, and the relationship between population dynamics and the branching structure of the network. Comparisons among the simulation, the simplified model, and the physical experiment data will also be discussed.
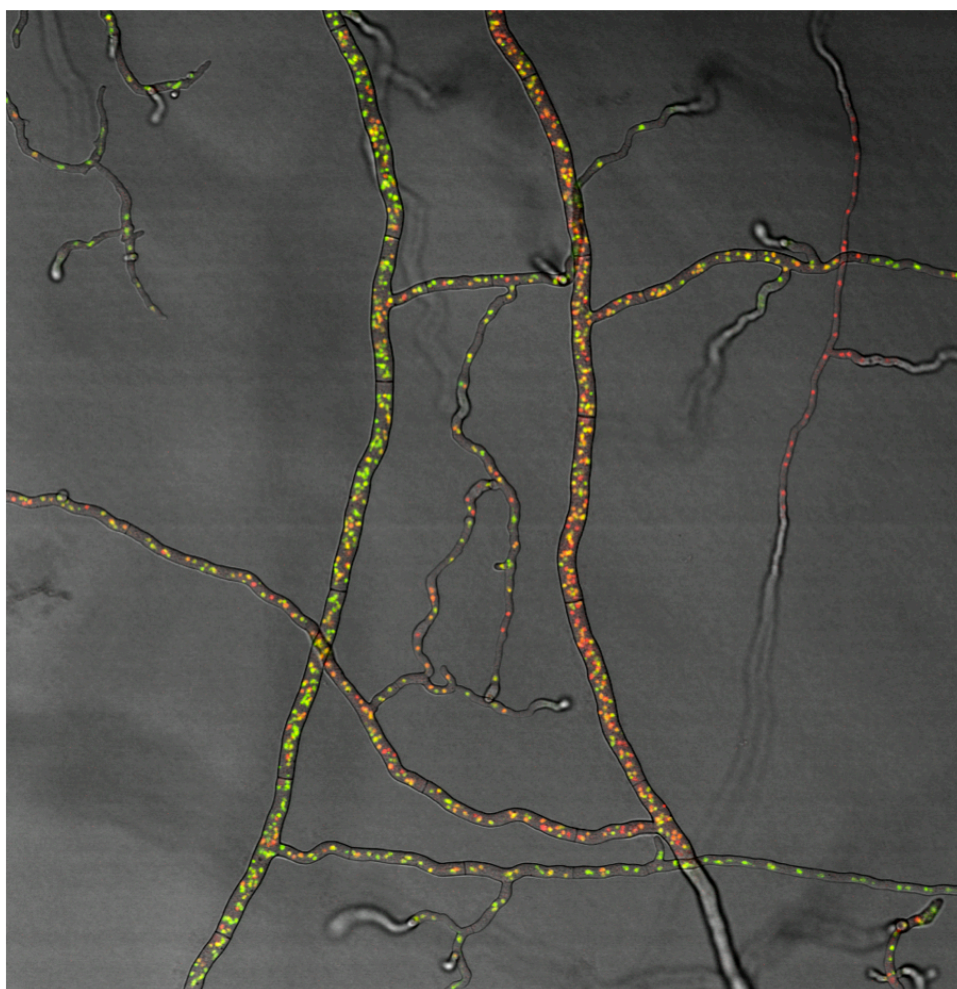
Figure 1.1: Chimeric *Nurospora crassa* mycelium containing hH1::gfp and hH1::DsRed labeled nuclei. [RSH13]

CHAPTER 2

Review of Previous Work

Throughout this thesis we will refer to two canonical models for genetic dynamics within asexual population - Moran model[Mor58] for a fixed-size, well-mixed population, and the stepping stone model[KAH10] that describes a linked chain of such populations in which there is migration between neighboring populations. Both of these models are renewed in [CK70]. In particular these models highlight the role of genetic drift - that is fluctuations in population make up due to randomness of division - in controlling the dynamics of these populations over time.

2.1   Moran Process: Genetic Drift in Finite Populations

Consider a well mixed population of fixed size n composed of two nuclear species. Denote k as the size of the first species. The dynamics of the Moran process is described as follows: During each step, one nucleus chosen randomly from the population divides, then to maintain this fixed size of the population, another nucleus chosen randomly from the original population is removed from the system. Previously, this model was used by [RSH13] as a model for the population of nuclei behind the growing apex of a fungal hypha.

If we denote $P(j|k)$ as the conditional probability(with respect to the current state) of the size of the first species changes from k to j in one time step, then we can write the

transition as follows:

$$P(k-1|k) = (1 - k/n)(k/n) = H(k,n)/2$$

$$P(k+1|k) = (k/n)(1 - k/n) = H(k,n)/2 \qquad (2.1)$$

$$P(k|k) = 1 - 2(k/n)(1 - k/n) = 1 - H(k,n)$$

where the heterozygosity is defined as $H(k,n) = 2(k/n)(1 - k/n)$. This quantity is the probability that two randomly selected individuals are from different species. (In ecology this quantity is called the Gini-Simpson diversity index. [Jos06] [Hur71] [GM62])

It can be shown that:

$$H(k-1,n) = H(k,n) + 2k/n^2 - (2/n)(1 - (k-1)/n) \qquad (2.2)$$

$$H(k+1,n) = H(k,n) - 2k/n^2 + (2/n)(1 - (k+1)/n) \qquad (2.3)$$

Now we can define a Markovian sequence of random variables $\{k_i,\ i = 0, 1, \cdots\}$, representing the sequence of the sizes of the first population during each division step. From 2.1, we see that k_i performs a symmetric random walk on the set $\{0, 1, \cdots, n\}$. Accordingly

$$E(k_{i+1}|k_i) = (k_i - 1)P(k_i - 1|k_i) + (k_i + 1)P(k_i + 1|k_i) + k_i P(k_i|k_i) = k_i \qquad (2.4)$$

and hence

$$E(k_{i+1}) = E(k_i) = \cdots = E(k_0) \qquad (2.5)$$

However, the boundary conditions at $k = 0$ and $k = n$ are absorbing, meaning if $k_i = 0$ or $k_i = n$ at some time point, then $k_j = k_i, \forall j > i$. We say then that one of the species has taken over, or the population has fixed to one of species.

From 2.5 we see the dynamics of fixation are not reflected in the first moment of k_i, but they do show up when we compute the evolution of the heterozygosity. Specifically, if we then denote $H(k_i, n) = H_i$, we have a conditional expectation for H_{i+1} given H_i, that is:

$$
\begin{aligned}
E(H_{i+1}|H_i) =& H(k_i - 1, n)P(k_i - 1|k_i) + H(k_i, n)P(k_i|k_i) + H(k_i + 1)P(k_i + 1|k_i) \\
=&(H_i + 2k_i/n^2 - (2/n)(1 - (k_i - 1)/n))(H_i/2) \\
&+ H_i(1 - H_i) + (H_i - 2k_i/n^2 + (2/n)(1 - (k+1)/n))(H_i/2) \\
=&(1 - 2/n^2)H_i
\end{aligned}
\qquad (2.6)
$$

Using law of total expectation, we then have

$$E(H_{i+1}) = E_{H_i}(E(H_{i+i}|H_i)) = (1 - 2/n^2)E(H_i) \tag{2.7}$$

and hence

$$E(H_{i+1}) = (1 - 2/n^2)^i H_0 \tag{2.8}$$

i.e, $\lim_{i \to \infty} E(H_i) = 0$, so a finite population loses diversity with time.

Assuming that $\tau(= dt)$ be the mean time between one nuclear division in the population and the next, and hence $\tau_g = n\tau = ndt$ be the generation time(taken for each nucleus in the generation to divide), a continuous version of this dynamics is:

$$d(f_t) = \sqrt{(2/n^2)f_t(1 - f_t)}dW = \sqrt{(2/(n\tau_g))f_t(1 - f_t)}dW \tag{2.9}$$

where $f(t)$ is the proportion of first species at time t. For E(f(t)), the expectation of $f(t)$ we have

$$\frac{dE(f(t))}{dt} = 0 \tag{2.10}$$

For $H(t) = E(f(t)(1 - f(t)))$, the expectation of heterozygosity we have

$$d(H(t)) = (-2/n^2)H(t) = (-2/(n\tau_g))H(t)dt \tag{2.11}$$

which also leads to an exponential decay: $H(t) = H(0)\exp(-2t/(n\tau_g))$.

2.2   The Stepping Stone Model

The Moran process model assumes a discrete well-mixed population; although nuclei do not directly interact, any nucleus may divide and any other nucleus may be ejected from the population. To understand sectoring and founder effects, some form of spatial structure must be incorporated. A standard model for space and time dynamics of a population is the stepping stone model.

Here we imagine that the population is partitioned into semi-isolated "islands" or demes. Within each deme, the population performs a version of the Moran process. But each deme

is connected to a subset of the other demes, with which it may exchange a certain number of nuclei at each time step. Because of this nuclear exchange the dynamics of adjacent demes are coupled - that is they can not fix to one species independently at each other.

We will focus on the one dimensional version presented in [KAH10] [KN11]. Korolev et al. used the stepping stone model to study the dynamics of sectoring; they solved the model on the cylinder $S_1 \times [0, \infty)$; treating each new time point as a new ring of demes being added at the growing edge of a diverse colony of cells.

To make things easier, we will not consider mutation and selection. That is, red nuclei divide at the same rate as green nuclei, and each division produces two identical copies of the parent nucleus.

After these assumptions, the dynamics of the stepping stone model is as follows: There are an infinite set of demes arranged on a line. Each deme has N fixed size populations and denote the proportion of species one in deme l by $f_l(t)$.

At each step, a deme exchanges $\tilde{m}N/2$ individuals with its right neighbor and $\tilde{m}N/2$ individuals with its left neighbor. Assume $\tilde{m} \ll 1$. Also assume the genetic drift within the l-th deme itself outweighs the drift caused by migration. Effectively this assumption is equivalent to assuming that if a fraction f_l of the nuclei in deme l are of species one then during the neighbor exchange step exactly $\frac{\tilde{m}N}{2} f_l$ species one nuclei are transferred to each neighbor. The neglect of fluctuations here a justified if $\tilde{m}N \gg 1$ so that the binomial variance of the selection step is negligible.

Therefore, the diffusion due to the migration between the neighbors are neglected. Then in one time step:

$$df_l = \frac{m}{2}(f_{l-1} + f_{l+1} - 2f_l) + \sqrt{D_g f(1-f)}dW \qquad (2.12)$$

where $m = \tilde{m}\tau_g^{-1}$, τ_g is the generation time. Taking the continuous space limit with $x = la$

$$df = D_s \frac{\partial^2 f}{\partial x^2} + \sqrt{D_g f(1-f)}dW \qquad (2.13)$$

where $D_s = ma^2/2$ and $D_g = a = 2a/(\tau_g N)$ Let $F(t, x) = E(f(t, x))$ and $H(t, x_1, x_2) = E(f(t, x_1)(1 - f(t, x_2)) + f(t, x_2)(1 - f(t, x_1)))$. H represents a pairwise heterozygosity, i.e,

the likelihood that two individuals randomly chosen from sites x_1 and x_2 are genetically different. Using Ito's Lemma, one can show

$$\frac{\partial H(t, x_1, x_2)}{\partial t} = D_s \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_1^2} \right) H(t, x_1, x_2) - D_g H(t, x_1, x_2) \delta(x_1 - x_2) \qquad (2.14)$$

By further assuming well-mixed, spatially homogeneous initial conditions, so the solution is translationally invariant, and define $x = x_1 - x_2$, we have

$$\frac{\partial H(t, x)}{\partial t} = 2 D_s \frac{\partial^2}{\partial x^2} H(t, x) - D_g H(t, 0) \delta(x) \qquad (2.15)$$

Equation 2.15 can be solved using Laplace transform. Specifically at $x = 0$, one have

$$H(t, 0) = H_0 \text{erfc} \left(\frac{D_g^2 t}{8 D_s} \right) e^{\frac{D_g^2 t}{8 D_s}} \qquad (2.16)$$

where $\text{erfc}(\cdot)$ is the complementary error function. And when t is sufficiently large, H has asymptotic behavior:

$$H(t, 0) = H_0 \left(\frac{\pi D_g^2 t}{8 D_s} \right)^{-1/2} + O(t^{-3/2}) \qquad (2.17)$$

2.3 Experimental Observations of Population Dynamics in Living Fungal Cells

Ma et al. 2015 [MSC15] studied experimentally the diversity dynamics of the nuclear populations present with single Neurospora crass cells, in which nuclear "species" could be distinguished by their different fluorescent labels. They analyzed diversity experimentally using two metrics:

- Sample diversity is given by the heterozygosity of any pair of nuclei isolated from within a 5mm diameter region of the fungal network, independently of whether or not they come from the same hyphal compartment.

- Spore diversity is given by the heterozygosity of individual spores, containing exactly two nuclei. It measures the diversity of individual hyphal compartments.

10

Here heterozygosity, which will be referred at many points in this thesis, is again defined to be the likelihood that two randomly chosen nuclei have different genotypes.

Ma et al. found that spore heterozygosity was less than sample heterozygosity, at all points within a growing fungal cell - that is, although both genotypes are typically present in identical proportions in a small area of the fungal network, they are not necessarily found in the same hyphal compartments. However, the data showed two surprising and unexplained features:

- Spore heterozygosities did not show detectable decrease in diversity with distance grown. Thus although hyphal compartments are genetically less rich than the entire mycelium, their diversity is stable with cell age (at least up to the time point where spores are formed).

- The spore diversity was sensitive to the geometry of cell growth - if cells were grown in race-tubes, and therefore grew only in one direction, then the spore diversity was indistinguishable from the sample diversity - that is, genotypes were mutually mixed even at the scale of individual hyphal compartments. By contrast, when cells were grown in petri dishes, and expanded in all directions as they grew, then spore heterozygosity was always significantly smaller than sample heterozygosity, suggesting that individual compartments had different genetic composition.

Ma et al's experiments were performed in wild type cells in *Neurospora crassa*, the cells form a densely connected network, in which nuclei can move on complex, multi-directional. However, Yang et al.(unpublished data, 2015, see Figure 2.1) has since shown that these same dynamics(in particular a reduced by stable spore heterozygosity) is also seen in *soft* cells. These mutant cells are genetically altered so that they can not form interconnected networks, they may branch but they can not fuse with each other. *soft* networks therefore have a tree like topology, with a single root(the spore that produced the network) feeding many different leaves(growing hyphal tips). It follows that branching alone suffices to produce the complex population dynamics seen in real fungal cells, and thus motivates us to consider networks

that have only branching in the rest of this thesis.
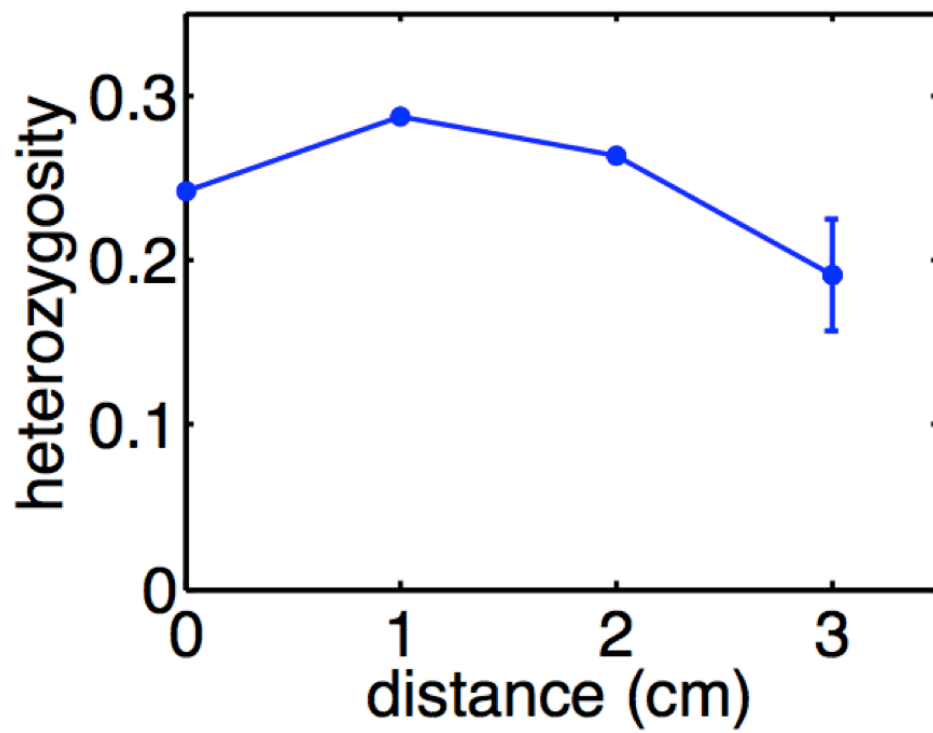
Figure 2.1: Heterozygosity in soft colony measured from experiments.

CHAPTER 3

Numerical Simulations

3.1 The Pinball Model of Nuclear Dynamics within a Filamentous Fungus

We introduce a model for the dynamics of division and migration of nuclei within a simply connected tree-like fungal network.

Consider a fungal network that contains two types of nuclei (with genotypes *hH1::gfp* and *hH1::DsRed*; in this thesis we will refer to them as red nuclei and green nuclei). We build up this network as a linked tree of demes(nuclear sub-populations) that branch outward from a single root deme to multiple leaf demes. Each deme represents a hyphal sub-compartment and the leaf demes are hyphal tips. The fungal growth occurs at the tips, so the capacity of tip demes increases with time. The space thereby created is kept filled with nuclei which are produced by divisions occurring through the network.

Besides the specific rules of our proposed simulations and models, we will make the following general assumptions:

- Hyphae grow at a constant rate(supported by experimental observation [MRB06]);

- When the number of nuclei of each leaf deme reaches to a certain amount N, a new leaf deme is generated below it in the network;

- The underlying graph of the network will have a tree structure (that is, there is a unique path between any pair of demes in the fungus). In general the networks built by *Neurospora crassa* are multi-connected because of the ability of hyphae to fuse with

each other. But unpublished experiment work shows that soft, a mutant strain of the fungus that is unable to undergo hyphal fusion[FSG08] still maintains genetic diversity at levels that are comparable to the wild type strain of the fungus.

- Within each deme(or hyphal compartment) the nuclei are well mixed. That is when a nucleus exits the compartment it can be chosen at random from the N nuclei within the deme. In real hypha nuclei are mixed up within each compartment by a combination of velocity gradients in the flow and by their collisions with other organelles, e.g, vesicles, and random motor - driven motions that often oppose the direction of bulk cyto-plasmic flow.[RLH15]

We call this model the pinball model because each division sets off a cascade of nuclear transfers at lower and lower nodes until a left node is reached. See figure 3.1 for an illustration.

For notation, let n_k be the number of red nuclei in deme k, $f_k = n_k/N$ be the proportion of the red nuclei, and the heterozygosity, which turns out to be a more useful statistics for diversity, be $H_k = 2f_k(1 - f_k)$. For most of the numerical experiments in this thesis, $N = 30$.

3.1.1 Direct Tree Simulation

For sufficiently small networks we can directly simulate both network growth and nuclear division on the network, and calculate the distributions for the random variables f_k, etc, by combining ensembles of these networks. The rules for these direct simulations are as follows:

- Binary tree growth: when one deme branches, it always has two descendent demes. Note: because of this structure, our demes can be ordered into generations, starting with the root(generation 0), the deme adjacent to it(generation 1), and so on.

- Each step occurs with exactly one nucleus dividing. Each nucleus is equally likely to divide.
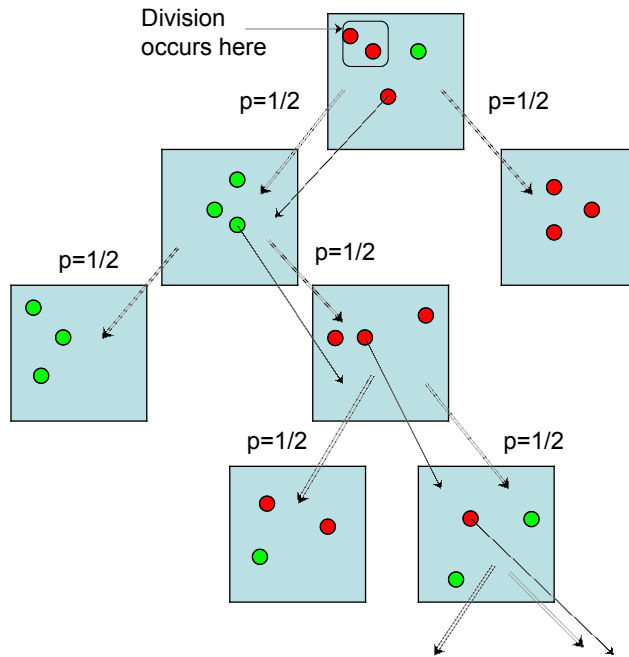
Figure 3.1: Diagram sketch of the pinball model. One particular realization is shown when there are four generations and $N = 3$. In this realization, the new division occurs at the root. And then it sets off a cascade of nuclear transfers throughout the network.

- There is also another nucleus get transferred out of the deme where the division occurs. This nucleus is equally likely to be any nucleus from choice of the original population (i.e, is chosen independently from the population of that deme).

- When a nucleus transfers through a branching point, it will go to either deme in the next generation with equal probability. The deme that is being "fed" then transfers out another nucleus from its original population. And the process goes on.

The simulation is illustrated in 3.2 and 3.3.

3.1.2   Single Lineage in a Full Branching Network

Note that in the previous simulation, because the number of demes grow exponentially with the number of generations, one can only directly simulate a very limited number of
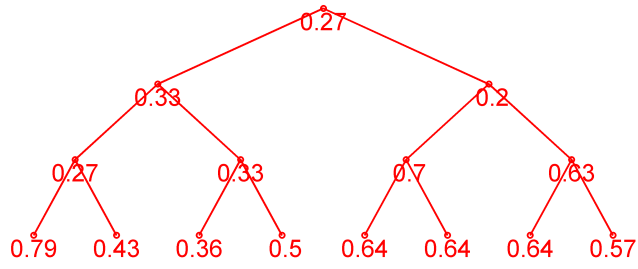
Figure 3.2: Diagram sketch of the direct simulation. One particular realization is shown when there are four generations. Each node represents a deme in the network. The number on each node represents f_k, the proportion of the red nuclei, of that deme at a particular time.
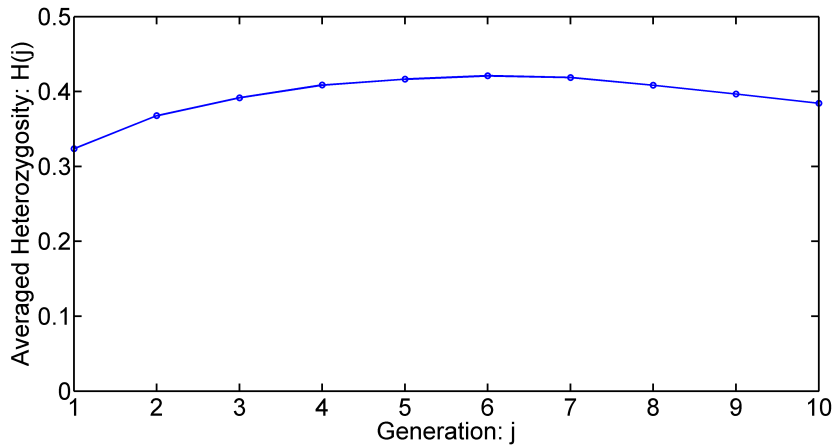
Figure 3.3: The direct simulation. Heterozygosity is averaged over each generations and over 100 independent realizations.

generations. Instead, one can simulate one chain that links the root of the tree to a growing tip. Although nuclei can leave this chain of demes, demes not on this path cannot affect the composition of any deme on the path. Hence the rule is: the probability of transferring from deme i to deme $j (i \leq j \leq m + 1)$ is $1/2^{j-i}$, where i is the deme that divides a nucleus.
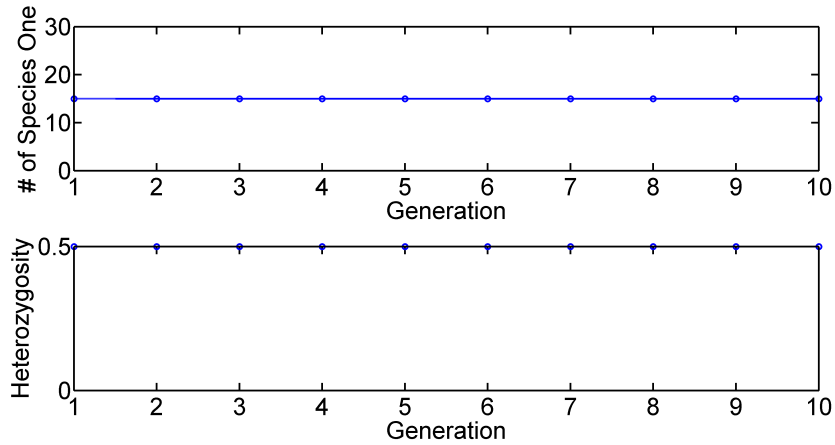
Figure 3.4: Initial condition for all the numerical simulations of single lineages in this chapter. We start with $m_0 = 10$ generations(or demes). $N = 30$. The initial number of species in each deme is $n_j^{(0)} = 15, \forall j$. Top figure: Number of species 1 in each generation. Bottom figure: The heterozygosity for each generation, corresponding to this realization.
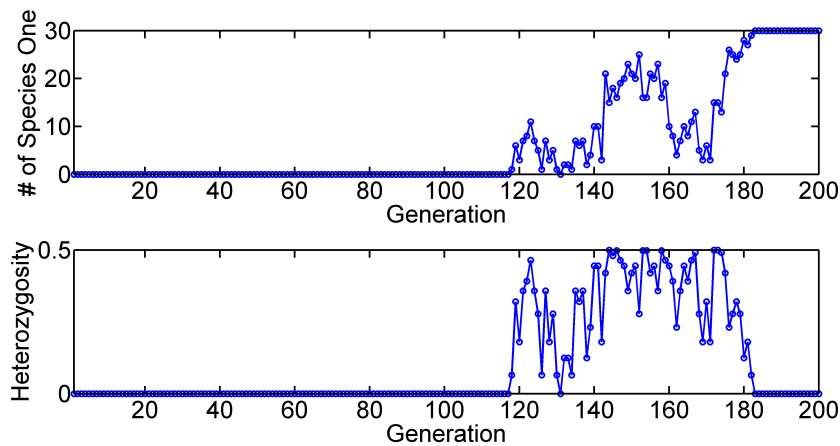
Figure 3.5: One realization of a single lineage with full branching. Here $H_0 = 0.5$, $N = 30$. We start with 10 demes. Top figure: Number of species 1 in each generation. Bottom figure: The heterozygosity for each generation, corresponding to this realization.

3.1.3   Single Lineage in a Full Branching Network with Local Transfers

One can further simplify the growth model by allowing only local nuclear transfers. Under this model the nuclei can be transferred no further than one deme within the chain. That

17
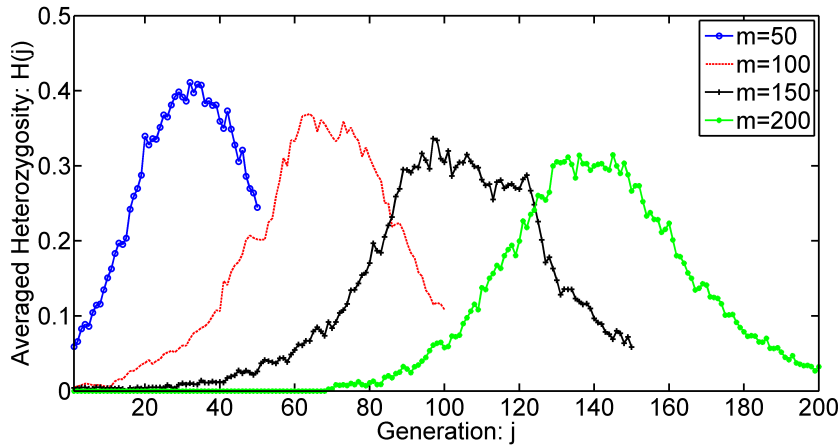
Figure 3.6: Simulation of a single lineage with full branching. In each case, heterozygosity is averaged over 100 independent realizations.

is, in a symmetric branching tree with probability

The rule is: when one nucleus is divided, the transfer is ended in the next one generation.
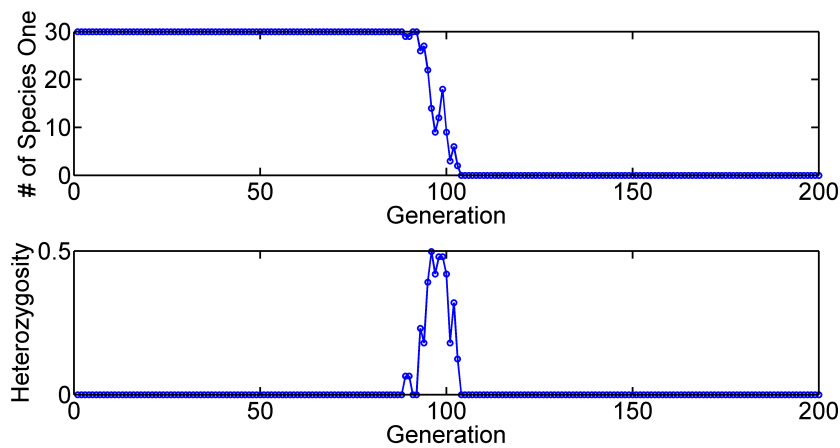
Figure 3.7: One realization of a single lineage with full branching and local transfers. Here $H_0 = 0.5$, $N = 30$. We start with 10 demes. Top figure: Number of species 1 in each generation. Bottom figure: The heterozygosity for each generation, corresponding to this realization.

The result is shown in 3.8.
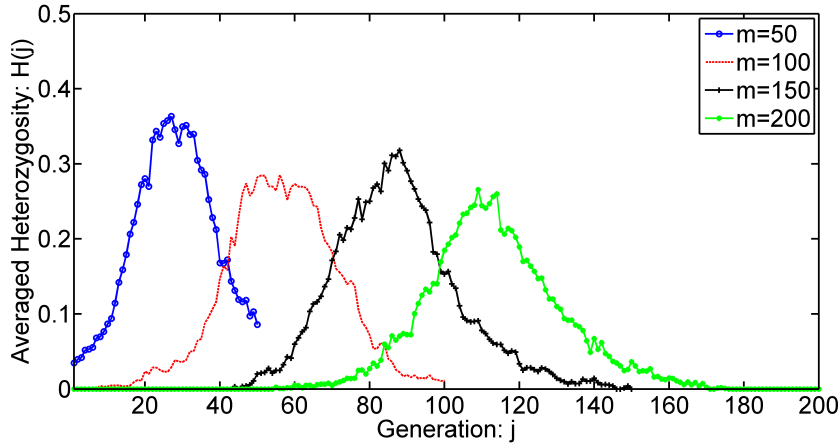
Figure 3.8: Simulation of a single lineage with full branching and local transfers. In each case, heterozygosity is averaged over 100 independent realizations.

3.1.4 No Branching Case

As the opposite case of full branching, we also consider the case where the network has no branching. In this case the pinball effect always proceeds along all linked demes until a nucleus is moved into the last deme. The result is shown in 3.10.
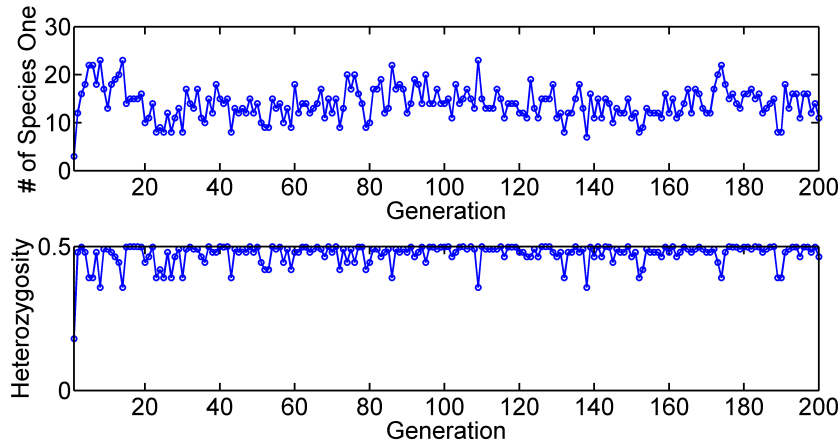
Figure 3.9: One realization of a single lineage with full branching and local transfers. Here $H_0 = 0.5$, $N = 30$. We start with 10 demes. Top figure: Number of species 1 in each generation. Bottom figure: The heterozygosity for each generation, corresponding to this realization.
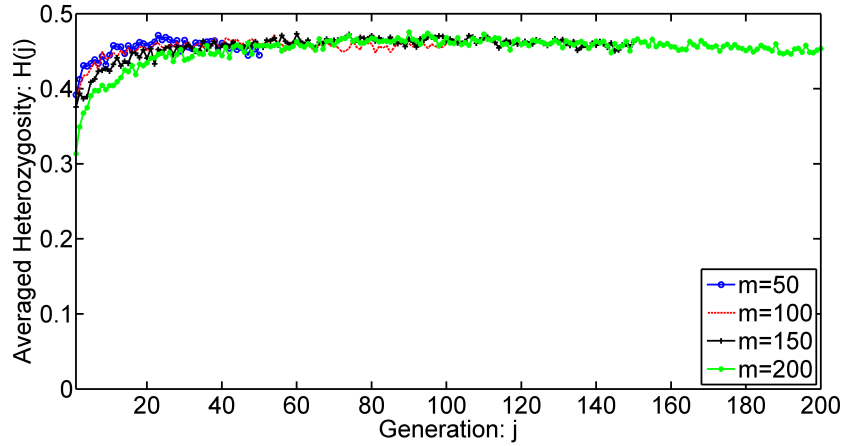
Figure 3.10: Simulation of a single lineage with no branching. In each case, heterozygosity is averaged over 100 independent realizations.

3.2 Observations of the Numerical Results

Here are a few things that one can observe from the numerical simulations:

- For both full branching networks, the expected heterozygosity profile seems to be a traveling "bump".

- For full branching networks, no matter how the specifics on nuclear transferring is, the peak of the heterozygosity in general decays with time. However, this decay is not as significant as the exponential decay in the Moran process.

- The no branching network seems to have a very different dynamics. While the heterozygosity of root deme decays with time, the tips maintain the diversity pretty well. In fact, the heterozygosity of the tip demes stays almost flat as the time increases.

All of these make it necessary and interesting to analyze the spatial distribution of heterozygosity.

20

CHAPTER 4

Dynamics of Heterozygosity

Although fixed size populations are homogenized eventually by genetic drift, a growing size population may maintain its diversity indefinitely.

Moreover in a linked chain of demes the expected diversity may be transiently increased by nuclear dynamics. In this chapter we first discuss two previously unreported models that expose the time scales of these effects.

The traveling bump solution observed in simulations of a full branching network is consistent with these two ideas; since at each point in the linked chain of demes, there is long time decay of heterozygosity, but as the bump travels along the chain, heterozygosity may locally increase. In this chapter we also analyze bump dynamics asymptotically, yielding a theory that can quantitatively explain the height, width and velocity of the bump.

4.1   Bucket Model

We will start with a non-spatial model that has a simple dynamics: During each step, one nucleus chosen randomly from the population divides. The system also starts with a well mixed population of fixed size n with two nuclear species. The key difference compared with a Moran process is that no nucleus is removed at any time, resulting a growing population size.

Variants of these models(i.e, a population having exponential growth) have been previously discussed by [SH91] [Hah66] - in the context of modeling a population that undergoes periodic cycles of exponential growth and then collapse(genetic bottlenecks). However these

works do not provide analytic formulas for the heterozygosity dynamics.

We call this model "the bucket model" since it corresponds to taking all of the nuclei from the fungus and putting them in a single bucket, with no tracking of where nuclei are distributed throughout the fungal network.

Denote $f_k = r_k/n_k$, where r_k is the size of the population of species one nuclei at the k-th step and $n_k = n_0 + k$ is the size of the population at the k-th step. Let the heterozygosity $H_k = 2f_k(1 - f_k)$ same as before. During the k-th to the $k+1$-th step, we have

$$E(f_{k+1}|k\text{-th step}) = \frac{1 + r_k}{1 + n_k}f_k + \frac{r_k}{1 + n_k}(1 - f_k)$$
$$= f_k$$
(4.1)

$$E((f_{k+1})^2|k\text{-th step}) = \left(\frac{1 + r_k}{1 + n_k}\right)^2 f_k + \left(\frac{r_k}{1 + n_k}\right)^2 (1 - f_k)$$
$$= f_k^2 + \frac{1}{2(1 + n_k)^2}H_k$$
(4.2)

Use the total law of expectation, we have

$$E(f_{k+1}) = E(f_k)$$
(4.3)

$$E(H_{k+1}) = \left(1 - \frac{1}{(1 + n_k)^2}\right)E(H_k)$$
(4.4)

Therefore, we have

$$E(f_k) = E(f_0), \forall k$$
(4.5)

and

$$E(H_k) = \left(\frac{k + 1 + n_0}{k + n_0}\frac{n_0}{1 + n_0}\right)E(H_0), \forall k$$
(4.6)

So we have

$$\lim_{k \to \infty} E(H_k) \to \frac{n_0}{1 + n_0}E(H_0)$$
(4.7)

that is, the heterozygosity, instead of decaying to 0, has a non-zero asymptote $\frac{n_0}{1+n_0}E(H_0)$. It follows that the population is not expected to fix to one specie. In other words, by allowing a growing population size, the diversity of the system is preserved.

22

4.2 Building Block: One division step for two demes

According to the bucket model the "global heterozygosity" of a population of nuclei that is not limited in size will not decay to 0 as the population grows. However, a fungal mycelium, though it may grow without bound, is made up of demes whose size is fixed. Moreover the relevant quantity to the diversity of the network and of the spores that it produces, is the diversity of individual demes.

To understand how diversity may change with time at the scale of individual demes, we first consider two linked demes. For the sake of simplicity we assume that the dynamics is as follows: The nuclear division occurs in Deme 1, and the randomly chosen nucleus from the original population of Deme 1 is transferred into Deme 2, and then one nucleus from the original population of Deme 2, chosen randomly, is dropped out. Denote n_1, n_2 as the number of red nuclei in each Deme 1 and Deme 2 and $f_1 = n_1/N$, $f_2 = n_2/N$ as the density of red nucleus for each deme.

Table 4.1 shows how the state variables (n_1, n_2) change during in one step.

State Variables	Description	Probability
$(n_1 + 1, n_2)$	red divides, green transferred, green dropped out	$f_1(1 - f_1)(1 - f_2)$
$(n_1 + 1, n_2 + 1)$		0
$(n_1, n_2 + 1)$	red divides, red transferred, green dropped out	$f_1^2(1 - f_2)$
$(n_1 - 1, n_2 + 1)$	green divides, red transferred, green dropped out	$(1 - f_1)f_1(1 - f_2)$
$(n_1 - 1, n_2)$	green divides, red transferred, red dropped out	$(1 - f_1)f_1 f_2$
$(n_1 - 1, n_2 - 1)$		0
$(n_1, n_2 - 1)$	green divides, green transferred, red dropped out	$(1 - f_1)^2 f_2$
$(n_1 + 1, n_2 - 1)$	red divides, green transferred, red dropped out	$f_1(1 - f_1)f_2$
(n_1, n_2)	red or green divides, transferred, dropped out	$f_1^2 f_2 + (1 - f_1)^2(1 - f_2)$

Table 4.1: Description of the one step dynamics.

We can also marginalize the joint distribution to track only the changes in deme 1 or in

deme 2, as needed. The marginal transition probabilities are then

$$n_1 \to n_1 - 1 : \quad (1 - f_1)f_1$$

$$n_1 \to n_1 : \quad f_1^2 + (1 - f_1)^2$$

$$n_1 \to n_1 + 1 : \quad f_1(1 - f_1)$$

$$n_2 \to n_2 - 1 : \quad (1 - f_1)f_2$$

$$n_2 \to n_2 : \quad f_1 f_2 + (1 - f_1)(1 - f_2)$$

$$n_2 \to n_2 + 1 : \quad f_1(1 - f_2)$$

Now let f_1^+ and f_2^+ be the density after one step, then conditioning on f_1, f_2(state variables at the current step), the expected properties obey:

$$< f_1^+ > = (f_1 - 1/N)(1 - f_1)f_1 + f_1(f_1^2 + (1 - f_1)^2) + (f_1 + 1/N)f_1(1 - f_1)$$

$$= f_1$$

$$< (f_1^+ - < f_1^+ >)^2 > = (1/N)^2(1 - f_1)f_1 + 0 + (1/N)^2 f_1(1 - f_1)$$

$$= (2/N^2)f_1(1 - f_1)$$

$$< f_2^+ > = (f_2 - 1/N)(1 - f_1)f_2 + f_2(f_1 f_2 + (1 - f_1)(1 - f_2)) + (f_2 + 1/N)f_1(1 - f_2)$$

$$= f_2 + (1/N)(f_1 - f_2)$$

$$< (f_2^+ - < f_2^+ >)^2 > = (1/N^2)[(-1 - f_1 + f_2)^2(1 - f_1)f_2 + (-f_1 + f_2)^2(f_1 f_2 + (1 - f_1)(1 - f_2))$$

$$+ (1 - f_1 + f_2)^2 f_1(1 - f_2)]$$

$$= (1/N^2)[(f_2 - f_1)^2 + (1 - f_1)f_2 + f_1(1 - f_2) + 2(f_2 - f_1)(f_1 - f_2)]$$

$$= (1/N^2)[-(f_2 - f_1)^2 + f_1 + f_2 - 2f_1 f_2]$$

$$= (1/N^2)(f_1(1 - f_1) + f_2(1 - f_2))$$

$$(4.8)$$

And also

$$< (f_1^2)^+ > = (f_1 - 1/N)^2(1 - f_1)f_1 + f_1^2(f_1^2 + (1 - f_1)^2) + (f_1 + 1/N)^2 f_1(1 - f_1)$$

$$= (1 - 2/N^2)f_1^2 + (2/N^2)f_1$$

$$< (f_2^2)^+ > = (f_2 - 1/N)^2(1 - f_1)f_2 + f_2^2(f_1 f_2 + (1 - f_1)(1 - f_2)) + (f_2 + 1/N)^2 f_1(1 - f_2)$$

$$= f_2^2 + (1/N^2)(f_1(1 - f_2) + f_2(1 - f_1)) + (2f_2/N)(f_1 - f_2)$$

$$< (f_1 f_2)^+ > = f_1 f_2 + (1/N)f_1(f_1 - f_2) - (1/N^2)f_1(1 - f_1)$$

$$(4.9)$$

where $< \cdot >$ is conditional expectation taken with resect to f_1, f_2. To get unconditional expectation and high moment information, we need to apply the law of total expectation i.e, $E(\cdot) = E_{f_1, f_2}(< \cdot >)$. We then have

$$E(f_1^+) = E(f_1)$$

$$E(f_2^+) = (1/N)E(f_1) + (1 - 1/N)E(f_2)$$

$$E((f_1^2)^+) = (1 - 2/N^2)E(f_1^2) + (2/N^2)E(f_1) \qquad (4.10)$$

$$E((f_2^2)^+) = (1 - 2/N)E(f_2^2) + 2(1/N - 1/N^2)E(f_1 f_2) + (1/N^2)(E(f_1) + E(f_2))$$

$$E((f_1 f_2)^+) = (1 - 1/N)E(f_1 f_2) + (1/N + 1/N^2)E(f_1^2) - (1/N^2)E(f_1)$$

where $^+$ denotes the quantity in the next step. This will serve as a building block for the analysis of our proposed model.

We may track the expectations using a matrix recursion relation of the form:

$$[\ E(f_1^+)\ \ E(f_2^+)\ \ E((f_1^2)^+)\ \ E((f_2^2)^+)\ \ E((f_1 f_2)^+)\ ]^T$$

$$= T[\ E(f_1)\ \ E(f_2)\ \ E(f_1^2)\ \ E(f_2^2)\ \ E(f_1 f_2)\ ]^T \qquad (4.11)$$

where

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/N & 1 - 1/N & 0 & 0 & 0 \\ 2/N^2 & 0 & 1 - 2/N^2 & 0 & 0 \\ 1/N^2 & 1/N^2 & 0 & 1 - 2/N & 2/N - 2/N^2 \\ -1/N^2 & 0 & 1/N + 1/N^2 & 0 & 1 - 1/N \end{bmatrix} \qquad (4.12)$$

T is a constant matrix, so its eigenvalues determine the behavior of the expected moments. The eigenvalues of T are the diagonals, i.e,

$$\lambda_{1,2,3,4,5} = 1, 1 - 2/N^2, 1 - 1/N, 1 - 1/N, 1 - 2/N.$$

Note if we consider other statistics(e.g, heterozygosity $H(2) = E(2(f_2 - f_2^2)$, $H(1,2) = E(f_1 + f_2 - 2f_1f_2)$, up to second moment, we could change to another set of basis. The eigenvalues stay the same under change of basis.

We could see the eigenvalues 1 and $1 - 2/N^2$ corresponds to the dynamics of Deme 1 (i.e, the Moran process; that is $E(f_1)$ being constant and $E(H_1)$ decaying exponentially in time).

Although the other eigenvalues give a decay rate of the two-deme system, it is not so clear how spatial dynamics(quantities like $H(2)$ and $H(1,2)$) evolve. We will proceed with more careful analysis below.

So the diversity of these two demes decreases exponentially with time. Although we introduce the two deme model to get a sense of the time scales and rate of diversity loss, it is quantitatively correct as a model for the nuclear populations in the first two demes of the fungal network. These results are suggestive that any linked chain of demes loses diversity exponentially.

However, according to the bucket model the entire chain of demes does not. Taken together these results suggest that diversity has spatial structure on the linked demes, with exponential loss on any single deme at large times, but also with transient growth.

4.3   Full Branching Local Transfer Case

To explore spatial dynamics we consider a simplified case based on the full branching model and the restriction that nuclei can be transferred no further than to the next deme within the lineage chain. Let $< \cdot >$ be the conditional expectation (conditioned with respect to the current state and that the newly divided nucleus is from deme i) of the quantity of the state

after one step division, we have

$$
< f_j^+ >= \begin{cases} f_j & i \neq j-1 \\ \dfrac{1}{2}f_j + \dfrac{1}{2}\left[f_j + \dfrac{1}{N}(f_{j-1} - f_j)\right] & i = j-1 \end{cases} \tag{4.13}
$$

$$
< (f_j^2)^+ >= \begin{cases} f_j^2 & i \neq j-1, j \\ \dfrac{1}{2}f_j^2 + \dfrac{1}{2}\left[f_j^2 + \dfrac{1}{N^2}(f_{j-1}(1-f_j) + f_j(1-f_{j-1})) + \dfrac{2}{N}f_j(f_{j-1} - f_j)\right] & i = j-1 \\ f_j^2 + \dfrac{2}{N^2}f_j(1-f_j) & i = j \end{cases} \tag{4.14}
$$

$$
< (f_{j-1}f_j)^+ >= \begin{cases} f_j f_{j-1} & i \neq j-1, j-2 \\ \dfrac{1}{2}f_j f_{j-1} + \dfrac{1}{2}\left[f_j f_{j-1} + \dfrac{1}{N}f_{j-1}(f_{j-1} - f_j)) - \dfrac{1}{N^2}f_{j-1}(1-f_{j-1})\right] & i = j-1 \\ f_j\left[\dfrac{1}{2}f_{j-1} + \dfrac{1}{2}(f_{j-1} + \dfrac{1}{N}(f_{j-2} - f_{j-1}))\right] & i = j-2 \end{cases} \tag{4.15}
$$

$$
< (f_{j-d}f_j)^+ >= \begin{cases} f_j f_{j-d} & i \neq j-1, j-d-1 \\ f_{j-d}\left[\dfrac{1}{2}f_j + \dfrac{1}{2}(f_j + \dfrac{1}{N}(f_{j-1} - f_j))\right] & i = j-1 \\ f_j\left[\dfrac{1}{2}f_{j-d} + \dfrac{1}{2}(f_{j-d} + \dfrac{1}{N}(f_{j-d-1} - f_{j-d}))\right] & i = j-d-1 \end{cases} \tag{4.16}
$$

where $d \geq 2$.

Now let $E(\cdot)$ be the unconditional expectation and $h(i, j) = f_i(1 - f_j) + f_j(1 - f_i) = f_i + f_j - 2f_i f_j = \frac{1}{2}(h(i) + h(j)) + (f_i - f_j)^2$, $h(j) = h(j, j) = 2f_j(1 - f_j)$, $H(i, j) = E(h(i, j))$, $H(j) = E(h(j))$, with the law of total expectation (e.g, $E_{\mathbf{f}}(< h(j)^+ >) = E_{\mathbf{f}}(E(h(j)^+|\mathbf{f})) = H(j)^+$, etc) and denote $c = \frac{1}{2(mN+k)}$, we have the following

$$
H(1)^+ = (1 - 4c/N)H(1) \tag{4.17}
$$

which is the Moran process. For $j \geq 2$, we have

$$H(j)^+ = (1 - 2c - 4c/N)H(j) + 2c(1 - 1/N)H(j-1, j) \tag{4.18}$$

Similarly, we have

$$H(1,2)^+ = (1 - c)H(1,2) + c(1 + 1/N)H(1) \tag{4.19}$$

and for $j \geq 3$

$$H(j-1, j)^+ = cH(j-2, j) + (1 - 2c)H(j-1, j) + c(1 + 1/N)H(j-1) \tag{4.20}$$

For $d \geq 2$, we have

$$H(j - d, j)^+ = cH(j - d - 1, j) + (1 - 2c)H(j - d, j) + cH(j - d, j - 1) \tag{4.21}$$

and for $j \geq 3$

$$H(1, j)^+ = (1 - c)H(1, j) + cH(1, j - 1) \tag{4.22}$$

4.4   Approximations of the Dynamics

4.4.1   Closure Approximations of $f_{j-2}f_j$

The big challenge of the above model, even with local transfer, is that although we are typically only interested in $H(j)$, i.e, the diversity at a single point, to evolve this quantity we need also to know $E(f_{j-1}f_j)$, the correlation information between neighboring demes. But to evolve $E(f_{j-1}f_j)$, we also need to have $E(f_{j-2}f_j)$ and so on. Therefore, our evolution of $H(j)$ requires keeping track of all m^2 cross-correlations.

Instead of tracking all those terms, one might be tempted to find a reasonable approximation of the terms $f_{j-2}f_j$ by using only a combination of f_j, f_j^2 and $f_{j-1}f_j$. The aim is to: (i) have fewer variables(i.e, reducing from $O(m^2)$ to $O(m)$) to track; (ii) focus on the local behavior while still preserving the dynamics to some accuracy. Below are some ad-hoc ideas to do this.

- Use the approximation: $h(j-2,j) \approx h(j-2,j-1) + h(j-1,j)$

  The idea here is that nuclear populations form homogeneous islands within the cell [HK09] [KAH10]. $h(j-1,j)$ therefore represents the probability that an interface between two island occurs between $j-1$ and j. Since islands seem to span multiple demes, so $h(j-2,j) = 0$ if $j-2$, j lie in the same island; and $h(j-2,j) = 1$ only if there is a boundary between $j-2$ and $j-1$ or between $j-1$ and j, and if we treat these as exclusive events, we could end up using the sums of $(j-2,j-1)$ and $(j-1,j)$ correlations as a proxy to $(j-2,j)$ correlation, i.e,

$$f_{j-2}f_j \approx -f_{j-1} + f_{j-1}f_j + f_{j-2}f_{j-1}$$

- Use the approximation: $f_j(f_{j-2} - f_{j-1}) \approx f_j(f_{j-1} - f_j)$

$$f_{j-2}f_j \approx 2f_{j-1}f_j - f_j^2$$

- $f_{j-2}f_j \approx f_{j-1}f_j$

- $f_{j-2}f_j \approx \frac{1}{2}(f_{j-1}f_j + f_{j-2}f_{j-1})$

All of these approximations except the first are analogous to making a continuum approximation in which all pairwise correlation are obtained by Taylor series expansions involving the correlation between immediately neighboring demes. The plurality of different approximations reflects different ways of upwinding these series expansions. The first approximation on the other hand assumes that in realizations of the dynamics $f_j \in \{0,1\}$. (Naturally this requires that all correlations vary only on a length scale much larger than the deme spacing.) However, even these conditions are met, it turns out that the approximations introduce instability artifacts. In particular each of the correlations must be derived from a real probability, which imposes the following inequality constraints:

$$0 \leq H(j) \leq 0.5 \tag{4.23}$$

$$0 \leq H(j-d,j) \leq 1 \tag{4.24}$$

Equivalently, the following constraints on $E(f_j)^2$ and $E(f_{j-d}f_j)$ must be satisfied:

$$0 \leq \max\{E(f_j)^2, E(f_j) - 0.25\} \leq E(f_j)^2 \leq E(f_j) \leq 1 \tag{4.25}$$

$$0 \leq \max\{0, \frac{1}{2}(E(f_{j-d}) + E(f_j)) - 1\} \leq E(f_{j-d}f_j) \leq \frac{1}{2}(E(f_{j-d}) + E(f_j)) \leq 1, \forall d \geq 1 \tag{4.26}$$

Figure 4.1 4.2 show those approximations fail to meet the constraints after a few iterations, when presented with initial data that violates long wavelength variation condition.
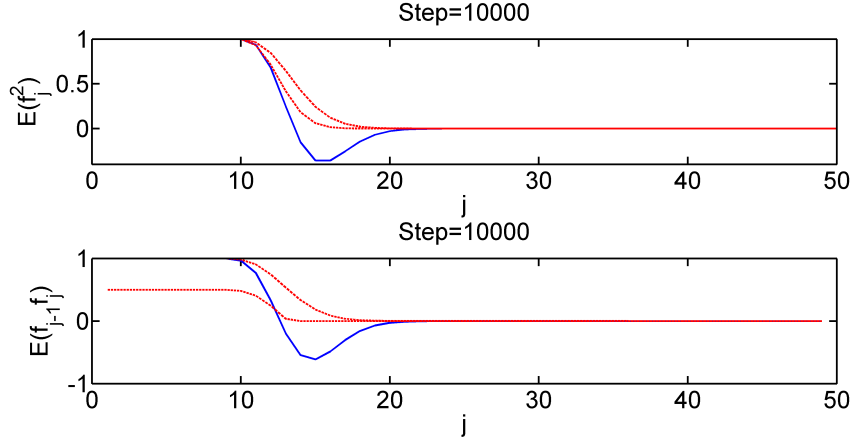
Figure 4.1: $h(j-2,j) \approx h(j-2,j-1) + h(j-1,j)$ violates the constraint. Blue lines: $E(f_j^2)$ and $E(f_{j-1}f_j)$ using this approximation in the update. Red dashed lines: lower and upper bounds of $E(f_j^2)$ and $E(f_{j-1}f_j)$ by equations 4.25 and 4.26. Initial Data: $f_j^{(0)} = 1_{\{j \leq 10\}}, 1 \leq j \leq 50$.

4.4.2 Approximations in the Heterozygosity Update

Another ad-hoc approximation that we shall explore in some detail is to ignore most of the $O(1/N)$ terms in our heterozygosity update. This idea would lead to the following update:

$$H(1)^+ = (1 - 4c/N)H(1)$$

$$H(j-d,j)^+ = cH(j-d-1,j) + (1 - 2c - \delta(d)(4c/N))H(j-d,j) + cH(j-d,j-1),$$

$$0 \leq d \leq j - 2$$

$$H(1,j)^+ = (1-c)H(1,j) + cH(1,j-1), \quad j \geq 2$$

$$\tag{4.27}$$
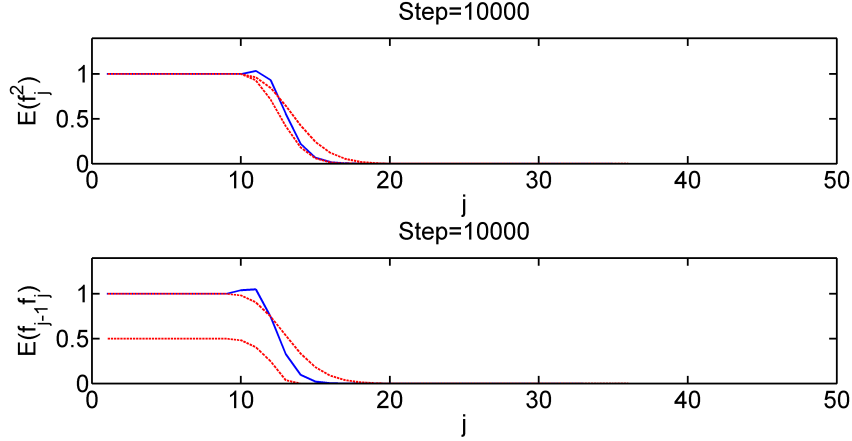
Figure 4.2: $f_j(f_{j-2} - f_{j-1}) \approx f_j(f_{j-1} - f_j)$ violates the constraint. Blue lines: $E(f_j^2)$ and $E(f_{j-1}f_j)$ using this approximation in the update. Red dashed lines: lower and upper bounds of $E(f_j^2)$ and $E(f_{j-1}f_j)$ by equations 4.25 and 4.26. Initial Data: $f_j^{(0)} = 1_{\{j \leq 10\}}, 1 \leq j \leq 50$.

By only keeping the leading order terms(except for $\delta(d)(4c/N)$), one essentially makes the assumption that genetic drift occurs only within the relevant deme itself, ignoring the drift due to advection from neighbors. This is the same assumption in the derivation of the stepping stone model, assuming that nucleus transferred between demes is not chosen at random, but that deme k contributes a deterministic f_k red nuclei to its downstream neighbor each time a transfer occurs.

4.4.3 PDE Perspective

The discrete update equations can be solved numerically, but to gain analytic insight into the dynamics of the network we start by seeking a PDE (continuous space and time) approximation to the update equation.

We would like to consider a continuous version of the update. Assume that for each nucleus, the time that it divides a new nucleus follows $Exp(\lambda)$. Note then the time it takes for $mN + k$ nuclei to divide a new one follows $Exp((mN + k)\lambda)$. And then if we re-scale one increment in time as the time to have a new nucleus division, we would take $\Delta t \approx \frac{1}{mN+k} = 2c$.

Let Δx be the space scale between two demes. With

$$H(j-d-1,j) \approx H(j-d,j) - \frac{\partial H}{\partial x_1}\Delta x + \frac{1}{2}\frac{\partial^2 H}{\partial x_1^2}(\Delta x)^2 \qquad (4.28)$$

$$H(j-d-1,j-1) \approx H(j-d,j) - \frac{\partial H}{\partial x_2}\Delta x + \frac{1}{2}\frac{\partial^2 H}{\partial x_2^2}(\Delta x)^2 \qquad (4.29)$$

And with $\Delta x = 1$, the "leading order" update becomes

$$\frac{\partial H(x_1, x_2, t)}{\partial t} = -\frac{1}{2}\left(\frac{\partial H}{\partial x_1} + \frac{\partial H}{\partial x_2}\right) + \frac{1}{4}\left(\frac{\partial^2 H}{\partial x_1^2} + \frac{\partial^2 H}{\partial x_2^2}\right) - (2/N)\delta(x_2 - x_1)H(x_1, x_2, t) \quad (4.30)$$

with proper initial and boundary conditions, to be discussed below.
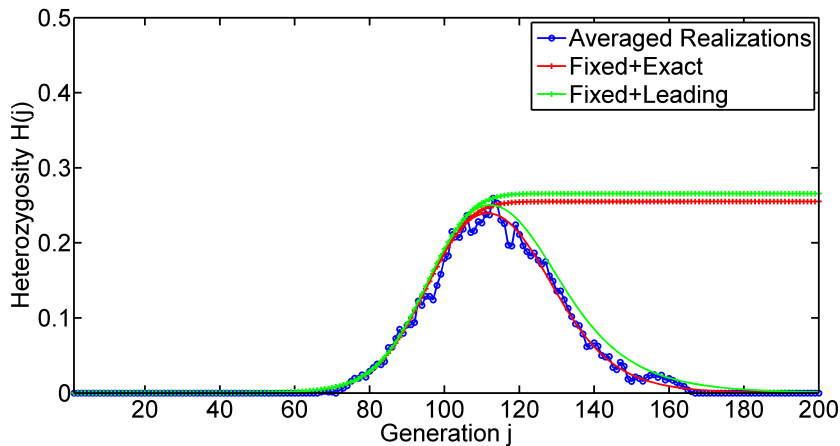
Figure 4.3: Comparison of heterozygosity: Average of realization, fixed-size exact update of H, fixed-size "leading order" update of H. Here the system grows from 10 demes to 200 demes, $T \approx 210$.

The PDE 4.30 is highly similar to the continuous stepping stone model [KAH10]; except that it includes an additional advection term, and that the boundary conditions(to be discussed below) are different. These differences will be shown to strongly alter the character of the PDE, in particular the hyperbolic terms mean that the dynamics must be treated using characteristics, that will be shown to travel more slowly than the growing edge of the fungus. The result is that the localized Moran dynamics in the first and last demes dominate the population genetic dynamics within the chain, as will be discussed below.

32
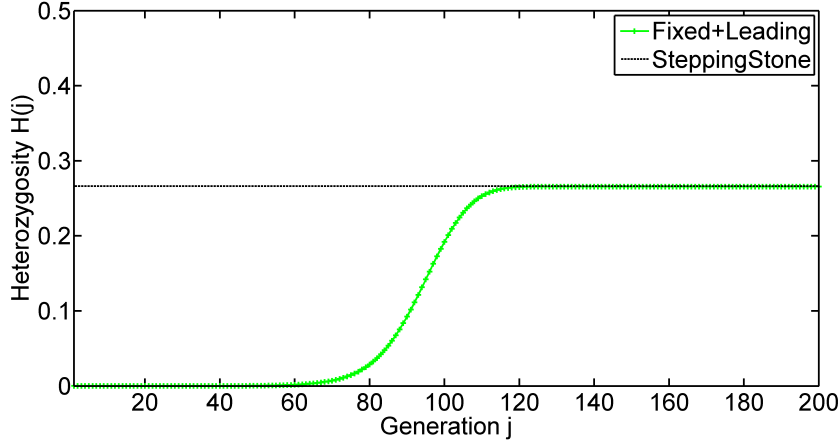
Figure 4.4: Comparison of heterozygosity: Fixed size "leading order" update of H, Stepping-Stone solution. Here the system grows from 10 demes to 200 demes, $T \approx 210$.

To empirically test what approximations are allowed for the dynamics, we consider simplifying the dynamics by removing the moving boundary. To do this we consider an infinite chain of demes. Since our update rule only requires that $f_{j-d}f_j, (1 \leq d \leq j-1)$ be known to compute $f_j \ f_j^2$, etc, we can truncate this chain at a finite point, i.e, consider only the first m demes of the network. When we compare the discrete time and space dynamics on this infinite network they quantitatively agree with the observed bump on its left slope. The absence of characteristics being generated from the moving boundary means that the right slope of the bump is not reproduced by these infinite domain calculations, which asymptote to some constant value as $j \rightarrow \infty$. But with this asymptotic value agrees very closely with the height of the bump.

Additionally the shape of the bump is in quantitative accord with the approximation that considers only leading order terms, i.e, with the PDE solution of 4.30 in the large x limit, since this solution becomes independent of x, we can neglect the advection terms in 4.30.

The height of the bump in the leading order update is the solution of the stepping stone model. This is expected because for the update on leading order terms, the only difference is that stepping stone model assumes homogeneous spatial data and has infinite boundary.

In fact, there are several ways to interpret the left half of the "bump".

If in equation 4.30 we assume H depends only on one spatial variable x with $x = x_2 - x_1$ (note this assumes well-mixed, spatially homogeneous data, same as in [KAH10]), with

$$\frac{\partial H}{\partial x} = -\frac{\partial H}{\partial x_1} = \frac{\partial H}{\partial x_2} \qquad (4.31)$$

$$\frac{\partial^2 H}{\partial x^2} = \frac{\partial^2 H}{\partial x_1^2} = \frac{\partial^2 H}{\partial x_2^2} \qquad (4.32)$$

Denote $\tilde{H}$ as the heterozygosity under this assumption. Then $\tilde{H}(x, t)$ would satisfy

$$\frac{\partial \tilde{H}(x,t)}{\partial t} = \frac{1}{2}\frac{\partial^2 \tilde{H}}{\partial x^2} - (2/N)\delta(x)\tilde{H}(x,t) \qquad (4.33)$$

which is exactly the 1-D stepping stone equation with constants $D_s = 1/4$ and $D_g = 2/N$. Solution of this equation is reported in Chapter 2 and is drawn in Figure 4.4.

Another view is also based on trying to find an approximation of $H(t, x = x_1 = x_2)$. We start with a system with only advection terms. Also we neglect exponential decay terms except for the first deme. That is

$$\frac{\partial H}{\partial t} = -\frac{1}{2}\frac{\partial H}{\partial x} \qquad (4.34)$$

with boundary conditions

$$H|_{x=0} = H_0 e^{-\frac{2}{N}t} \qquad (4.35)$$

$$H|_{t=0} = H_0 \qquad (4.36)$$

Using the method of characteristics(see Figure 4.6), The solution to this is

$$H(t,x) = \begin{cases} H_0 e^{-\frac{2}{N}(t-2x)}, & x < t/2 \\ H_0, & x \geq t/2 \end{cases} \qquad (4.37)$$

Note this solution captures the wave speed $1/2$, but does not reflect the decaying scale. In fact, the right thing to do is to take the stepping stone solution into consideration. Specifically, let

$$A(t) = \text{erfc}\left(\frac{D_g^2 t}{8D_s}\right) e^{\frac{D_g^2 t}{8D_s}} \qquad (4.38)$$

34

with $D_s = 1/4$ and $D_g = 2/N$. And the right thing to do is to apply this decaying scale along the characteristics lines, instead of remaining constant. We will then have

$$H(t,x) = \begin{cases} H_0 e^{-\frac{2}{N}(t-2x)} A(2x), & x < t/2 \\ H_0 A(t), & x \geq t/2 \end{cases} \tag{4.39}$$

Up to the diffusion terms, this is a very good approximation of the left half of the bump. See Figure 4.5.
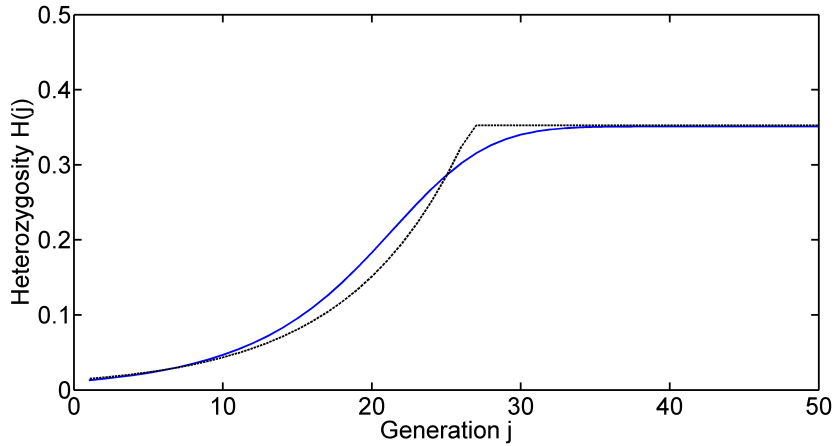
Figure 4.5: Blue line: Fixed Size Leading Order Update. Black line: Update without diffusions, i.e, use 4.39. Here $m = 50$.

The difference between 4.39 and the left half of the "bump" can be attributed to the diffusion terms. Our discrete update essentially gives a numerical scheme of the PDE description. In fact, our discrete update with leading order terms gives the upwind scheme of the advection equation The upwind scheme is known to introduce numerical diffusion. The effect of the diffusion terms could also be observed by modified equation analysis, as in [LL92].

4.4.4 Maximal Principles for Heterozygosity

For a spatially extended network of demes, the bucket model indicates that under some conditions the global diversity of any pair of nuclei isolated from anywhere in the network
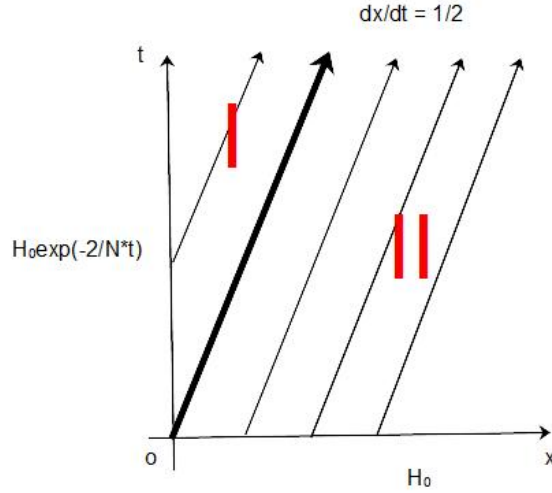
Figure 4.6: Illustration of the characteristics lines when we solve the left half of the "bump" using advection terms. Here $I = \{(x,t)|x < t/2\}$, $II = \{(x,t)|x > t/2\}$.

may not decay to 0. But in fact the individual entries of $H(i,j)$ are not clearly bounded by this constraint. We examine our update rule for evidence of constraints of the heterozygosity of individual demes.

For the discrete leading order term update 4.27, we have

$$\max_{i,j}|H(i,j)|^+ \leq \max_{i,j}|H(i,j)| \tag{4.40}$$

i.e, the maximal of $|H(i,j)|$ is non-increasing.

For our exact update, however, $|H(j-1,j)|$ could have a transient growth despite the system has a decay. Specifically, if we have $H(j-1) = H(j) = H(j-1,j) = H_0$, then in the next step, $H(j-1,j)^+ = (1+c/N)H_0$. As an exact upper bound, we have

$$\max_{i,j}|H(i,j)|^+ \leq (1+c/N)\max_{i,j}|H(i,j)| \tag{4.41}$$

36

4.5 The General Case

We would like to extend the analysis into a more general framework, allowing different types of branching networks to be modeled. Let k_{ij} be the probability that deme $j(j > i)$ is fed by the previous generation(that is by deme $j - 1$) given a division occurs in deme i. Let d_i be the probability that division occurs in deme i that starts the cascade. Then our transition probabilities must be modified to:

$$
< f_j^+ > = \begin{cases} f_j & i \geq j \\ (1 - \dfrac{k_{ij}}{N})f_j + \dfrac{k_{ij}}{N}f_{j-1} & i < j \end{cases} \tag{4.42}
$$

$$
< (f_j^2)^+ > = \begin{cases} f_j^2 & i > j \\ f_j^2 + \dfrac{2}{N^2}f_j(1 - f_j) & i = j \\ (1 - k_{ij})f_j^2 + k_{ij}\left[f_j^2 + \dfrac{1}{N^2}h(j - 1, j) + \dfrac{2}{N}f_j(f_{j-1} - f_j)\right] & i < j \end{cases} \tag{4.43}
$$

$$
< (f_{j-1}f_j)^+ > = \begin{cases} f_{j-1}f_j & i \geq j \\ \\ (1 - k_{j-1,j})f_{j-1}f_j \\ \quad + k_{j-1,j}\left[f_{j-1}f_j + \dfrac{1}{N}f_{j-1}(f_{j-1} - f_j)) - \dfrac{1}{N^2}f_{j-1}(1 - f_{j-1})\right] & i = j - 1 \\ \\ (1 - k_{i,j-1})f_{j-1}f_j \\ \quad + (k_{i,j-1} - k_{ij})\left[f_{j-1} + \dfrac{1}{N}(f_{j-2} - f_{j-1})\right]f_j \\ \quad + k_{ij}[(1 - \dfrac{1}{N})^2 f_{j-1}f_j + (\dfrac{1}{N} - \dfrac{1}{N^2})f_{j-2}f_j \\ \quad + \dfrac{1}{N^2}f_{j-2}f_{j-1} + \dfrac{1}{N}f_{j-1}^2 - \dfrac{1}{N^2}f_{j-1}] & i < j - 1 \end{cases}
$$
$$\tag{4.44}$$

$$< (f_{j-d}f_j)^+ > = \begin{cases} f_{j-d}f_j & i \geq j \\ f_{j-d}\left[(1-k_{ij})f_j + k_{ij}(f_j + \frac{1}{N}(f_{j-1}-f_j))\right] & j-d \leq i < j \\ (1-k_{i,j-d})f_{j-d}f_j & \\ + (k_{i,j-d}-k_{ij})\left[f_{j-d} + \frac{1}{N}(f_{j-d-1}-f_{j-d})\right]f_j & \\ + k_{ij}\left[f_{j-d} + \frac{1}{N}(f_{j-d-1}-f_{j-d})\right]\left[f_j + \frac{1}{N}(f_{j-1}-f_j)\right] & i < j-d \end{cases}$$

$$(4.45)$$

Let $c_j = \frac{1}{N}\sum_{i<j}d_i k_{ij} = c_{j,d} + \hat{c}_{j,d}$, where $c_{j,d} = \frac{1}{N}\sum_{i<j-d}d_i k_{ij}$, $\hat{c}_{j,d} = \frac{1}{N}\sum_{j-d\leq i<j}d_i k_{ij}$, we have

$$E(f_j^+|\text{current state}) = (1-c_j)f_j + c_j f_{j-1}$$

$$E((f_j^2)^+|\text{current state}) = (1-2c_j)f_j^2 + \frac{1}{N^2}d_j h(j) + 2c_j f_{j-1}f_j + \frac{1}{N}c_j h(j-1,j)$$

$$E((f_{j-1}f_j)^+|\text{current state}) = (1-c_{j-1}-c_j+\frac{1}{N}c_{j,1})f_{j-1}f_j + (c_{j-1}-\frac{1}{N}c_{j,1})f_{j-2}f_j$$
$$+ (c_j(1+\frac{1}{N})-\frac{1}{N}c_{j,1})f_{j-1}^2 - \frac{1}{N}c_j f_{j-1} + \frac{1}{N}c_{j,1}f_{j-2}f_{j-1}, (d \geq 2)$$

$$E((f_{j-d}f_j)^+|\text{current state}) = (1-c_{j-d}-c_j+\frac{1}{N}c_{j,d})f_{j-d}f_j + (c_{j-d}-\frac{1}{N}c_{j,d})f_{j-d-1}f_j$$
$$+ (c_j - \frac{1}{N}c_{j,d})f_{j-d}f_{j-1} + \frac{1}{N}c_{j,d}f_{j-d-1}f_{j-1}, (d \geq 2)$$

$$(4.46)$$

So for first moments we have

$$F(j)^+ = (1-c_j)F(j) + c_j F(j-1) \tag{4.47}$$

Whereas for heterozygosity we have

$$H(j)^+ = (1-2c_j-2d_j/N^2)H(j) + 2c_j(1-1/N)H(j-1,j) \tag{4.48}$$

$$H(j-1,j)^+ = (1-c_{j-1}-c_j+c_{j,1}/N)H(j-1,j)$$
$$+ (c_{j-1}-c_{j,1}/N)H(j-2,j) + (c_j(1+1/N)-c_{j,1}/N)H(j-1) \tag{4.49}$$
$$+ (c_{j,1}/N)H(j-2,j-1)$$

$$H(j-d,j)^+ = (1 - c_{j-d} - c_j + c_{j,d}/N)H(j-d,j)$$
$$+ (c_{j-d} - c_{j,d}/N)H(j-d-1,j) + (c_j - c_{j,d}/N)H(j-d,j-1) \quad (4.50)$$
$$+ (c_{j,d}/N)H(j-d,j-1) \quad d \geq 2$$

If we assume that each deme is equally likely to divide, i.e, $d_i = N/(mN + k), \forall i$, all previously described branching models can be obtained as special cases of the above. In particular:

- Full branching local kernel

$k_{ij} = \frac{1}{2}\chi_{\{j-i=1\}}, j > i,$

$c_j = \frac{1}{2}\frac{1}{mN+k}, \forall j,$

$c_{j,d} = 0, d \geq 1.$

- Full branching full kernel

$k_{ij} = \frac{1}{2^{j-i}}, j > i,$

$c_j = \left(1 - 2^{-(j-1)}\right)\frac{1}{mN+k},$

$c_{j,d} = \left(2^{-d} - 2^{-(j-1)}\right)\frac{1}{mN+k}.$

- No branching

$k_{ij} = 1, j > i,$

$c_j = (j-1)\frac{1}{mN+k},$

$c_{j,d} = (j-d-1)\frac{1}{mN+k}.$

4.6   No Branching Case

Note in previous section, we just showed that a lineage without branching is a just special case following the general update rule of H with:

$$k_{ij} = 1, j > i \tag{4.51}$$

$$c_j = (j-1)\frac{1}{mN+k} \tag{4.52}$$

$$c_{j,d} = (j-d-1)\frac{1}{mN+k}. \tag{4.53}$$

Together with the boundary update rule described in the next section, we have an exact update for the expectation of heterozygosity in this case. The result is in Fig 4.7 Note the dynamics is quite different from the full branching case. Although the demes near the root tend to decay with time, the tip demes remain very diverse(heterozygosity≈ 0.45). Furthermore, if we consider the global diversity of the system, which can be calculated as $(\sum_{i,j\leq m} H(i,j))/m^2$, strictly follows the bucket model. This should be expected since no nuclei are lost from the chain of demes in the absence of branching. See Fig 4.8.
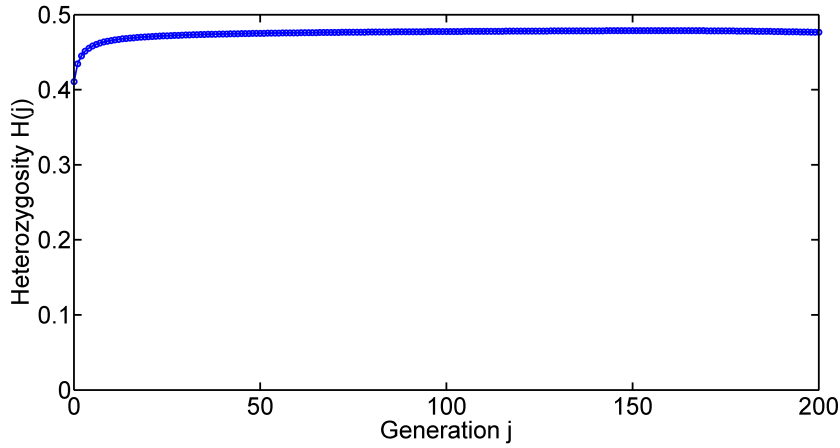
Figure 4.7: Exact update on $H(i,j)$ when there is no branching. Here the system grows from 10 demes to 200 demes.

4.7 Moving Boundary

4.7.1 Boundary Speed

At stage (m, N, k) (i.e, when the system is growing from deme m to deme $m+1$ and k, $0 \leq k \leq N-1$ nuclei are at the new site), For local kernel, we have the probability of a newly division adds to the $(k+1)$-th nucleus equals to $\frac{k+N/2}{mN+k}$, (note for general kernels, we
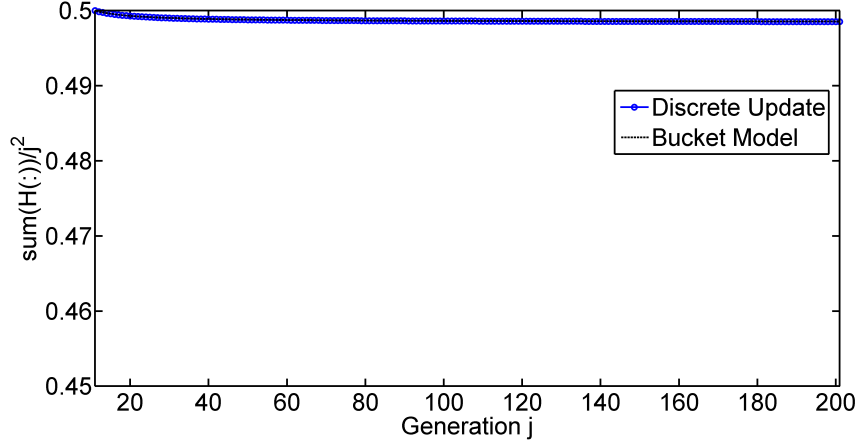
Figure 4.8: The exact update on $H(i,j)$ when there is no branching means $(\sum_{i,j \leq m} H(i,j))/m^2$ follows the bucket model. Here the system grows from 10 demes to 200 demes.

have this probability equals to $\frac{k}{mN+k} + \sum_{i=1}^{m} d_i k_{i,m+1}$ using the notation before), hence the expected number of divisions to grow one more nucleus in the considered chain is $\frac{mN+k}{k+N/2}$. (i.e, follows a geometric distribution).

Note also that we assume the time it takes to divide one more nucleus is proportional to $1/(mN+k)$, i.e, inversely proportional to the total number of nuclei in the system. Therefore, the expected time it takes to grow one more nucleus is $\frac{1}{k+N/2}$. The expected time it takes to grow one more deme is $\sum_{k=0}^{N-1} \frac{1}{k+N/2}$ (note this is independent of m). When N is large enough, this sum is approximately $\int_0^1 \frac{1}{x+1/2} dx = \ln 3$.

Note in the general case, we have the time to grow from m demes to $m+1$ demes equals to the following sum

$$t_{m \to (m+1)} = \sum_{k=0}^{N-1} \frac{1}{k + Nc_{m+1}} \approx \ln\left(\frac{1 + c_{m+1}}{c_{m+1}}\right) \tag{4.54}$$

This derivation gives an estimate of the average speed of the moving boundary. Specifically in the case of full branching local transfer, we have the speed of the boundary approximately equals to $1/\ln 3 \approx 0.91 \approx 1$. That is, the boundary moves faster than the heterozygosity "wave", which travels at speed $1/2$.

41

4.7.2   Update of H on the Boundary in No Branching Case

When there is no branching, we'll make another assumption that no divisions occur in the $(m+1)$-th deme. This simplifies the calculations involved. Under this set of assumptions the number of nuclei in the last deme increases monotonically by 1, with each division occurring anywhere in the first m demes in the linked chain. Under previous notations we have at state (m, N, k), i.e, with probability f_m, $f_{m+1}^+ = (1 + kf_{m+1})/(1+k)$; with probability $1 - f_m$, $f_{m+1}^+ = (kf_{m+1})/(1+k)$.

Hence, fixing m and N, for $k = 0, 1, 2, ..., N-1$, we have

$$< (f_m f_{m+1})^+ > = \begin{cases} \dfrac{1}{1+k} f_m(kf_{m+1} + (1 + 1/N)f_m - 1/N) & i = m \\ \dfrac{1}{1+k}(kf_{m+1}(f_{m-1}/N + (1 - 1/N)f_m) + f_m(f_{m-1}/N + f_m - 1/N)) & i < m \end{cases} \tag{4.55}$$

$$< (f_j f_{m+1})^+ > = \begin{cases} \dfrac{1}{1+k} f_j(f_m + kf_{m+1}) & i \geq j \\ \dfrac{1}{1+k}(f_{j-1}/N + (1 - 1/N)f_j)(f_m + kf_{m+1}) & i < j \end{cases} \tag{4.56}$$

$$\begin{aligned} E(f_{m+1}^+|\text{current state}) &= \frac{1}{1+k}(f_m + kf_{m+1}) \\ E((f_{m+1}^2)^+|\text{current state}) &= \left(\frac{1}{1+k}\right)^2 ((kf_{m+1})^2 + f_m(1 + 2kf_{m+1})) \\ E((f_m f_{m+1})^+|\text{current state}) &= \frac{1}{m}\frac{1}{1+k}f_m(kf_{m+1} + (1 + 1/N)f_m - 1/N) \\ &\quad + \frac{m-1}{m}\frac{1}{1+k}(kf_{m+1}(f_{m-1}/N + (1 - 1/N)f_m) \\ &\quad + f_m(f_{m-1}/N + f_m - 1/N)) \\ E((f_j f_{m+1})^+|\text{current state}) &= \frac{1}{1+k}\left(\frac{j-1}{mN}f_{j-1} + \left(1 - \frac{j-1}{mN}\right)f_j\right)(f_m + kf_{m+1}) \end{aligned} \tag{4.57}$$

So for first order moments we have

$$F(m+1)^+ = \frac{1}{1+k}F(m) + \frac{k}{1+k}F(m+1) \tag{4.58}$$

42

For heterozygosity we have

$$H(m+1)^+ = \frac{k^2}{(1+k)^2} H(m+1) + \frac{2k}{(1+k)^2} H(m, m+1) \tag{4.59}$$

$$
\begin{aligned}
H(m, m+1)^+ = {} & \left(1 - \frac{m-1}{mN}\right) \frac{k}{1+k} H(m, m+1) \\
& + \frac{m-1}{mN} \frac{k}{1+k} H(m-1, m+1) + \left(1 + \frac{1}{mN}\right) \frac{1}{1+k} H(m) \\
& + \frac{m-1}{mN} \frac{1}{1+k} H(m-1, m)
\end{aligned}
\tag{4.60}
$$

While for $j < m$,

$$
\begin{aligned}
H(j, m+1)^+ = {} & \left(1 - \frac{j-1}{mN}\right) \frac{k}{1+k} H(j, m+1) \\
& + \frac{j-1}{mN} \frac{k}{1+k} H(j-1, m+1) + \left(1 - \frac{j-1}{mN}\right) \frac{1}{1+k} H(j, m) \\
& + \frac{j-1}{mN} \frac{1}{1+k} H(j-1, m)
\end{aligned}
\tag{4.61}
$$

4.7.3 Approximations of the Boundary

From the numerical simulations, we observe $H(j, m) \approx H(j, m+1)$ on the moving boundary. This corresponds to the boundary condition

$$\frac{\partial H}{\partial x_1} = \frac{\partial H}{\partial x_2} = 0 \tag{4.62}$$

for the continuous model. We enforce this condition in our update on the expectation of heterozygosity, and the result fits very well with the numerical simulations. See Figure 4.12.

4.7.4 "Bump" Revisited

Now that if the boundary condition is valid, we can actually explain the right half of the bump. We claim that approximately the heterozygosity on the right end of the moving boundary decays exponentially, but with a decay rate equals to $1/N$, i.e, half of the decay rate at the origin. Too see this, let the heterozygosity on the right boundary be $H_{bdd}(t) =$

$H(t, L_0 + vt, L_0 + vt)$, where L_0 is the initial length of the system, v be the speed of the boundary. Consider the this quantity with the PDE update without diffusion,

$$\frac{\partial H}{\partial t} = -\frac{1}{2}\left(\frac{\partial H}{\partial x_1} + \frac{\partial H}{\partial x_2}\right) - (2/N)\delta(x_2 - x_1)H(x_1, x_2, t) \tag{4.63}$$

with

$$\frac{dH_{bdd}}{dt} \approx v\frac{\partial H}{\partial x_1} \approx v\frac{\partial H}{\partial x_2} \tag{4.64}$$

approximately true on the boundary, we see $H_{bdd}(t)$ satisfies

$$\frac{dH_{bdd}}{dt} = -\frac{2}{N}\frac{1}{1 + 1/v}H_{bdd}(t) \tag{4.65}$$

From the previous derivation on the speed of the boundary, we know for full branching local transfer systems v is a constant and $v \approx 1$. Therefore,

$$\frac{dH_{bdd}}{dt} = -\frac{1}{N}H_{bdd}(t) \tag{4.66}$$

and $H_{bdd}(t) = e^{-(t/N)}H_{bdd}(0)$. See Figure 4.9 for a justification for this claim.
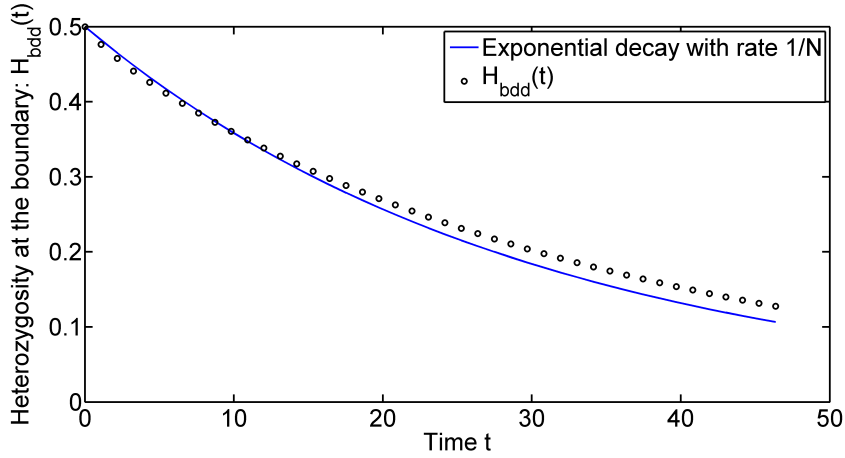
Figure 4.9: Black dots: Heterozygosity decay at the boundary. Blue line: Exponential decay $e^{-(t/N)}H_0$.

We can then explain the right half of the bump.

Let us say the initial boundary is $[0, L_0]$. The key is that the moving boundary, which travels with speed 1, is also generating new characteristic lines(which travels with speed

44
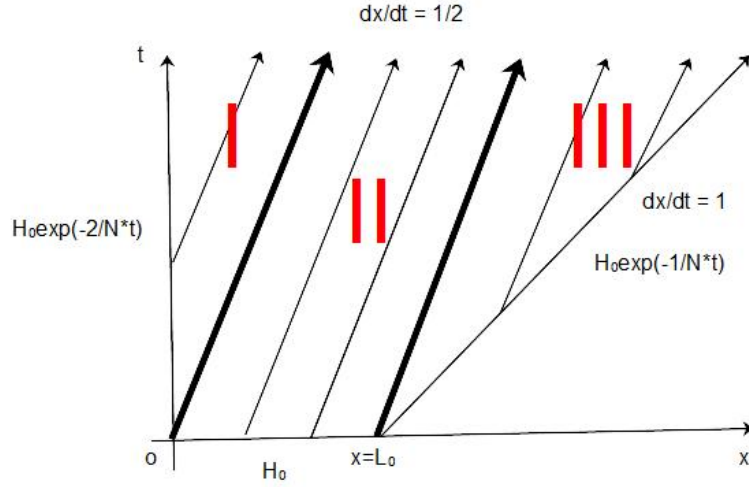
Figure 4.10: Illustration of the characteristics lines when we solve the "bump" using advection terms. Here $I = \{(x,t)|0 < x < t/2\}$, $II = \{(x,t)|t/2 < x < L_0 + t/2\}$, $III = \{(x,t)|L_0 + t/2 < x < L_0 + t\}$.

1/2). We then apply the same idea as the analysis the left half of the bump. See 4.10 for an illustration of the characteristics.

$$H(t,x) = \begin{cases} H_0 e^{-\frac{2}{N}(t-2x)} A(2x), & 0 \leq x \leq t/2 \\ H_0 A(t), & t/2 < x < L_0 + t/2 \\ H_0 e^{-\frac{1}{N}t^*} A(t - t^*), & L_0 + t/2 \leq x \leq L_0 + t \end{cases} \quad (4.67)$$

where $t^* = 2t - 2(x - L_0)$.

See Figure 4.11 and 4.12. for a comparison with the exact solution. Again, the difference is because we ignore the diffusion terms.
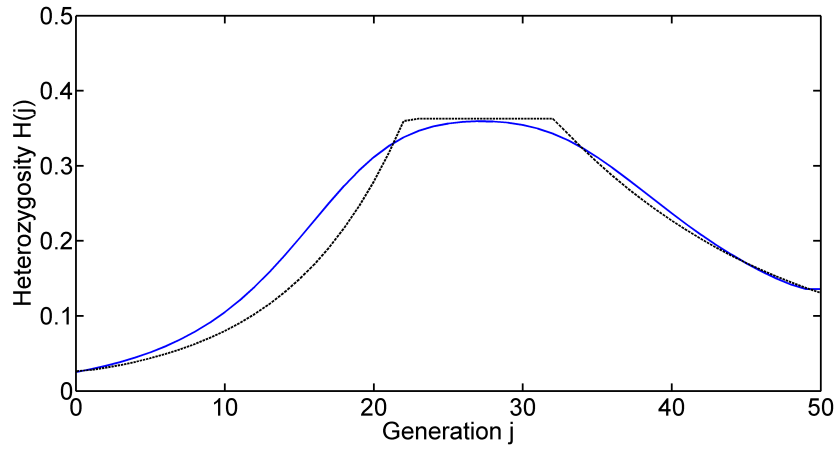
Figure 4.11: Blue line: Leading Order Update. Black line: Update without diffusions. Here the system grows from 10 demes to 50 demes, $T \approx 44$.

Figure 4.12: Comparison: Average of realization, fixed size update, free boundary update, Stepping Stone solution. Here the system grows from 10 demes to 200 demes, $T \approx 210$.

CHAPTER 5

Coupling Branching and Growth

5.1 Connections with Experiment Data

In Chapter 4, we showed that when there is no branching, the heterozygosity does not decay either locally or globally; whereas when there is full branching, the heterozygosity travels with a "bump" profile, with the peak of the bump decays like $t^{-1/2}$.

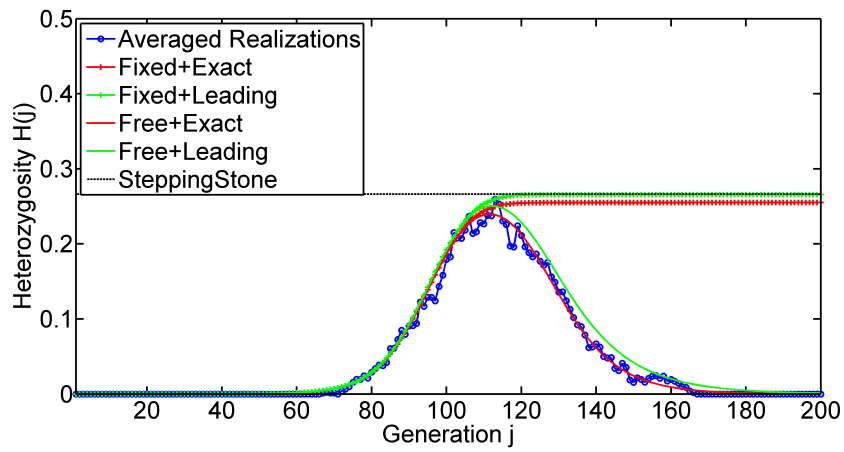Real hyphal networks are somewhere intermediate between these two cases. Interestingly the amount of branching that occurs is necessarily dependent on whether the mycelial network has an expanding periphery, or grows in one dimension along a racetube, which were the two experiment geometries used by [MSC15].

If the mycelium grows in one-dimension then no new hyphal tips need to be generated to maintain uniform density with growth, so in theory no branching is needed. For a mycelium growing in two-dimensions, for the hp-to-bp separation to stay constant with distance growth, new tips must be continuously produced via hyphal branching.

Ma et al. [MSC15] found that the spore diversity was sensitive to the geometry of cell growth. Specifically, for the race-tube growth, which corresponds to the "no branching" case in our model, the experiment shows the genotypes were well mixed even at the scale of individual hyphal compartments, this is in general the same prediction of our "no branching" model. For fungal that grow in petri dishes, where branching is unavoidable, the experiment shows that individual compartments had different genetic composition, this also is in line with the prediction of our "full branching" model.

However to populate a linearly growing front with a uniform density of hps branching

does not need to occur uniformly since this would produce exponential proliferation of hps. In fact the number of branches needs only increase as the logarithm of the number of demes, a case that we have not previously considered.

In this chapter we use numerical simulations to determine how finite or logarithmic branching densities affect population dynamics, with the goal of making a quantitative comparison with the experimental data of [MSC15].

5.2   Exploring Different Network Structures

In real fungal systems like *Neurospora crassa*, we observe that the network structure is somewhere between the "full branch" case and the "no branching" case. Therefore, we would like to know how the dynamics of heterozygosity are in those intermediate cases. Here we show two numerical results regarding two structures that may of interests:

- Only branching around the tips. That is, only the last m_t generations are allowed to branch. Here m_t could be a constant or could vary with chain length(e.g, $m_t \sim O(\log m)$).

- Only branching around the root. That is, only the first m_r generations are allowed to branch. Here m_r could be a constant or could vary with chain length(e.g, $m_r \sim O(\log m)$).

We simulate these two cases with our discrete exact update of heterozygosity on a single lineage. The result is shown in Figure 5.2 and 5.2.

We observe that (i) branching around the tips would result a loss of heterozygosity; (ii) branching around the root seems to preserve the heterozygosity. In fact, this case is comparable to the "no branching" case, which we showed in the Chapter 4.

To gain insight into the heterozygosity dynamics on these two cases, analysis is needed using the mechanism developed in Chapter 4 to better understand these observations. This will be our future work.
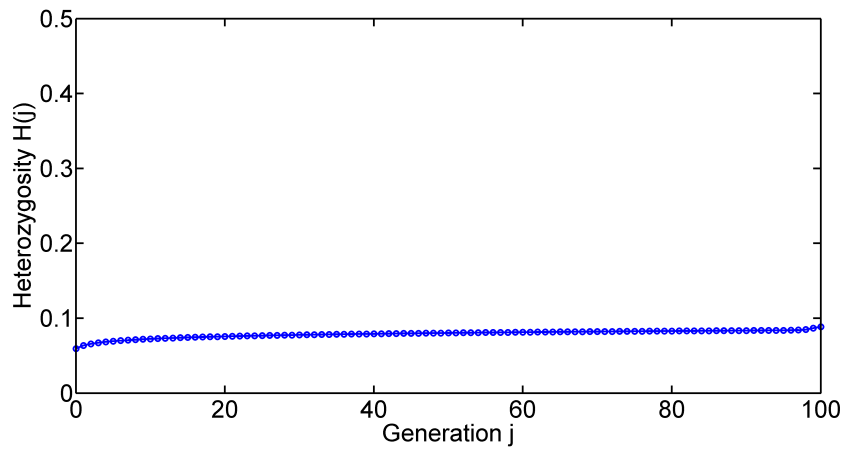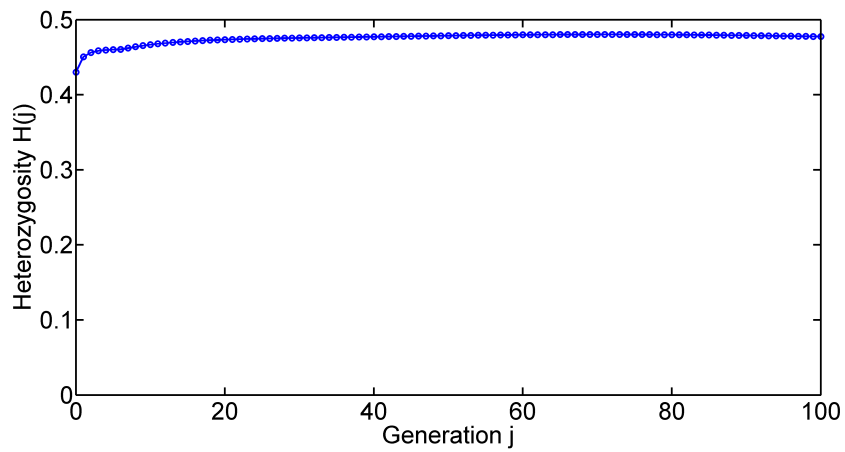
Figure 5.1: Only branching around the tips. $m_t = 6$.

Figure 5.2: Only branching around the root. $m_r = 6$.

REFERENCES

[ARK09] Thomas E Angelini, Marcus Roper, Roberto Kolter, David A Weitz, and Michael P Brenner. "Bacillus subtilis spreads by surfing on waves of surfactant." *Proceedings of the National Academy of Sciences*, **106**(43):18109–18113, 2009.

[BB91]   Elena O Budrene and Howard C Berg. "Complex patterns formed by motile cells of Escherichia coli." 1991.

[Ber08] Howard C Berg. *E. coli in Motion*. Springer Science & Business Media, 2008.

[Bla11] Meredith Blackwell. "The Fungi: 1, 2, 3 5.1 million species?" *American journal of botany*, **98**(3):426–438, 2011.

[Bus14] Leo W Buss. *The evolution of individuality*. Princeton University Press, 2014.

[CK70]   James F Crow, Motoo Kimura, et al. "An introduction to population genetics theory." *An introduction to population genetics theory.*, 1970.

[DMS03] Charles R Doering, Carl Mueller, and Peter Smereka. "Interacting particles, the stochastic Fisher–Kolmogorov–Petrovsky–Piscounov equation, and duality." *Physica A: Statistical Mechanics and its Applications*, **325**(1):243–259, 2003.

[DTM05] Willow R DiLuzio, Linda Turner, Michael Mayer, Piotr Garstecki, Douglas B Weibel, Howard C Berg, and George M Whitesides. "Escherichia coli swim on the right-hand side." *Nature*, **435**(7046):1271–1274, 2005.

[FLR09] André Fleissner, Abigail C Leeder, M Gabriela Roca, Nick D Read, and N Louise Glass. "Oscillatory recruitment of signaling proteins to cell tips promotes coordinated behavior during cell fusion." *Proceedings of the National Academy of Sciences*, **106**(46):19387–19392, 2009.

[FSG08] André Fleißner, Anna R Simonin, and N Louise Glass. "Cell fusion in the filamentous fungus, Neurospora crassa." In *Cell Fusion*, pp. 21–38. Springer, 2008.

[FSJ05] André Fleißner, Sovan Sarkar, David J Jacobson, M Gabriela Roca, Nick D Read, and N Louise Glass. "The so locus is required for vegetative cell fusion and post-fertilization events in Neurospora crassa." *Eukaryotic Cell*, **4**(5):920–930, 2005.

[GD06]   N Louise Glass and Karine Dementhon. "Non-self recognition and programmed cell death in filamentous fungi." *Current opinion in microbiology*, **9**(6):553–558, 2006.

[GFL04] Van D Gooch, Laura Freeman, and Patricia L Lakin-Thomas. "Time-lapse analysis of the circadian rhythms of conidiation and growth rate in Neurospora." *Journal of biological rhythms*, **19**(6):493–503, 2004.

[GFV00] Naama Goren-Inbar, Craig S Feibel, Kenneth L Verosub, Yoel Melamed, Mordechai E Kislev, Eitan Tchernov, and Idit Saragusti. "Pleistocene milestones on the out-of-Africa corridor at Gesher Benot Ya'aqov, Israel." *Science*, **289**(5481):944–947, 2000.

[GGR13] Jeff Gole, Athurva Gore, Andrew Richards, Yu-Jui Chiu, Ho-Lim Fung, Diane Bushman, Hsin-I Chiang, Jerold Chun, Yu-Hwa Lo, and Kun Zhang. "Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells." *Nature biotechnology*, **31**(12):1126–1132, 2013.

[GHP06] Amy S Gladfelter, A Katrin Hungerbuehler, and Peter Philippsen. "Asynchronous nuclear division cycles in multinucleated cells." *The Journal of cell biology*, **172**(3):347–362, 2006.

[GJS00] N Louise Glass, David J Jacobson, and Patrick KT Shiu. "The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi." *Annual review of genetics*, **34**(1):165–186, 2000.

[Gla04] NL Glass et al. "in filamentous fungi. Trends in Microbiology 9: 553-558." *Trends in Microbiology*, **12**(3):135–141, 2004.

[Gla06] Amy S Gladfelter. "Nuclear anarchy: asynchronous mitosis in multinucleated fungal hyphae." *Current opinion in microbiology*, **9**(6):547–552, 2006.

[GM62] Jack P Gibbs and Walter T Martin. "Urbanization, technology, and the division of labor: International patterns." *American Sociological Review*, pp. 667–677, 1962.

[Gou02] Stephen Jay Gould. *The structure of evolutionary theory.* Harvard University Press, 2002.

[Hah66] George M Hahn. "State Vector Description of the Proliferation of Mammalian Cells in Tissue Culture: I. Exponential Growth." *Biophysical journal*, **6**(3):275, 1966.

[HC13] Dan Hu and David Cai. "Adaptation and optimization of biological transport networks." *Physical review letters*, **111**(13):138701, 2013.

[Hil73] Mark O Hill. "Diversity and evenness: a unifying notation and its consequences." *Ecology*, **54**(2):427–432, 1973.

[HK09] Oskar Hallatschek and KS Korolev. "Fisher waves in the strong noise limit." *Physical Review Letters*, **103**(10):108103, 2009.

[HKR98] Michael F Hammer, Tatiana Karafet, Arani Rasanayagam, Elizabeth T Wood, Tasha K Altheide, Trefor Jenkins, Robert C Griffiths, Allan R Templeton, and Stephen L Zegura. "Out of Africa and back again: nested cladistic analysis of human Y chromosome variation." *Molecular Biology and Evolution*, **15**(4):427–441, 1998.

[HN08]   Oskar Hallatschek and David R Nelson. "Gene surfing in expanding populations." *Theoretical population biology*, **73**(1):158–170, 2008.

[Hur71]   Stuart H Hurlbert. "The nonconcept of species diversity: a critique and alternative parameters." *Ecology*, **52**(4):577–586, 1971.

[Jos06]   Lou Jost. "Entropy and diversity." *Oikos*, **113**(2):363–375, 2006.

[KAH10]   KS Korolev, Mikkel Avlund, Oskar Hallatschek, and David R Nelson. "Genetic demixing and evolution in linear stepping stone models." *Reviews of modern physics*, **82**(2):1691, 2010.

[KKB06]   Jörg Kämper, Regine Kahmann, Michael Bölker, Li-Jun Ma, Thomas Brefort, Barry J Saville, Flora Banuett, James W Kronstad, Scott E Gold, Olaf Müller, et al. "Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis." *Nature*, **444**(7115):97–101, 2006.

[KN11]   KS Korolev and David R Nelson. "Competition and cooperation in one-dimensional stepping-stone models." *Physical Review Letters*, **107**(8):088103, 2011.

[KW64]   Motoo Kimura and George H Weiss. "The stepping stone model of population structure and the decrease of genetic correlation with distance." *Genetics*, **49**(4):561, 1964.

[Lew05]   Roger R Lew. "Mass flow and pressure-driven hyphal extension in Neurospora crassa." *Microbiology*, **151**(8):2685–2692, 2005.

[LL92]   Randall J LeVeque and Randall J Le Veque. *Numerical methods for conservation laws*, volume 132. Springer, 1992.

[Mah05]   Ramesh Maheshwari. "Nuclear behavior in fungal hyphae." *FEMS microbiology letters*, **249**(1):7–14, 2005.

[Mor58]   Patrick Alfred Pierce Moran. "Random processes in genetics." In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 54, pp. 60–71. Cambridge Univ Press, 1958.

[MR63]   Panchanan Maheshwari and NS Rangaswamy. "Plant tissue and organ culture-a symposium." *Int. Soc. Pl. Morph., Delhi. Review article General article, Tissue culture (PMBD, 185306652)*, 1963.

[MRB06]   Rosa R Mouriño-Pérez, Robert W Roberson, and Salomon Bartnicki-García. "Microtubule dynamics and organization during hyphal growth and branching in Neurospora crassa." *Fungal Genetics and Biology*, **43**(6):389–400, 2006.

[MSC15]   Linda Ma, Boya Song, Thomas Curran, Nhu Phong, Emilie Dressaire, and Marcus Roper. "Defining individual size in the model filamentous fungus Neurospora crassa." *arXiv preprint arXiv:1510.08965*, 2015.

[OHL06] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. "A simple rule for the evolution of cooperation on graphs and social networks." *Nature*, **441**(7092):502–505, 2006.

[PG44] G Pontecorvo and AR Gemmell. "Colonies of Peniciluum Notatum and other Moulds as Models for the Study of Population Genetics." *Nature*, **154**:532–534, 1944.

[Pon59] Guido Pontecorvo et al. "Trends in genetic analysis." *Trends in genetic analysis.*, 1959.

[Pon75] Guido Pontecorvo. "Production of mammalian somatic cell hybrids by means of polyethylene glycol treatment." *Somatic cell genetics*, **1**(4):397–400, 1975.

[RET11] Marcus Roper, Chris Ellison, John W Taylor, and N Louise Glass. "Nuclear and genome dynamics in multinucleate ascomycete fungi." *Current Biology*, **21**(18):R786–R793, 2011.

[RLH15] Marcus Roper, ChangHwan Lee, Patrick C Hickey, and Amy S Gladfelter. "Life as a moving fluid: fate of cytoplasmic macromolecules in dynamic fungal syncytia." *Current opinion in microbiology*, **26**:116–122, 2015.

[RSH13] Marcus Roper, Anna Simonin, Patrick C Hickey, Abby Leeder, and N Louise Glass. "Nuclear dynamics in a fungal chimera." *Proceedings of the National Academy of Sciences*, **110**(32):12875–12880, 2013.

[SH91] Montgomery Slatkin and Richard R Hudson. "Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations." *Genetics*, **129**(2):555–562, 1991.

[SOW06] Rob Sturman, Julio M Ottino, and Stephen Wiggins. *The mathematical foundations of mixing: the linked twist map as a paradigm in applications: micro to macro, fluids to solids*, volume 22. Cambridge University Press, 2006.

[SRY10] Anna R Simonin, Carolyn G Rasmussen, Mabel Yang, and N Louise Glass. "Genes encoding a striatin-like protein (ham-3) and a forkhead associated protein (ham-4) are required for hyphal fusion in Neurospora crassa." *Fungal Genetics and Biology*, **47**(10):855–868, 2010.