AN ABSTRACT OF THE DISSERTATION OF

Torrey Johnson for the degree of Doctor of Philosophy in Mathematics presented on June 7, 2012.

Title: Branching Random Walk and Probability Problems from Physics and Biology

Abstract approved: _____

Edward C Waymire

This thesis studies connections between disorder type in tree polymers and the branching random walk and presents an application to swarm site-selection. Chapter two extends results on tree polymers in the infinite volume limit to critical strong disorder. Almost sure (a.s.) convergence in the infinite volume limit is obtained for weak disorder by standard theory on multiplicative cascades or the branching random walk. Chapter three establishes results for a simple branching random walk in connection with a related tree polymer. A central limit theorem (CLT) is shown to hold regardless of polymer disorder type, and a.s. connectivity of the support is established in the asymmetric case. Chapter four contains a model for site-selection in honeybee swarms. Simulations demonstrate a trade-off between speed and accuracy, and strongly suggest that increasing the quorum threshold at which the process terminates usually improves decision performance.

Branching Random Walk and Probability Problems from Physics
and Biology

by

Torrey Johnson

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 7, 2012
Commencement June 2013

Doctor of Philosophy dissertation of Torrey Johnson presented on June 7, 2012.

APPROVED:

_____

Major Professor, representing Mathematics

_____

Chair of the Department of Mathematics

_____

Dean of the Graduate School

_____

Torrey Johnson, Author

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

Branching Random Walk and Probability Problems from Physics and Biology

1   Introduction

This thesis presents results for a class of branching random walks and explores connections with a related polymer model known as the tree polymer. A model for honey bee site-selection is also presented. I want to emphasize that the latter is not a branching random walk model, but does incorporate some elements from branching processes in general (and random spatial movement), and these connections will facilitate some results.

Branching processes have a long and rich history in both pure and applied areas, and they have been applied extensively to models in the life sciences. Indeed, the subject began with Galton's study of the extinction of noble family lines [22]. I do not recount it all here. However, Kingman's work on the branching random walk in [27] is very relevant, as are certain later developments. Taking this as starting point, I provide a brief account of the early branching random walk theory initiated by Kingman and substantially carried forward by Biggins and others, and discuss some recent and pertinent results.

The basic object is the one-dimensional branching random walk in discrete time. A single particle, the *ancestor*, is initially at the origin. At time 1 the particle dies, simultaneously giving birth to some number of new particles according to some offspring distribution. Consistent with the language of genealogy, particles that

give rise to other particles are called *parent* particles or simply *parents*, and the particles they generate *child* particles or *children*. Each of the child particles position themselves on the real line independently according to some point process describing displacements from the parent. The child particles repeat this process, and the process continues as long as particles are alive.

The event that a positive number of particles are alive in every generation is called the survival set or ultimate survival. By a standard theorem, this event has positive probability if and only if the underlying Galton-Watson process, the branching process obtained by ignoring the spatial distribution of particles, is supercritical, i.e. the mean number of offspring per particle exceeds one. Super-criticality will be frequently assumed.

One of the main investigations on the branching random walk concerns the asymptotics of its spreading speed. To state results, suppose $\mathbb{P}$ is the probability measure for the process, $\mathbb{E}$ its expectation, and Z the point process in $\mathbb{R}$. Let $Z^{(n)}$ be the set of positions of the nth generation particles and let $B^{(n)}$ denote the rightmost particle in $Z^{(n)}$. Biggins [11] generalized the work in [27] to prove a law of large numbers for $B^{(n)}$,

$$\frac{B^{(n)}}{n} \to \Gamma \quad \text{a.s.,} \tag{1.0.1}$$

where Γ is a constant, under the condition that $\kappa(\phi) := \log \int e^{\phi z} \mathbb{E} Z(dz) < \infty$ for some $\phi > 0$. The speed Γ can be computed by

$$\Gamma = \sup\{a : \kappa^*(a) < 0\} = \inf_\theta \left\{ \frac{\kappa(\theta)}{\theta} \right\}$$

where $\kappa^*(a) = \sup_\theta\{\theta a - \kappa(\theta)\}$ is the Fenchel dual of κ [8]. An early version of this result was also shown by Hammersley in [20].

In a paper published the following year, Biggins [5] proved a Chernoff-type theorem characterizing particle growth in the intervals $[na, \infty)$. Assume there is a $\phi > 0$ such that $\kappa(\phi) < \infty$. Then for all $a \neq \Gamma$, $\frac{1}{n}\log(Z^{(n)}[na, \infty))$ converges to either $-\kappa^*(a)$ (if $\kappa(a) \leq 0$) or $-\infty$ (if $\kappa(a) > 0$). In mean, the convergence is to $-\kappa^*(a)$, except possibly at one a. The previous result on the speed of the rightmost particle follows immediately from this theorem and the properties of κ^*. Both this result and (1.0.1) generalize to the multitype case, although for the Chernoff-type result this was established much later, again by Biggins [9].

The rate of convergence to Γ has also been extensively studied. Biggins [7] proved that $B^{(n)} - n\Gamma \to -\infty$ almost surely if and only if there is some $0 < \nu < \infty$ such that $\kappa(\nu) = \nu\Gamma$, assuming displacements larger that Γ have positive probability. An important special case not covered by this result is considered by Dekking and Host [18]: if $\{X_n\}$ is the minimal displacement of a supercritical branching random walk with non-negative, integer-valued displacements, then X_n converges to an a.s. finite random variable X. Under a moment condition, they prove that the centered sequence $\{X_n - \mathbb{E}(X_n)\}$ is tight. In the case of the simple branching random walk with 0-1 displacements, each having probability one-half, Grill [19] has proven that the support of the walk is eventually, almost surely, an interval $[K, n - L]$, for finite random variables K and L. This result will be extended in chapter 3.

Certain martingales arise frequently in the analysis of the branching random walk. The so-called *additive martingale* is $W_n = \sum_{|x|=n} \exp(-z_x)$, where z_x denotes the position of a child labeled x and $|x| = n$ means x is in generation n. Since it is non-negative, it converges to an a.s. finite random variable. For convergence in mean, there is a Kesten-Stigum type result shown by Biggins in [6]: if $\mathbb{E}W_1 > 0$, then W_n converges in L_1 if and only $\mathbb{E}(W_1 \log^+ W_1) < \infty$. In this case, $\mathbb{E}(W) = 1$, so the limit W is non-zero. The *derivative martingale*, introduced in [10], is defined by $D_n = \sum_{|x|=n} z_x \exp(-z_x)$. The branching random walk is said to be in the *boundary* case if $\mathbb{E}W_1 = 1$ and $\mathbb{E}D_1 = 0$. In this case, a very recent result in [2] has established convergence in probability, under integrability conditions, for the ratio of the additive and derivative martingales:

$$\lim_{n\to\infty} n^{-\frac{1}{2}} \frac{W_n}{D_n} = \left(\frac{2}{\pi\sigma^2}\right)^{1/2} \quad \text{in probability,}$$

for the positive constant $\sigma^2 := \mathbb{E}\left[\sum_{|x|=1} z_x^2 e^{-z_x}\right]$. The result cannot be strengthened to almost sure convergence [2]. This will be the key to proving the main result in chapter 2.

For tree polymers, self-contained introductions to the notation and precise definitions are found in the relevant chapters, so I do not repeat them here in detail, but simply outline the main ideas. The vertices of a complete binary tree $T = \cup_{n=0}^{\infty}\{-1, 1\}^n$ are equipped with i.i.d. copies of a strictly positive mean-one random variable X. Take products of these variables down a single path of the tree to level n, and let Z_n be the sum of all such products scaled by 2^{-n}. Z_n

converges a.s. to a random variable Z, and by a theorem established by Kahane and Peyrière [26], $Z_n \to Z > 0$ a.s. if and only if $\mathbb{E}(X \ln X) < \ln 2$, and $Z_n \to Z = 0$ a.s. otherwise. Bolthausen [13] introduced the terms *weak disorder* and *strong disorder*, respectively, to describe this dichotomy.

These notions have analogues in the theory of multiplicative cascades. Suppose T is the metric space consisting of the complete b-ary tree, $b \geq 2$, equipped with the ultrametric that assigns the distance between two tree paths to be b^{-m}, where m is the smallest tree level at which the two paths no longer coincide. Multiplicative cascades are the positive T-martingales (introduced in [25]) represented multiplicatively via martingale ratios [45]. The factors are non-negative mean one random variables, referred to as *generators* of the cascade. The Kahane-Peyrière theorem applies to multiplicative cascades with i.i.d. generators, the *homogeneous independent cascades*. If X denotes the generator, as above with tree polymers, then the cascade is said to a.s. *survive* if and only if the condition for weak disorder holds, and to a.s. *die out* otherwise (strong disorder). Using size-biasing techniques, Waymire and Williams provide an entirely probabilistic proof of the Kahane-Peyrière theorem, and provide an extension to dependent cascades, see [43], [44], [45].

Returning to the tree polymer, a sequence of random probability measures $prob_n$ on the space of tree polymer paths will be defined by norming the products down paths in a natural way, leading immediately to convergence questions and the extent of dependence on disorder type. It is well-known that $prob_n$ converges to an a.s. unique limit in the case of weak disorder [46]. For strong disorder the question

is open, but the special case of critical strong disorder, i.e. $\mathbb{E}(X \ln X) = \ln 2$, is taken up in chapter 2. Critical strong disorder for the tree polymer corresponds to the boundary case for the branching random walk. Defining $(S)_n$ to be the path position at level n (sum of the ± 1 path coordinates), another natural problem is to prove a central limit theorem similar to the classical theorem for the simple symmetric random walk, but with respect to the measures $prob_n$. In [46], under a mild additional assumption, this was shown in the full range of weak disorder: $n^{-\frac{1}{2}}(S)_n$ a.s. converges weakly to a standard normal distribution. In chapter 3, an a.s. central limit theorem is established for a simple branching random walk that is easily related to a certain tree polymer, and there is no dependence on disorder type in this very special case.

Tree polymers are naturally related to branching random walks in the following manner. A product of i.i.d. positive random variables down a tree polymer path to level n can always be expressed as the exponential of the sum of the logarithms of those variables. The sum in the exponent may be regarded as the sum of n i.i.d. displacements. In the discrete-time branching random walk, this is the position of one of the nth generation offspring particles. Via this connection, results for the branching random walk may be applied to tree polymers and vice versa.

More generally, polymer models usually refer to directed lattice polymers on the $d+1$ dimensional integer lattice, which are modeled by random path distributions in $\mathbb{N} \times \mathbb{Z}^d$ [46]. The first coordinate directs the path, ensuring that the polymer paths are self-avoiding. For tree polymers, the directed paths are the polygonal paths of a binary tree. Tree polymers were considered in [14] in the analysis of

directed lattice polymer models [46]. More information in this connection can be found in [46].

The organization of this thesis is as follows. The second chapter is based on a paper published in the *Journal of Applied Probability* [24]. It investigates convergence of probability measures on tree polymers. Almost sure convergence to a polymer probability in the infinite volume limit is easily obtained under weak disorder. The main goal is to prove, by exploiting certain connections with the branching random walk, convergence in probability for the critical strong disorder case. The third chapter, based on a preprint manuscript, establishes a central limit theorem and a result on connectivity of the support of a simple branching random walk, and relates the results to tree polymer models in an effort to understand the effects of disorder type. The fourth chapter contains a branching process model for site-selection in honey bee swarms. Its goal is to gain insight into the collective-decision making process, and investigate conditions under which the swarm makes optimal choices. The model involves branching growth in time and random movement in space, similar to a branching random walk, but the usual branching random walk requirement of i.i.d. displacements to determine spatial movement is replaced by a more realistic assumption.

2   Tree Polymers in the Infinite Volume Limit at Critical Strong Disorder

2.1   Introduction and Preliminaries

Polymers are abstractions of chains of molecules embedded in a solvent by non-self-intersecting polygonal paths of points whose probabilities are themselves random (reflecting impurities of the solvent). In this connection, tree polymers take advantage of a particular way to determine path structure and their probabilities as follows.

Three different references to paths occur in this formulation. An ∞-*tree path* is a sequence $s = (s_1, s_2, \ldots) \in \{-1, 1\}^{\mathbb{N}}$ emanating from a root 0. A *finite tree path* or *vertex* v is a finite sequence $v = s|n = (s_1, \ldots, s_n)$, read "path s restricted to level n", of length $|v| = n$. The symbol $*$ denotes concatenation of finite tree paths; if $v = (v_1, \ldots, v_n)$ and $t = (t_1, \ldots, t_m)$, then $v * t = (v_1, \ldots, v_n, t_1, \ldots, t_m)$. Vertices belong to $T := \bigcup_{n=0}^{\infty} \{-1, 1\}^n$, and can be viewed as unique finite paths to the root of the directed binary tree T equipped with the obvious graph structure. We also write

$$\partial T = \{-1, 1\}^{\mathbb{N}}$$

for the boundary of T. The third type of path, and the one of main interest to polymer questions, is that of the *polygonal tree path* defined by $n \to (s)_n := \sum_{j=1}^{n} s_j$,

$n \geq 0$, with $(s)_0 := 0$, for a given $s \in \partial T$.

∂T is a compact, topological Abelian group for coordinate-wise multiplication and the product topology. The *uniform distribution* on ∞-tree paths is the Haar measure on $(\partial T, \mathcal{B})$, i.e.

$$\lambda(ds) = \left(\frac{1}{2}\delta_+(ds) + \frac{1}{2}\delta_-(ds)\right)^{\mathbb{N}}.$$

Let $\{X_v : v \in T\}$ be an i.i.d. family of positive random variables on $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}X < \infty$; we denote a generic random variable with the common distribution of X_v by X. Without loss of generality we may assume that $\mathbb{E}X = 1$. Define a sequence of *random probability measures* $\mathrm{prob}_n(ds)$ on $(\partial T, \mathcal{B})$ by the prescription that

$$\mathrm{prob}_n(ds) << \lambda(ds)$$

with

$$\frac{d\mathrm{prob}_n}{d\lambda}(s) = Z_n^{-1}\prod_{j=1}^n X_{s|j}$$

where

$$Z_n = \int_{\partial T}\prod_{j=1}^n X_{s|j}\lambda(ds) = \sum_{|s|=n}\prod_{j=1}^n X_{s|j}2^{-n}.$$

Observing that $\{Z_n : n = 1, 2\ldots\}$ is a positive martingale, it follows that

$$Z_\infty := \lim_{n\to\infty} Z_n$$

exists a.s. in $(\Omega, \mathcal{F}, P)$. According to a classic theorem of Kahane and Peyrière

[26] in the context of multiplicative cascades, and Biggins [11] in the context of branching random walks, one has the following dichotomy:

$$P(Z_\infty > 0) = 1 \quad \Longleftrightarrow \quad \mathbb{E}X \ln X < \ln 2$$

$$P(Z_\infty = 0) = 1 \quad \Longleftrightarrow \quad \mathbb{E}X \ln X \geq \ln 2.$$

The a.s. occurence of the event $[Z_\infty > 0]$ is refered to as *weak disorder* and that of $[Z_\infty = 0]$ as *strong disorder*; see Bolthausen [13]. In particular, the critical case $\mathbb{E}X \ln X = \ln 2$ is strong disorder. In the case of tree polymers one may view the notions of weak/strong in terms of a disorder parameter defined by $\mathbb{E}X \ln X$ and relative to the branching rate, $\ln 2$.

In this chapter some new insights into a few delicate problems for the case of strong disorder are provided.

2.2   Tree Polymers under Weak Disorder

To set the stage for contrast, we record a rather robust consequence of weak disorder.

Theorem 2.2.1. *Under weak disorder, there is a random probability measure* $\mathrm{prob}_\infty(ds)$ *on* $(\partial T, \mathcal{B})$ *such that a.s.*

$$\mathrm{prob}_n(ds) \Rightarrow \mathrm{prob}_\infty(ds)$$

where $\Rightarrow$ *denotes weak convergence.*

Proof. Define $\lambda_n(ds) = Z_n\mathrm{prob}_n(ds)$, $n = 1, 2, \ldots$. By Kahane's T-martingale theory, e.g. [25], $\lambda_n(ds)$ converges vaguely to a non-zero random measure $\lambda_\infty(ds)$ on $(\partial T, \mathcal{B})$ with probability one. By definition of weak disorder $Z_n \to Z_\infty > 0$ a.s., thus we obtain

$$\mathrm{prob}_n(ds) = Z_n^{-1}\lambda(ds) \Rightarrow Z_\infty^{-1}\lambda_\infty(ds) \quad \text{a.s.}$$

$\square$

Notice that in the case of no disorder, i.e. $X = 1$ a.s., one has

$$\mathrm{prob}_n(ds) = \lambda(ds) \quad \forall n = 1, 2, \ldots.$$

Moreover, under $\lambda(ds)$, the polygonal paths are simply symmetric simple random walk paths, where the probability theory is quite will-known and complete. For example, the central limit theorem takes the form

$$\lim_{n \to \infty} \lambda\left(\left\{s \in \partial T : \frac{(s)_n}{\sqrt{n}} \leq x\right\}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\xi^2/2}d\xi.$$

For probability laws involving convergence in distribution, one may ask if the CLT continues to hold a.s. with $\lambda(ds)$ replaced by $\mathrm{prob}_n(ds)$. This form of universality was answered in the affirmative by Waymire and Williams in [46] for weak disorder under the additional assumption that $\mathbb{E}X^{1+\delta} < \infty$ for some $\delta > 0$. Problems involving limit laws such as a.s. strong laws, a.s. laws of the iterated logarithm, etc, however, require an infinite volume probability $\mathrm{prob}_\infty(ds)$ for their formulation. While the preceding theorem answers this in the case of weak disorder, the

problem is open for strong disorder. Moreover, it has been speculated by Yuval Peres (private communication) that $\text{prob}_n(ds)$ will a.s. have infinitely many weak limit points under strong disorder. However, in the case of critical strong disorder we show that a natural infinite volume polymer exists and is related to the finite volume polymers through limits in probability.

2.3   Tree Polymers at Critical Strong Disorder

In this section we show the existence under critical strong disorder, i.e., assuming $\mathbb{E}X \ln X = \ln 2$, of an infinite volume polymer probability $\text{prob}_\infty(ds)$ that may be viewed as the weak limit in probability of the sequence $\text{prob}_n(ds), n \geq 1$, in the sense that its characteristic function is the limit in probability of the corresponding sequence of characteristic functions of $\text{prob}_n(ds), n \geq 1$.

For $v \in T$, $v = (v_1, \ldots, v_m)$, say, let

$$\Delta_m(v) = \{s \in \partial T : s_i = v_i, \ i = 1, \ldots, m\}, \qquad |v| = m.$$

Since T is countable there are countably many such finite-dimensional rectangles in ∂T.

For $m > n$, note that

$$\text{prob}_n(\Delta_m(v)) = \int_{\Delta_m(v)} \frac{d\text{prob}_n}{d\lambda}(s)\lambda(ds)$$

$$= \int_{\Delta_m(v)} Z_n^{-1} \prod_{j=1}^{n} X_{s|j}\lambda(ds)$$

$$= Z_n^{-1} \int_{\Delta_m(v)} \prod_{j=1}^{n} X_{v|j}\lambda(ds)$$

$$= Z_n^{-1} \prod_{j=1}^{n} X_{v|j} \cdot 2^{-m}.$$

For example,

$$\text{prob}_1(\Delta_m(v)) = Z_1^{-1} X_{v|1} 2^{-m}, \qquad Z_1 = \frac{X_+ + X_-}{2}$$

$$= \frac{X_{v|1} 2^{-(m-1)}}{X_+ + X_-}$$

$$= \begin{cases} \frac{X_+ 2^{-(m-1)}}{X_+ + X_-}, & v|1 = +1 \\ \frac{X_- 2^{-(m-1)}}{X_+ + X_-}, & v|1 = -1. \end{cases}$$

$\sum_{|v|=m} \text{prob}_1(\Delta_m(v)) = 1$ since there are 2^m such v's, half of which have $v_1 = +1$ and the other half have $v_1 = -1$.

For $m \leq n, |v| = m$, we have

$$\text{prob}_n(\Delta_m(v)) = Z_n^{-1} \int_{\Delta_m(v)} \prod_{j=1}^{n} X_{s|j}\lambda(ds)$$

$$= Z_n^{-1} \prod_{j=1}^{m} X_{v|j} \sum_{|t|=n-m} \prod_{j=1}^{n-m} X_{(v*t)|j} 2^{-n}$$

$$= Z_n^{-1} \left(\prod_{j=1}^{m} X_{v|j} 2^{-m} \right) Z_{n-m}(v),$$

where

$$Z_0(v) = 1, \quad Z_{n-m}(v) = \sum_{|t|=n-m} \prod_{j=1}^{n-m} X_{(v*t)|j} 2^{-(n-m)}.$$

In particular, $Z_n = Z_n(0)$, where $0 \in T$ is the root.

Note that

$$Z_n = \sum_{|u|=m} \sum_{|t|=n-m} \prod_{j=1}^{m} X_{u|j} 2^{-m} \prod_{j=1}^{n-m} X_{(u*t)|j} 2^{-(n-m)}$$

$$= \sum_{|u|=m} Z_{n-m}(u) \prod_{j=1}^{m} X_{u|j} 2^{-m}.$$

Thus, letting $a_k = 1/\sqrt{k}, k \geq 1$,

$$\text{prob}_n(\Delta_m(v)) = \frac{D_{n-m}(v) \prod_{j=1}^{m} X_{v|j} 2^{-m} \frac{Z_{n-m}(v)}{a_{n-m} D_{n-m}(v)}}{\sum_{|u|=m} D_{n-m}(u) \left(\prod_{j=1}^{m} X_{v|j} 2^{-m} \right) \frac{Z_{n-m}(u)}{a_{n-m} D_{n-m}(u)}}$$

$$\longrightarrow \frac{D_\infty(v) \prod_{j=1}^{m} X_{v|j} 2^{-m}}{\sum_{|u|=m} D_\infty(u) \left(\prod_{j=1}^{m} X_{v|j} 2^{-m} \right)}$$

where (i) the convergence to $D_\infty(v)$ is the almost sure limit of the *derivative martingale* obtained by Biggins and Kyprianou in [10], and (ii) $\lim_{n\to\infty} \frac{Z_{n-m}(v)}{a_{n-m} D_{n-m}(v)} = c > 0$ is the limit in probability at critical strong disorder recently obtained by Aidékon and Shi in [2]. The constant $c = (\frac{2}{\pi\sigma^2})^{1/2}$, for $\sigma^2 = \mathbb{E}\{X(\ln(X))^2\} - (\mathbb{E}\{X \ln(X)\})^2 > 0$, does not depend on $v \in T$. Aidékon and Shi [2] also point out that the almost sure positivity of $D_\infty(v)$ follows from [10] and [1]. The sequence $a_k = k^{-\frac{1}{2}}, k \geq 1$, is referred to as the Seneta-Heyde scaling.

Remark 2.3.1. *For each $v \in T$, there is a set $N(v)$ of probability zero such that*

$$D_\infty(v, \omega) = \lim_{n \to \infty} D_n(v, \omega), \quad \omega \in \Omega \backslash N(v).$$

Since T is countable, the set $N = \bigcup_{v \in T} N(v)$ is still a P-null subset of Ω. The almost sure convergence of the derivative martingales is essential to the construction of prob_∞ *given in the lemma below.*

We now define

$$\mathrm{prob}_\infty(\Delta_m(v), \omega) = \frac{D_\infty(v, \omega) \prod_{j=1}^m X_{v|j}(\omega) 2^{-m}}{\sum_{|u|=m} D_\infty(u, \omega) \left(\prod_{j=1}^m X_{u|j}(\omega) 2^{-m} \right)}$$

for $\omega \in \Omega \backslash N$.

Lemma 2.3.2. $\mathrm{prob}_\infty(\Delta_m(v), \omega)$ *extends to a unique probability on $(\partial T, \mathcal{B})$ for each $\omega \in \Omega \backslash N$.*

Proof. We use Caratheodory extension, taking careful advantage of the fact that the sets $\Delta(v)$, $v \in T$, are both open and closed subsets of the compact set ∂T. For $\omega \in \Omega \backslash N$, $\mathrm{prob}_\infty(\cdot, \omega)$ extends to the algebra generated by $\{\Delta(v) : v \in T\}$ by addition. Since ∂T is compact and the rectangles are both open and closed, countable additivity on this algebra must hold as a consequence of finite additivity; i.e. if $\bigcup_{i=1}^\infty \Delta(v_i)$ is contained in the algebra generated by $\{\Delta(v) : v \in T\}$, then $\bigcup_{i=1}^\infty \Delta(v_i)$ is closed, hence compact, and its own open cover, i.e. $\bigcup_{i=1}^\infty \Delta(v_i) = \bigcup_{i=1}^l \Delta(v_{i_l})$ for some finite subsequence $\{i_j\}_{j=1}^l$ of $\{1, 2, \ldots\}$. $\qquad \square$

Theorem 2.3.3. *At critical strong disorder, for each finite set $F \subseteq \mathbb{N}$*

$$\widehat{\mathrm{prob}_n}(F) \quad \Rightarrow \quad \widehat{\mathrm{prob}_\infty}(F) \qquad \text{in probability,}$$

where $\widehat{\mathrm{prob}}_n, n \geq 1, \widehat{\mathrm{prob}}_\infty$ *denote their respective Fourier transforms as probabilities on the compact abelian multiplicative group ∂T for the product topology.*

Proof. The continuous characters of the group ∂T are given by

$$\chi_F(t) = \prod_{j \in F} t_j \quad \text{for finite sets } F \subseteq \mathbb{N}.$$

In particular there are only countably many characters of ∂T. From standard Fourier analysis it follows that we need only show that $\lim_{n \to \infty} \mathbb{E}_{\mathrm{prob}_n} \chi_F = \mathbb{E}_{\mathrm{prob}_\infty} \chi_F$ in probability for each finite set $F \subseteq \mathbb{N}$. Let $m = \max\{k : k \in F\}$. Then for $n > m$,

$$
\begin{aligned}
\mathbb{E}_{\mathrm{prob}_n} \chi_F &= \int_{\partial T = \cup_{|v|=m} \Delta_m(v)} \chi_F(s) \frac{d\mathrm{prob}_n}{d\lambda}(s) \lambda(ds) \\
&= \sum_{|v|=m} \left(\prod_{j \in F} v_j \right) Z_n^{-1}(0) \prod_{j=1}^m X_{v|j} 2^{-m} \sum_{|t|=n-m} \prod_{j=1}^{n-m} X_{(v\star t)|j} 2^{-(n-m)} \\
&= \sum_{|v|=m} \left(\prod_{j \in F} v_j \right) \prod_{j=1}^m X_{v|j} 2^{-m} \frac{Z_{n-m}(v)}{Z_n(0)} \\
&= \sum_{|v|=m} \left(\prod_{j \in F} v_j \right) \prod_{j=1}^m X_{v|j} 2^{-m} \frac{D_{n-m}(v) \frac{Z_{n-m}(v)}{a_{n-m} D_{n-m}(v)}}{\sum_{|u|=m} \prod_{j=1}^m X_{u|j} 2^{-m} D_{n-m}(u) \frac{Z_{n-m}(u)}{a_{n-m} D_{n-m}(u)}} \\
&\longrightarrow \mathbb{E}_{\mathrm{prob}_\infty} \chi_F,
\end{aligned}
$$

where the convergence is almost sure for terms of the form D_{n-m} and in probability

for those of the form $Z_{n-m}/(a_{n-m}D_{n-m})$ as $n \to \infty$. $\qquad\qquad\square$

2.4 Diffusivity Problems at Strong Disorder

With regard to the aforementioned a.s. limits in distribution of polygonal tree paths, Waymire and Williams in [46] also obtained a.s. limits of the form

$$\lim_{n\to\infty} \frac{\ln E_{prob_n} e^{r(S)_n}}{n} = F(r)$$

under both weak and strong disorder. Let us refer to these as almost sure *Laplace rates* in reference to the Laplace principle of large deviation theory.

In the case of weak disorder the universal limit is $F(r) = \ln \cosh(r)$, in a neighborhood of the origin, otherwise independent of the distribution of X. In addition to being independent of the distribution of X within the range of weak disorder, this universality of Laplace rates is manifested in the coincidence with the same limit obtained for $X \equiv 1$, i.e., for simple symmetric random walk.

For an illustrative case of strong disorder, consider $X = e^{\beta Z - \frac{\beta^2}{2}}$, where Z is standard normal and $\beta \geq \beta_c = \sqrt{2 \ln 2}$. Then from [46], it follows that a.s. in a neighborhood of the origin that

$$F(r) = r \tanh(rh(r)) + \beta^2 h(r) - \beta\beta_c,$$

where $h(r)$ is the uniquely determined solution to

$$\beta^2 h^2(r) + 2rh(r)\tanh(rh(r)) - 2\ln\cosh(rh(r)) = \beta_c^2;$$

also see [46] (Sec 6, Cor 2, 2010) for the general formulae in the case of strong disorder. In particular, the universality of the Laplace rates breaks down, even at critical strong disorder. A graph of $F(r)$ computed from MATLAB is indicated in Figure 2.1 for the strong disorder case of $\beta = 2\beta_c$.
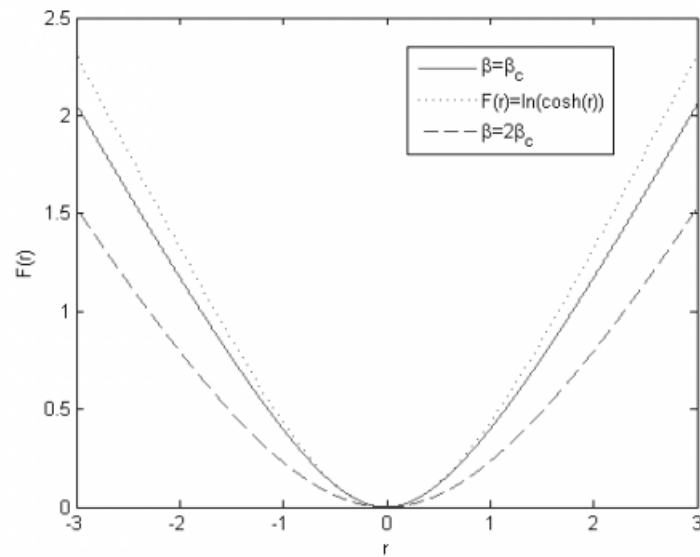
Figure 2.1: Graph of the function F for various β.

Using the equations defining $F(r)$ one may easily verify that $F(0) = 0, F'(0) = 0$ and $F''(0) = \frac{2\beta\beta_c - \beta_c^2}{\beta^2}$. While these specific calculations follow directly from the general results of [46], from here one is naturally lead to speculate that the asymptotic

variance under strong disorder is obtained under diffusive scaling by $\sqrt{n}$ precisely as

$$\sigma^2(\beta) = \frac{2\beta\beta_c - \beta_c^2}{\beta^2}, \quad \beta \geq \beta_c.$$

(To avoid potential confusion, let us mention that other forms of polymer scalings appear in the recent probability literature under which the polymer is referred to as "superdiffusive" even in the context of weak disorder; e.g., in reference to wandering exponents in [4].) In particular this formula continuously extends the weak disorder variance $\sigma^2(\beta) \equiv 1, \beta < \beta_c$, across $\beta = \beta_c$. In any case, this quantity is a basic parameter of the rigorously proven limit $F(r)$.

3 Simple Branching Random Walk and an Associated Tree Polymer

3.1 Introduction

In this chapter we present results for the simple branching random walk and an associated tree polymer model. The simple branching random walk is a discrete time, discrete space branching random walk on the space $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ of non-negative integers. A single particle is initially at zero, and at each time step a deterministic number of particles $b \in \{2, 3, \dots\}$ is generated per particle. b is fixed throughout the process. Each particle moves independently, either remaining at its parent's location with probability q, or moving right one unit with probability $p = 1 - q$.

Denote by $W_{n,k}$ the number of particles (walkers) at location k at time n. The evolution of the distribution (counts) of particles $\mathbf{W}_n := \{W_{n,k} : k = 0, 1, \dots, n\}, n = 0, 1, 2, \dots$, can be described by the following simple recursion. Suppose that for each $n \geq 0$, $\mathbf{Y}_{n+1} = \{Y_{n+1,k} : k = 1, 2, \dots\}$, and $\mathbf{Y}'_{n+1} = \{Y'_{n+1,k} : k = 1, 2, \dots\}$ denote mutually independent i.i.d. sequences of Bin(2,q), and Bin(2,p), respectively, distributed random variables and independent of $\mathbf{W}_n$. Then, letting $W_{0,0} = 1$, and

making the the convention that $W_{n,-1} := 0$ for any n, one has

$$W_{n+1,k} = \sum_{j=1}^{bW_{n,k}} Y_{n+1,j} + \sum_{j=1}^{bW_{n,k-1}} Y'_{n+1,j}. \tag{3.1.1}$$

Notice that in the case $p = q = 1/2$, letting $w_{n+1} = b^{-(n+1)}W_{n+1}$ the prediction $\hat{w}_{n+1,k} = \mathbb{E}(w_{n+1,k}|w_{n,k} : k = 0, 1, \ldots, n)$ satisfies

$$\hat{w}_{n+1,k} = w_{n,k} + w_{n,k-1}. \tag{3.1.2}$$

More generally, it also follows easily for any $n, b \geq 2, p \in [0, 1]$, that the expected value of the random probability distribution $Q_n := \{w_{n,k} \equiv \frac{W_{n,k}}{b^n} : k = 0, 1, \ldots n\}$, is a (deterministic) binomial distribution with parameters n, p, not depending on $b \geq 2$. Building on the familiar triangle recursion

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} \tag{3.1.3}$$

for the binomial coefficients, (3.1.2) provides an interesting version of a *noisy Pascal's triangle (SPT) property* for the (normalized) numbers of branching random walkers.

The motivation for this chapter is based on connections of the simple branching random walk with tree polymers and multiplicative cascades. Such connections are generally well-known, but this particular choice has virtues that may be useful in efforts to understand effects of disorder. To set this up will require a minimal amount of notation and a few basic facts and results about tree polymers.

For the context of tree polymers we fix $b = 2$. Denote a directed binary (genealogical) tree by $T = \bigcup_{n=0}^{\infty} \{-1, 1\}^n$. Define the *boundary* of T by $\partial T = \{-1, 1\}^{\mathbb{N}}$, with the product topology. An ∞-*tree path* is denoted by $s = (s_1, s_2, \ldots) \in \partial T$. We will also refer to *finite tree paths* $s = (s_1, \ldots, s_n) \in T \setminus \{0\}$ of length $|s| = n$, and for $s = (s_1, s_2, \ldots) \in \partial T$, we use the notation $s|n := (s_1, s_2, \ldots, s_n)$, read "s restricted to n", to truncate.

We will define a sequence of probability measures for ∞-tree paths in ∂T which are dependent on the realization of a random *environment*, defined on the vertices in T by a collection $\{X_v : v \in T\}$ of i.i.d., mean one, strictly positive random variables on a probability space $(\Omega, \mathcal{F}, P)$. Denote by X a generic random variable having the common distribution of each X_v; X is referred to as the *polymer weights*. The polymer probability measures are defined with respect to the infinite product measure $\lambda(ds) = \left(\frac{1}{2} \delta_+(ds) + \frac{1}{2} \delta_-(ds)\right)^{\mathbb{N}}$ on $(\partial T, \mathcal{B})$, by the sequence of Radon-Nikodym derivatives given by

$$\frac{d\mathrm{prob}_n}{d\lambda}(s) = Z_n^{-1} \prod_{j=1}^{n} X_{s|j},$$

where the *partition function* Z_n normalizes to a probability by

$$Z_n = \int_{\partial T} \prod_{j=1}^{n} X_{s|j} \lambda(ds) = \sum_{|s|=n} \prod_{j=1}^{n} X_{s|j} 2^{-n}.$$

The sequence of probability measures $\mathrm{prob}_n(ds)$ is called the *tree polymer* on $(\partial T, \mathcal{B})$.

The specific model to be analyzed in the present chapter is the tree polymer

defined by path weights distributed as

$$
X = \begin{cases}
e^{\beta} & \text{with probability } p_{\beta} = \frac{1}{1+e^{\beta}} \\
e^{-\beta} & \text{with probability } q_{\beta} = \frac{e^{\beta}}{1+e^{\beta}}
\end{cases} \tag{3.1.4}
$$

Here β is a fixed real number parameter but, since X is invariant under $\beta \to -\beta$, there is no loss of generality in restricting attention to $\beta \geq 0$. The partition function Z_n of the tree polymer with this choice of polymer weights is related to the simple branching random walk $\mathbf{W}_n = \{W_{n,k} : k = 0, 1, \ldots, n\}, n \geq 0$, with probabilities $p = p_{\beta} \leq 1/2$ and $q = q_{\beta} \geq 1/2$ by the following proposition whose verification is left to the reader.

Proposition 3.1.1. *For $\beta \in \mathbb{R}$ and the polymer weight* (3.1.4), *one has for any $n \geq 0$ that*

$$
prob_n(ds) = W_n^{-1} e^{2\beta H_n(s)} \lambda(ds),
$$

where $H_n(s) = \#\{i \in \{1, \ldots, n\} : X(s|i) = e^{\beta}\}$ is the position of the walker defined by genealogy s at time n, and

$$
W_n = \sum_{k=0}^{n} W_{n,k} e^{2\beta k} = 2^n e^{\beta n} Z_n.
$$

The notions of *weak disorder* and *strong disorder*, e.g., see Bolthausen [13], provide a well-known dichotomy between tree polymer environments. To describe them, first note that it is easily verified that the sequence of partition functions $Z_1, Z_2, \ldots$ is a positive martingale, so by martingale convergence the limiting ran-

dom variable $Z_\infty := \lim_{n\to\infty} Z_n$ exists a.s. in $(\Omega, \mathcal{F}, P)$. By positivity of the environment, $[Z_\infty = 0]$ is a tail event, so by Kolmogorov's zero-one law, $P(Z_\infty = 0)$ must equal zero or one. Kahane and Peyrière [26] for multiplicative cascades, and (independently) Biggins [11] for branching random walks, established the following dichotomy:

$$P(Z_\infty > 0) = 1 \quad \Longleftrightarrow \quad \mathbb{E}X \ln X < \ln 2$$

$$P(Z_\infty = 0) = 1 \quad \Longleftrightarrow \quad \mathbb{E}X \ln X \geq \ln 2.$$

In the first case where $[Z_\infty > 0]$ a.s., the environment is said to be in a state of *weak disorder*, whereas if $[Z_\infty = 0]$ a.s., the environment is in a state of *strong disorder*. Note that the deterministic environment $X \equiv 1$ a.s., i.e., $\beta = 0$, can be regarded informally as the "weakest" of the weak disorder regimes. The tree polymer paths in this case are just the simple symmetric random walk paths distributed as $prob_n(ds) \equiv \lambda(ds)$.

Remark 3.1.2. *As a matter of terminology, multiplicative cascades are said to survive or die out according to whether the martingale limit Z_∞ is positive or not. For branching random walks, the case of critical strong disorder $\mathbb{E}X \ln X = \ln b$ is referred to as the boundary case.*

In general one may view the selection of a (finite) path S of length n from $prob_n(ds)$ as follows: Let $\mathcal{P}_{n,j}$ denote the (random) collection of all branching random walk paths that end at j at time n, for $j = 0, 1, \ldots, n$. Select the position $H_n(S) = j$ with probability $W_{n,j} e^{2\beta j}/W_n$, $j = 0, 1, 2, \ldots, n$ and then, given $H_n(S) =$

j, pick a path $S = s$ from $\mathcal{P}_{n,j}$ uniformly with probability $1/W_{n,j}$.

The disorder of the tree polymer defined by (3.1.4) is easily computed as

$$\mathbb{E}X \ln X = \beta(1 - 2p_\beta) = \beta \tanh\left(\frac{\beta}{2}\right). \qquad (3.1.5)$$

In particular, there is a critical value $0 < \beta_c < \infty$ such that weak disorder occurs for $0 \leq \beta < \beta_c$ and strong disorder for $\beta \geq \beta_c \approx 1.25$.

While disorder type has obvious consequences for the asymptotic behavior of the partition functions Z_n, and polymer path distributions, see [2], [21], [46], [24], for example, one of the main results of this chapter show that it is of no consequence to an almost sure central limit theorem approximation to the (normalized) random probability distribution Q_n of walkers.

The closest work to that presented in this chapter that we know is that of Grill [19] establishing the almost sure connectivity of the support of the simple symmetric branching random walk, i.e., $p = 1/2$ with $b > 2$, (supercritical branching). In fact [19] permits a random offspring distribution, so long as the genealogy is supercritical as a Galton-Watson branching process. In the context of multiplicative cascades one may check that for $p = 1/2$ and $b > 2$, the positive martingales $M_n(\lambda) = \left(\frac{b}{2}\right)^{-n}(1 + e^\lambda)^{-n} \sum_{k=0}^{n} e^{\lambda k} W_{n,k}$ have the survival property that $M_n(\lambda) \to M_\infty(\lambda) > 0$ a.s. for any choice of $\lambda \in \mathbb{R}$, or 'weak disorder' property.

3.2 The Support of Z

In [19] the connectivity of the support of the simple (symmetric) branching random walk is established. In this section we provide an extension to asymmetric simple branching random walk. In particular we consider a (deterministic) branching number $b \geq 3$ and $p \neq 1/2$, as [19] covers the case $p = 1/2$. While this extends to supercritical Galton-Watson trees with offspring distribution having mean b, treatment of this generality only complicates the notation for the purposes here.

The technique is to exploit the recursive properties of the branching random walk in ways similar to that of [19], and others. Using the notation from [19], define $K_n := \inf\{k : W_{n,k} > 0\}$ and $L_n := \inf\{\ell : W_{n,n-\ell} > 0\}$. Monotonicity in n over the sets on which the infimum is taken implies that both limits

$$K = \lim_{n \to \infty} K_n \quad \text{and} \quad L = \lim_{n \to \infty} L_n$$

exist, but may be infinite. In [19] it is observed that by symmetry $(p = q = 1/2)$, K_n and L_n have the same distribution, and therefore both K and L are a.s. finite since extinction of the genealogical tree is impossible. In fact, it follows from standard results on the position of the rightmost particle, e.g., see [11], [27], [20] that $\lim_n L_n = \infty$ a.s. for $\beta \neq 0$, so that $L = \infty$ in the non-symmetric case. Nonetheless, as we will now show, one may still obtain the a.s. connectivity of the support as follows.

Theorem 3.2.1. *Suppose $p \in (0,1)$ and $\max\{p,q\} < \alpha_b := \frac{b-1}{b-1+b^{\frac{1}{b}}}$. Then almost*

surely, eventually,

$$\{k : W_{n,k} > 0\} = [K_n, n - L_n].$$

Remark 3.2.2. *Note the hypothesis* $\max\{p,q\} < \alpha_b$ *implies* $b \geq 3$ *since* $\alpha_2 = \sqrt{2} - 1 < \frac{1}{2}$. *For* $b \geq 3$, *one can easily verify that* $1 - \alpha_b > \frac{1}{b}$, *so* $\max\{p,q\} < \alpha_b$ *also implies* $\min\{bp, bq\} > 1$. *As in [19], both the number of particles at zero and the number of particles at the rightmost possible location* n *in generation* n *are supercritical Galton-Watson branching processes. In the context of multiplicative cascades (3.1.4) one has* $\mathbb{E}X \ln X = \beta e^\beta p_\beta - \beta e^{-\beta} q_\beta = \beta(q_\beta - p_\beta) < \beta < \ln b$, *since* $b/(1 + e^\beta) > 1$. *That is, the cascade survives, or in the language of polymers, the disorder is weak. The extent of connectivity with respect to the full range of weak disorder and strong disorder remains open.*

We begin with some lemmas.

Define, as in [19], $p_{n,k} = \mathbb{P}(W_{n,k} = 0)$, and $q_{n,k} = \mathbb{P}(W_{n,j} = 0 \text{ for all } j \leq k)$. Let $F(z)$ denote the generating function of the offspring distribution. By a well-known recursion, also used in [19], we have

$$p_{n+1,k} = F(qp_{n,k} + pp_{n,k-1}) \tag{3.2.1}$$

and the same recursion holds for the probabilities $\{q_{n,k}\}$.

Lemma 3.2.3 (Spatial Unimodality). *For any* $p \in (0,1)$ *and any integer* $n \geq 2$, *there exists an integer* $k_n, 0 < k_n < n$, *satisfying*

$$p_{n,0} \geq \ldots \geq p_{n,k_n} \leq \ldots \leq p_{n,n}. \tag{3.2.2}$$

The sequence $\{k_n\}$ is non-decreasing.

Proof. For any generating function $F(z)$, one easily computes $p_{2,0} = F(qF(p) + p)$, $p_{2,1} = F(qF(p) + pF(q))$, and $p_{2,2} = F(q + pF(q))$. Since $F(z) \in (0, 1)$ for $z \in (0, 1)$, we obtain $p_{2,0} \geq p_{2,1} \leq p_{2,2}$.

Now suppose (3.2.2) holds for an arbitrary integer $n \geq 2$ with associated location k_n. The recursion (3.2.1) immediately implies both

$$p_{n+1,0} \geq \ldots \geq p_{n+1,k_n} \qquad \text{and} \qquad p_{n+1,k_n+1} \leq \ldots \leq p_{n+1,n+1}.$$

If $p_{n+1,k_n} \geq p_{n+1,k_n+1}$, set $k_{n+1} = k_n + 1$. Otherwise set $k_{n+1} = k_n$. The result follows by induction. $\qquad\square$

Lemma 3.2.4 (Eventual Temporal Monotonicity). *Let $p \in (0, 1)$. For each integer k, there exists an integer $n_0(k)$ depending on k such that $p_{n,k}$ is non-decreasing in n for $n \geq n_0(k)$.*

Proof. Since $p_{n,-1} = 1$ for any $n \geq 0$, it follows immediately from the recursion (3.2.1) that $p_{0,0} \leq p_{1,0} \leq p_{2,0} \leq \ldots$. Moreover, if $n_0(k)$ is such that $p_{n_0(k),k} \leq p_{n_0(k)+1,k} \leq \ldots$ for a given k, and there exists some $n_0(k+1) \geq n_0(k)$ such that $p_{n_0(k+1),k+1} \leq p_{n_0(k+1)+1,k+1}$, then (3.2.1) also implies $p_{n_0(k+1),k+1} \leq p_{n_0(k+1)+1,k+1} \leq \ldots$. This will establish the result by induction.

We now show that the required time $n_0(k+1)$ exists. If not, then there is a k and a time $n_0(k)$ such that $p_{n_0(k),k+1} > p_{n_0(k)+1,k+1} > p_{n_0(k)+2,k+1} > \ldots$ and $p_{n_0(k),k} \leq p_{n_0(k)+1,k} \leq \ldots$. By monotonicity, $p_k := \lim_{n\to\infty} p_{n,k}$ exists, and therefore

$h(m) := p_{n_0(k)+m,k} - p_{n_0(k)+m-1,k} \downarrow 0$ as $m \to \infty$. Then for $m \geq 2$,

$$p_{n_0(k)+m,k+1} = F\big(qp_{n_0(k)+m-1,k+1} + pp_{n_0(k)+m-1,k}\big)$$

$$< F\big(qp_{n_0(k)+m-2,k+1} + pp_{n_0(k)+m-2,k} + ph(m-1)\big)$$

$$= p_{n_0(k)+m-1,k+1} + \big[F\big(qp_{n_0(k)+m-2,k+1} + pp_{n_0(k)+m-2,k} + ph(m-1)\big)$$

$$- F\big(qp_{n_0(k)+m-2,k+1} + pp_{n_0(k)+m-2,k}\big)\big]$$

$$:= p_{n_0(k)+m-1,k+1} + \epsilon_m.$$

Since F is strictly increasing and continuous, the numbers ϵ_m are positive and $\epsilon_m \downarrow 0$ as $m \to \infty$. Thus for sufficiently large m we have $p_{n_0(k)+m,k+1} \leq p_{n_0(k)+m-1,k+1}$, contrary to hypothesis. $\qquad \square$

The following result is not used in the proof of Theorem 3.2.1, but is recorded for possible independent value.

Proposition 3.2.5 (Infinite Monotone Sequences)**.** *Let $p \in (0,1)$ and $n \geq 2$. Let k_n be the integer obtained from Lemma 3.2.3. Then there exists a minimal nonnegative integer m such that $p_{n+m,k_n} \geq p_{n+m,k_n+1}$.*

Proof. If $k_{n+1} = k_n$, then $m = 0$.

By Lemma 3.2.4 there is an integer $\tilde{n} = n_0(k_n)$ such that $p_{\tilde{n},k_n} \leq p_{\tilde{n}+1,k_n} \leq p_{\tilde{n}+2,k_n} \leq \dots$. If there is a non-negative integer $n' \leq \tilde{n}$ such that $p_{n+n',k_n} \geq p_{n+n',k_n+1}$, then m is the smallest such integer.

Otherwise, we can find a positive integer m' such that

$$p_{n+n'+\tilde{m},k_n} < p_{n+n'+\tilde{m},k_n+1} \qquad \text{for } \tilde{m} = 1, \ldots, m'. \tag{3.2.3}$$

Applying the recursion (3.2.1) repeatedly, using (3.2.3), and applying properties of generating functions, we obtain

$$p_{n+n'+m',k_n+1} = F\left(q p_{n+n'+m'-1,k_n+1} + p p_{n+n'+m'-1,k_n}\right)$$

$$< F\left(p_{n+n'+m'-1,k_n+1}\right)$$

$$< \ldots$$

$$< F^{(m')}\left(p_{n+n',k_n+1}\right).$$

For sufficiently large m', one therefore obtains

$$p_{n+n'+m',k_n+1} - p_{n+n'+m',k_n} < F^{(m')}\left(p_{n+n',k_n+1}\right) - p_{n+n',k_n} \leq 0. \tag{3.2.4}$$

Thus the set of all possible integers m' satisfying (3.2.3) must be finite. Denoting the largest such integer by $\tilde{m}$, we obtain $m = \tilde{n} + \tilde{m}$. $\qquad\square$

Proof of Theorem 3.2.1. We first prove the following bound: Suppose $n \geq 2$ and $0 \leq k \leq k_n$, where $\{k_n\}$ are the locations obtained from Lemma 3.2.3. Then there is a constant $\rho < 1$ such that

$$p_{n,k} - q_{n,k} \leq \rho^n.$$

Define $d_{n,k} = p_{n,k} - q_{n,k}$. By the recursion (3.2.1) and the corresponding recursion

for $\{q_{n,k}\}$, we obtain

$$d_{n+1,k} = F(qp_{n,k} + pp_{n,k-1}) - F(qq_{n,k} + pq_{n,k-1}).$$

The mean value theorem produces a z for which

$$d_{n+1,k} = F'(z)(qd_{n,k} + pd_{n,k-1}), \quad qq_{n,k} + pq_{n,k-1} \le z \le qp_{n,k} + pp_{n,k-1}.$$

Using Lemmas 3.2.3 and 3.2.4, noting in particular that $n_0(0) = 0$ by the first part of the proof of Lemma 3.2.4, we have

$$z \le qp_{n,k} + pp_{n,k-1} \le qp_{n,k-1} + pp_{n,k-1}$$

$$= p_{n,k-1} \le p_{n,0} \le p_0 := \lim_{n \to \infty} p_{n,0}.$$

The process we consider here has generating function $F(z) = z^b$. By monotonicity of $F'(z)$ for positive z, we obtain

$$d_{n+1,k} \le F'(p_0)(qd_{n,k} + pd_{n,k-1}). \tag{3.2.5}$$

p_0 is the extinction probability for the Galton-Watson branching process $W_{n,0}$ that records the number of particles *remaining at zero*. The probability generating function for this process is

$$g(z) = (p + qz)^b.$$

By a well-known result, p_0 is the minimum solution of the fixed-point equation

$g(z) = z$. By taking the bth root and rearranging, we get the equation

$$qz - z^{\frac{1}{b}} + p = 0.$$

Let $y = z^{\frac{1}{b}}$ and write the equation in the form

$$qy^b - y + p = 0.$$

Factoring yields

$$(y - 1)(qy^{b-1} + qy^{b-2} + \ldots + qy - p) = 0.$$

Set the second factor equal to zero and write it in the form

$$\sum_{j=1}^{b-1} y^j = \frac{p}{q}. \tag{3.2.6}$$

A non-trivial extinction probability p_0 exists by supercriticality. (Note that $bq > 1$ by the remark after the theorem.) Since $0 < p_0 < 1$ and $y = p_0^{\frac{1}{b}}$ (for the root corresponding to p_0), we also have $0 < y < 1$. Together with (3.2.6), this gives the bound

$$(b - 1)y^{b-1} < \frac{p}{q},$$

which implies

$$p_0 < \left(\frac{p}{(b-1)q} \right)^{\frac{b}{b-1}}.$$

Using the fact that $z \leq p_0$ and $F'(z) = bz^{b-1}$, we obtain

$$F'(z) < b\left(\frac{p}{(b-1)q}\right)^b.$$

One can verify that the right-hand side evaluates to 1 when $p = \alpha_b$ and $q = 1 - \alpha_b$. Thus for $q > 1 - \alpha_b$, one has $p < \alpha_b$ and therefore $\gamma_1 := F'(z) < 1$.

We now claim that

$$d_{n,k} \leq \binom{n}{k} \gamma_1^{n-k} \tag{3.2.7}$$

for each fixed $k \leq n$ and all n. Note that $d_{0,k} \equiv 0$ for any k. By induction, we have

$$d_{n+1,k} \leq \gamma_1 \left(q d_{n,k} + p d_{n,k-1}\right)$$

$$\leq \gamma_1 \left(q \binom{n}{k} \gamma_1^{n-k} + p \binom{n}{k-1} \gamma_1^{n-k+1} \right)$$

$$\leq \gamma_1 \left(\binom{n}{k} \gamma_1^{n-k} + \binom{n}{k-1} \gamma_1^{n-k} \right)$$

$$= \gamma_1^{n+1-k} \binom{n+1}{k}.$$

We used the fact that $0 < \gamma_1, p, q < 1$ to drop those factors in the third line; a standard identity for binomial coefficients gives the fourth.

The theorem can now be established essentially using arguments in [19], which we present here (with appropriate modifications) for completeness.

By (3.2.7), we can find $\alpha > 0$ and $\gamma_2 < 1$ such that for $k < \alpha n$,

$$d_{n,k} \leq \gamma_2^n.$$

For $\alpha n \leq k \leq k_n$, iterate to obtain

$$p_{n,k} = \left(q p_{n-1,k} + p p_{n-1,k-1}\right)^b \leq \left(p_{n-1,k-1}\right)^b \leq \left(p_{n-k,0}\right)^{bk} \leq p_0^{bk} \leq p_0^{b\alpha n}.$$

Thus for $\alpha n \leq k \leq k_n$ we have

$$d_{n,k} \leq p_{n,k} \leq \gamma_3^n,$$

where $\gamma_3 := p_0^{b\alpha} < 1$. Hence we may take $\rho = \max\{\gamma_2, \gamma_3\}$ to conclude $d_{n,k} = p_{n,k} - q_{n,k} \leq \rho^n$ for $n \geq 2$ and $0 \leq k \leq k_n$.

We now use this bound to complete the proof. Let A_n denote the (possibly null) event that $W_{n,k} = 0$ for some k satisfying $K_n \leq k \leq k_n$. In probability, A_n is equal to the event B_n that $W_{n,k} = 0$ and $W_{n,k-1} > 0$ for some k satisfying $0 \leq k \leq k_n$. Thus

$$\mathbb{P}(A_n) = \mathbb{P}(B_n) \leq \sum_{0 \leq k \leq k_n} (p_{n,k} - q_{n,k}) \leq n\rho^n.$$

This implies that $\sum_n \mathbb{P}(A_n) < \infty$, so by Borel-Cantelli I $\mathbb{P}(A_n \text{ i.o.}) = 0$.

We have shown that $[K_n, k_n]$ is a.s eventually occupied. By the remark after the statement of the theorem, $bp > 1$. Considering the process in which the roles of p and q are reversed, the corresponding process recording the number of particles

remaining at zero is also supercritical. The left-most occupied position of this process is L_n, and $n - L_n$ is the right-most occupied position in the original process. Moreover, the two-sided monotone property in Lemma 3.2.3 also holds for this process: the two sides reverse roles, and the central locations $\{k_n\}$ are maintained. The proof that $[K_n, k_n]$ is a.s. eventually occupied can therefore be repeated (essentially) to conclude that $[k_n, n - L_n]$ is also a.s. eventually occupied. $\square$

3.3   A Central Limit Theorem

The goal of this section is to prove a CLT for the random distributions Q_n. To this end we introduce the following notation for the moment-generating functions:

$$\hat{Q}_n(t) := \sum_{k=0}^{n} e^{kt} \frac{W_{n,k}}{b^n}, \quad t \in \mathbb{R}.$$

Theorem 3.3.1. *For any $p \in (0,1)$, there is a t-neighborhood of the origin such that, P-almost surely,*

$$e^{-npt} \hat{Q}_n \left(\frac{t}{\sqrt{npq}} \right) \longrightarrow e^{\frac{t^2}{2}}.$$

That is, the distributions Q_n, centered and scaled, converge weakly to to a standard normal distribution with P-probability one.

Remark 3.3.2. *The moment-generating function of the binomial distribution with parameters n and p is $m_n(t) = \sum_{k=0}^{n} e^{kt} \binom{n}{k} p^k q^{n-k}$. The usual central limit theorem*

may be expressed accordingly as

$$e^{-npt} m_n \left(\frac{t}{\sqrt{npq}} \right) \longrightarrow e^{\frac{t^2}{2}}$$

and clearly holds for the expected polymer (i.e., annealed model).

We now define

$$M_n(t) := \frac{\hat{Q}_n(t)}{(q + pe^t)^n}.$$

It will be useful to write $\hat{Q}_n$ in the form

$$\hat{Q}_n(t) = M_n(t)(q + pe^t)^n.$$

The following proposition is easily checked and left to the reader.

Proposition 3.3.3. *For any $t \in \mathbb{R}$,*

1. *$M_n(t)$ is a mean-one non-negative martingale with respect to $\{\mathcal{F}_n\}$, and*

2. *$M_n'(t)$ is a mean-zero martingale with respect to $\{\mathcal{F}_n\}$.*

Since $M_n(t)$ is non-negative for each t, it converges a.s. and for each t to a non-negative random variable, say $M_\infty(t)$. If $M_n(t)$ is also uniformly integrable, then by a standard result (see for example Theorem 3.5 of [12]) we conclude that $M_\infty(t)$ is integrable and in fact

$$1 \equiv \mathbb{E} M_n(t) \longrightarrow \mathbb{E} M_\infty(t),$$

so that $\mathbb{E}M_\infty(t) = 1$.

We now prove the following sufficient condition for uniform integrability of $M_n(t)$.

Theorem 3.3.4. *For any $0 \le p \le 1$, $\sup_n \|M_n(t)\|_{L^2(\Omega)} < \infty$ for all t in a neighborhood of the origin.*

Proving the theorem requires analysis of the cross-moments $\mathbb{E}[W_{n,k}W_{n,j}b^{-2n}]$, which we carry out in the following lemmas. For convenience we introduce the following notations:

1. We define
$$\mu_{kj}^{(n)} := \mathbb{E}[W_{n,k}W_{n,j}b^{-2n}].$$

   Note that we clearly have symmetry in $\mu_{kj}^{(n)}$: $\mu_{k,j}^{(n)} = \mu_{j,k}^{(n)}$.

2. If g is any function on $\mathbb{N}_0 \times \mathbb{N}_0$, then let

$$\mathbb{B}g(k,j) := \mathbb{E}g(\mathcal{K}_k, \mathcal{J}_j) \equiv q^2 g(k,j) + qpg(k-1,j) + pqg(k,j-1) + p^2 g(k-1,j-1)$$

   denote the induced "branched average" of g with respect to the indicated distribution of the pair of random variables $(\mathcal{K}_k, \mathcal{J}_j), k, j \ge 0$.

3. We use the notation $b(k; n, p) := \binom{n}{k}p^k q^{n-k}$ for binomial probabilities and we define
$$h_n(k,j) := b(k; n, p)b(j; n, p). \tag{3.3.1}$$

   It is convenient to define $b(k; n, p) = 0$ for $k < 0$ and $k > n$.

We begin by developing a recursive formula for the cross-moments $\mu_{kj}^{(n)}$ via

$$\mu_{kj}^{(n)} = \mathbb{E}\big[\mathbb{E}(W_{n,k}W_{n,j}b^{-2n}|\mathcal{F}_{n-1})\big].$$

Considering the cases $k = j$, $|k - j| = 1$, and $|k - j| \geq 2$ separately leads to the following lemma.

Lemma 3.3.5. *Let μ_n denote cross moments $\mu_{kj}^{(n)}$ as a function of k, j. For all $n \geq 1$ and all $k, j \in \mathbb{N}_0$,*

$$\mu_{kj}^{(n)} = \mathbb{B}\mu_{n-1}(k, j) + f_n(k, j)$$

where

$$f_n(k, j) := \begin{cases} b^{-n}pq(b(k; n-1, p) + b(k-1; n-1, p)) & \text{if } k = j \\ -b^{-n}pqb(k \wedge j; n-1, p) & \text{if } |k - j| = 1 \\ 0 & \text{if } |k - j| \geq 2. \end{cases}$$

Proof. Case 1. Assume $|k - j| \geq 2$. In this case $W_{n,k}$ and $W_{n,j}$ are conditionally independent given $\mathcal{F}_{n-1}$, since particle displacements per time step are at most one

unit. Thus

$$
\begin{aligned}
\mathbb{E}[W_{n,k}W_{n,j}b^{-2n}|\mathcal{F}_{n-1}] &= b^{-2n}\mathbb{E}[W_{n,k}|\mathcal{F}_{n-1}]\mathbb{E}[W_{n,k}|\mathcal{F}_{n-1}] \\
&= b^{-2n}(bqW_{n-1,k}+bpW_{n-1,k-1})(bqW_{n-1,j}+bpW_{n-1,j-1}) \\
&= b^{-2(n-1)}(q^2 W_{n-1,k}W_{n-1,j}+qpW_{n-1,k}W_{n-1,j-} \\
&\quad +pqW_{n-1,k-1}W_{n-1,j}+p^2 W_{n-1,k-1}W_{n-1,j-1}).
\end{aligned}
$$

Then

$$
\mu_{kj}^{(n)} = \mathbb{E}[\mathbb{E}[W_{n,k}W_{n,j}b^{-2n}|\mathcal{F}_{n-1}]] = \mathbb{B}\mu_{n-1}(k,j).
$$

Case 2. Let $|k-j| = 1$, $k \geq 1$, and suppose without loss of generality that $j = k-1$. Let $Y_1, Y_2, \ldots$ be i.i.d. $b(1,q)$ random variables and let $Y_1', Y_2', \ldots$ be i.i.d. $b(1,p)$ random variables that are independent of the variables $\{Y_i\}$. Then $W_{n,k-1}$ can be expressed as

$$
W_{n,k-1} = \sum_{i=1}^{bW_{n-1,k-1}} Y_i + \sum_{r=1}^{bW_{n-1,k-2}} Y_r'.
$$

Unlike the previous case, $W_{n,k}$ and $W_{n,k-1}$ are not conditionally independent. However, the number of particles $\zeta(n,k)$ displaced rightward to k in the time step $n-1$ to n is related to the number remaining at $k-1$ in the same time step by the relation

$$
\zeta(n,k) = bW_{n-1,k-1} - \sum_{i=1}^{bW_{n-1,k-1}} Y_i.
$$

Suppose Y_s'' are i.i.d. $b(1,q)$ random variables that are independent of $\{Y_i\}$ and

$\{Y'_r\}$. Then we can express

$$W_{n,k} = \sum_{s=1}^{bW_{n-1,k}} Y''_s + \zeta(n,k).$$

The product $W_{n,k}W_{n,k-1}$ can then be expressed as

$$
\begin{aligned}
W_{n,k}W_{n,k-1} &= \sum_{i=1}^{bW_{n-1,k-1}}\sum_{s=1}^{bW_{n-1,k}} Y_i Y''_s + bW_{n-1,k-1}\sum_{i=1}^{bW_{n-1,k-1}} Y_i \\
&\quad - \sum_{i=1}^{bW_{n-1,k-1}} Y_i^2 - 2\sum\sum_{i_1<i_2} Y_{i_1}Y_{i_2} + \sum_{r=1}^{bW_{n-1,k-2}}\sum_{s=1}^{bW_{n-1,k}} Y'_r Y''_s \\
&\quad + bW_{n-1,k-1}\sum_{r=1}^{bW_{n-1,k-2}} Y'_r - \sum_{r=1}^{bW_{n-1,k-2}}\sum_{i=1}^{bW_{n-1,k-1}} Y'_r Y_i.
\end{aligned}
$$

Taking conditional expectation with respect to $\mathcal{F}_{n-1}$ gives

$$
\begin{aligned}
\mathbb{E}[W_{n,k}W_{n,k-1}|\mathcal{F}_{n-1}] &= b^2q^2 W_{n-1,k-1}W_{n-1,k} + b^2q Z_{n-1,k-1}^2 - bqW_{n-1,k-1} \\
&\quad -q^2(b^2 Z_{n-1,k-1}^2 - bW_{n-1,k-1}) + b^2pq W_{n-1,k-2}W_{n-1,k} \\
&\quad +b^2p W_{n-1,k-1}W_{n-1,k-2} - b^2pq W_{n-1,k-2}W_{n-1,k-1} \\
&= b^2q^2 W_{n-1,k}W_{n-1,k-1} + b^2pq Z_{n-1,k-1}^2 + b^2pq W_{n-1,k}W_{n-1,k-2} \\
&\quad b^2p^2 W_{n-1,k-1}W_{n-1,k-2} - bpq W_{n-1,k-1}.
\end{aligned}
$$

Finally, multiply by b^{-2n} and take expectations to get

$$\mu_{k,k-1}^{(n)} = \mathbb{B}\mu_{n-1}(k,k-1) - b^{-n}pq\,b(k-1;n-1,p).$$

Case 3. Suppose $k = j \geq 1$. Let the random variables $\{Y_i\}$ and $\{Y'_r\}$ be defined

as in case 2. Write $W_{n,k}$ as

$$W_{n,k} = \sum_{i=1}^{bW_{n-1,k}} Y_i + \sum_{r=1}^{bW_{n-1,k-1}} Y'_r.$$

Squaring this expression gives

$$
\begin{aligned}
W_{n,k}^2 &= \sum_{i=1}^{bW_{n-1,k}} Y_i^2 + 2\sum\sum_{i_1<i_2} Y_{i_1} Y_{i_2} + 2\sum_{i=1}^{bW_{n-1,k}}\sum_{r=1}^{bW_{n-1,k-1}} Y_i Y'_r \\
&+ \sum_{r=1}^{bW_{n-1,k-1}} (Y'_r)^2 + 2\sum\sum_{r_1<r_2} Y'_{r_1} Y'_{r_2}.
\end{aligned}
$$

Taking conditional expectations with respect to $\mathcal{F}_{n-1}$ gives

$$
\begin{aligned}
\mathbb{E}[W_{n,k}^2|\mathcal{F}_{n-1}] &= bqW_{n-1,k} + (b^2 W_{n-1,k}^2 - bW_{n-1,k})q^2 + 2b^2 qp W_{n-1,k} Z_{n-1,k-1} \\
&+ bpW_{n-1,k-1} + (b^2 W_{n-1,k-1}^2 - bW_{n-1,k-1})p^2.
\end{aligned}
$$

Multiply by b^{-2n}, take expectations and apply the symmetry $\mu_{k,k-1}^{(n)} = \mu_{k-1,k}^{(n)}$ to get

$$\mu_{kk}^{(n)} = \mathbb{B}\mu_{n-1}(k,k) + b^{-n}(q-q^2)b(k;n-1,p) + b^{-n}(p-p^2)b(k-1;n-1,p).$$

Since $q - q^2 = q(1-q) = qp$ and, similarly, $p - p^2 = pq$, we have

$$\mu_{kk}^{(n)} = \mathbb{B}\mu_{n-1}(k,k) + b^{-n}pq(b(k;n-1,p) + b(k-1;n-1,p)).$$

This is the formula claimed for $k \geq 1$. When $k = 0$, $b(k-1;n-1,p) = 0$, so the formula also holds for $k = 0$. $\qquad\square$

The recursions established in Lemma 3.3.5 lead to a general formula for the cross-moments $\mu_{kj}^{(n)}$ as follows.

Lemma 3.3.6. *For any $n \geq 0$ and any $k, j \in \mathbb{N}_0$, we have*

$$\mu_{kj}^{(n)} = h_n(k, j) + \gamma_{kj}^{(n)},$$

where

$$\begin{aligned}
\gamma_{kj}^{(n)} &= Cov(W_{n,k}b^{-n}, W_{n,j}b^{-n}) \\
&= \sum_{m=0}^{n-1} \sum_{l_1=0}^{m} \sum_{l_2=0}^{m} h_m(l_1, l_2) f_{n-m}(k - l_1, j - l_2) \\
&= \sum_{m=0}^{n-1} f_{n-m} * h_m(k, j),
\end{aligned}$$

for the functions f_n defined in Lemma 3.3.5, h_n at (3.3.1), and provided the middle sum is defined to be zero when $n = 0$.

Proof of Lemma 3.7. The proof is by induction on n, making repeated use of the Pascal triangle recursion (3.1.3) for binomial coefficients. The base cases are trivial, since for $n = 0$ we have, by definition, $\mu_{kj}^{(0)} = 1$ if $k = j = 0$ and $\mu_{kj}^{(0)} = 0$ otherwise, and these also hold for the claimed formula.

Suppose $n \geq 1$ and assume the formula holds for $n - 1$ and all $k, j \in \mathbb{N}_0$. Let γ denote the function of n, k, j with values $\gamma_{kj}^{(n)}$. By Lemma 3.3.5,

$$\begin{aligned}
\mu_{kj}^{(n)} &= \mathbb{B}\mu_{n-1}(k, j) + f_n(k, j) \\
&= \mathbb{B}h_{n-1}(k, j) + \mathbb{B}\gamma_{n-1}(k, j) + f_n(k, j).
\end{aligned}$$

One may check that the functions h_n have the following property with respect to branched averaging:

$$h_n(k,j) = \mathbb{B}h_{n-1}(k,j) \tag{3.3.2}$$

This accounts for the first term in the claimed formula. Using this and some reindexing in l_1 and l_2, one has

$$
\begin{aligned}
\mathbb{B}\gamma_{n-1}(k,j) &= \sum_{m=0}^{n-2}\sum_{l_1=-1}^{m}\sum_{l_2=-1}^{m} q^2 h_m(l_1+1, l_2+1) f_{n-1-(m+1)}(k-(l_1+1), j-(l_2+1)) \\
&\quad + \sum_{m=0}^{n-2}\sum_{l_1=-1}^{m}\sum_{l_2=-1}^{m} qp h_m(l_1, l_2+1) f_{n-1-(m+1)}(k-(l_1+1), j-(l_2+1)) \\
&\quad + \sum_{m=0}^{n-2}\sum_{l_1=-1}^{m}\sum_{l_2=-1}^{m} pq h_m(l_1+1, l_2) f_{n-1-(m+1)}(k-(l_1+1), j-(l_2+1)) \\
&\quad + \sum_{m=0}^{n-2}\sum_{l_1=-1}^{m}\sum_{l_2=-1}^{m} p^2 h_m(l_1, l_2) f_{n-1-(m+1)}(k-(l_1+1), j-(l_2+1)) \\
&= \sum_{m=0}^{n-2}\sum_{l_1=-1}^{m}\sum_{l_2=-1}^{m} \mathbb{B}h_m(l_1+1, l_2+1) f_{n-1-(m+1)}(k-(l_1+1), j-(l_2+1)) \\
&= \sum_{m=0}^{n-2}\sum_{l_1=-1}^{m}\sum_{l_2=-1}^{m} h_{m+1}(l_1+1, l_2+1) f_{n-1-(m+1)}(k-(l_1+1), j-(l_2+1)) \\
&= \gamma_{kj}^{(n)} - f_n(k,j).
\end{aligned}
$$

The asserted formula now follows immediately. $\qquad\square$

The convolution representation of the covariances will be particularly useful for the proof of Theorem 3.3.4, when expressed in terms of bivariate-moment generating functions, denoted by $\hat{h}_m(s,t)$ and $\hat{f}_{n-m}(s,t)$, using the following lemma.

Lemma 3.3.7.

$$\hat{h}_m(s,t) = (q + pe^s)^m (q + pe^t)^m$$

$$\hat{f}_{n-m}(s,t) = pq(1 - e^s - e^t + e^{s+t})b^{-(n-m)}(q + pe^{s+t})^{n-m-1}$$

$$+q(e^{-(s+t)} - 1)b^{-(n-m)}(pe^{2t})^{n-m}.$$

Proof. For $\hat{h}_m(s,t) = \sum_{k,j=0}^m e^{ks}e^{jt}h_m(k,j)$ we have

$$\sum_{k,j=0}^m e^{ks}e^{jt}h_m(k,j) = \left(\sum_{k=0}^m e^{ks}b(k;m)\right)\left(\sum_{k=0}^m e^{kt}b(k;m)\right) = (q + pe^s)^m (q + pe^t)^m.$$

For the computation of $\hat{f}_{n-m}(s,t)$ we we take advantage of the fact that it is supported on the set

$$L_{n-m} := \{(k,k),(k+1,k),(k,k+1) : k = 0,1,\ldots,n-m-1\}\bigcup\{(n-m,n-m)\}.$$

Then, with a little algebra, one has

$$\hat{f}_{n-m}(s,t) = \sum_{k,j=0}^{n-m-1} e^{ks}e^{jt}f_{n-m}(k,j) + e^{(n-m)(s+t)}f_{n-m}(n-m,n-m)$$

$$= \sum_{k=0}^{n-m} e^{k(s+t)}f_{n-m}(k,k) + \sum_{k=0}^{n-m-1} e^t e^{k(s+t)}f_{n-m}(k,k+1)$$

$$+ \sum_{k=0}^{n-m-1} e^s e^{k(s+t)}f_{n-m}(k+1,k) + e^{(n-m)(s+t)}f_{n-m}(n-m,n-m)$$

$$= pq(1 - e^s - e^t + e^{s+t})b^{-(n-m)}(q + pe^{s+t})^{n-m-1}$$

$$+q(e^{-(s+t)} - 1)b^{-(n-m)}(pe^{2t})^{n-m}.$$

$\square$

Proof of Theorem 3.3.4. Reverse the order of summation and apply properties of convolution to obtain

$$
\begin{aligned}
\|M_n(t)\|^2_{L_2(\Omega)} &= \frac{\mathbb{E}\hat{Q}_n^2(t)}{(q+pe^t)^{2n}} \\
&= \frac{\sum_{k,j=0}^n e^{(k+j)t} h_n(k,j)}{(q+pe^t)^{2n}} + \frac{\sum_{k,j=0}^n e^{(k+j)t}\gamma_{kj}^{(n)}}{(q+pe^t)^{2n}} \\
&= \frac{\left(\sum_{k=0}^n e^{kt}b(k;n)\right)^2}{(q+pe^t)^{2n}} + \frac{\sum_{k,j=0}^n \sum_{m=0}^{n-1} e^{(k+j)t} f_{n-m} * h_m(k,j)}{(q+pe^t)^{2n}} \\
&= 1 + \frac{\sum_{m=0}^{n-1} \widehat{f_{n-m} * h_m}(t,t)}{(q+pe^t)^{2n}} \\
&= 1 + \frac{\sum_{m=0}^{n-1} \widehat{f_{n-m}}(t,t)\widehat{h_m}(t,t)}{(q+pe^t)^{2n}}.
\end{aligned}
$$

From Lemma 3.3.7 we have

$$
\begin{aligned}
\hat{h}_m(t,t) &= (q+pe^t)^{2m} \\
\hat{f}_{n-m}(t,t) &= pq(1-e^t)^2 b^{-(n-m)}(q+pe^{2t})^{n-m-1} + q(e^{-2t}-1)b^{-(n-m)}(pe^{2t})^{n-m}.
\end{aligned}
$$

Substituting these transforms into the previous expression gives

$$
\begin{aligned}
\|M_n(t)\|^2_{L_2(\Omega)} &= 1 + \frac{pq(1-e^t)^2}{q+pe^{2t}} \sum_{m=0}^{n-1} \left(\frac{q+pe^{2t}}{b(q+pe^t)^2}\right)^{n-m} \\
&\quad + q(e^{-2t}-1)\sum_{m=0}^{n-1}\left(\frac{pe^{2t}}{b(q+pe^t)^2}\right)^{n-m} \\
&< 1 + C'\sum_{m=1}^\infty r^m + C''\sum_{m=1}^\infty \left(\frac{pe^{2t}}{b(q+pe^t)^2}\right)^m,
\end{aligned}
$$

where $r := \frac{q+pe^{2t}}{b(q+pe^t)^2}$, $C' := \sup_t \left\{ \frac{pq(1-e^t)^2}{q+pe^{2t}} \right\}$, and $C'' := \sup_t \left\{ q(e^{-2t}-1) \right\}$. C' and C'' are both finite for t in a bounded interval. Since $\frac{pe^{2t}}{b(q+pe^t)^2} < r$, we will have $\|M_n(t)\|^2_{L_2(\Omega)} < \infty$ if $r < 1$ for t restricted to a bounded interval. If $p = 1$ then $r = \frac{1}{b} < 1$, and we may restrict t to any bounded neighborhood of the origin. Suppose $p < 1$. Let c be a positive integer and consider t in the interval $[-2^{-c}, 2^{-c}]$. Then

$$r < \frac{q + pe^{-(c-1)}}{b(q + pe^{-2^{-c}})^2} < \frac{1 - p(1 - e^{-2^{-(c-1)}})}{be^{-2^{-(c-1)}}} < \frac{e^{2^{-(c-1)}} + p}{b}.$$

The second inequality holds since $(q + pe^{-2^{-c}})^2 > ((q+p)e^{-2^{-c}})^2 = e^{-2^{-(c-1)}}$. Since $p < 1$ and $b \geq 2$, we may choose c sufficiently large so that $r < 1$. Thus

$$\|M_n(t)\|^2_{L_2(\Omega)} < 1 + (C' + C'') \sum_{m=1}^{\infty} r^m < \infty.$$

The middle quantity is independent of n, so we obtain $\sup_n \|M_n(t)\|_{L_2(\Omega)} < \infty$. $\square$

We now establish a similar bound for the derivative martingale.

Theorem 3.3.8. *For any $0 \leq p \leq 1$, $\sup_n \|M_n'(t)\|_{L_2(\Omega)} < \infty$ for all t in a neighborhood of the origin.*

Proof. Write the derivative martingale in the form

$$M_n'(t) = \frac{\hat{Q}_n'(t) - n\alpha \hat{Q}_n(t)}{(q + pe^t)^n},$$

where $\alpha := \frac{pe^t}{q+pe^t}$. Then

$$\mathbb{E}(M_n'(t))^2 = \frac{\mathbb{E}(\hat{Q}_n'(t))^2 - 2n\alpha\mathbb{E}\hat{Q}_n'(t)\hat{Q}_n(t) + n^2\alpha^2\mathbb{E}(\hat{Q}_n(t))^2}{(q+pe^t)^{2n}}.$$

To compute this expectation we use the bivariate moment-generating function

$$\widehat{f_{n-m} * h_m}(s,t) = \sum_{k,j=0}^{n} e^{ks}e^{jt}f_{n-m} * h_m(k,j) = \hat{f}_{n-m}(s,t)\hat{h}_m(s,t).$$

By straightforward computations similar to those used in the proof of the previous theorem, we have the following:

$$\mathbb{E}(\hat{Q}_n'(t))^2 = n^2\alpha^2(q+pe^t)^{2n} + \sum_{m=0}^{n-1} \frac{\partial^2}{\partial s\partial t}\widehat{f_{n-m} * h_m}(t,t)$$

$$\mathbb{E}\hat{Q}_n'(t)\hat{Q}_n(t) = n\alpha(q+pe^t)^{2n} + \sum_{m=0}^{n-1} \frac{\partial}{\partial s}\widehat{f_{n-m} * h_m}(t,t)$$

$$\mathbb{E}(\hat{Q}_n(t))^2 = (q+pe^t)^{2n} + \sum_{m=0}^{n-1} \widehat{f_{n-m} * h_m}(t,t)$$

Now, since

$$n^2\alpha^2(q+pe^t)^{2n} - 2n\alpha\left(n\alpha(q+pe^t)^{2n}\right) + n^2\alpha^2(q+pe^t)^{2n} = 0,$$

we can express $\mathbb{E}(M_n'(t))^2$ as

$$\frac{\sum_{m=0}^{n-1}\left(\frac{\partial^2}{\partial s\partial t}\widehat{f_{n-m} * h_m}(t,t) - 2n\alpha\frac{\partial}{\partial s}\widehat{f_{n-m} * h_m}(t,t) + n^2\alpha^2\widehat{f_{n-m} * h_m}(t,t)\right)}{(q+pe^t)^{2n}}.$$

Applying properties of convolution and differentiating,

$$\frac{\partial^2}{\partial s \partial t} \widehat{f_{n-m} * h_m} - 2n\alpha \frac{\partial}{\partial s} \widehat{f_{n-m} * h_m} + n^2\alpha^2 \widehat{f_{n-m} * h_m}$$

$$= \frac{\partial^2 \hat{f}_{n-m}}{\partial s \partial t} \hat{h}_m + \frac{\partial \hat{f}_{n-m}}{\partial s} \frac{\partial \hat{h}_m}{\partial t} + \frac{\partial \hat{f}_{n-m}}{\partial t} \frac{\partial \hat{h}_m}{\partial s} + \hat{f}_{n-m} \frac{\partial^2 \hat{h}_m}{\partial s \partial t}$$

$$- 2n\alpha \frac{\partial \hat{f}_{n-m}}{\partial s} \hat{h}_m - 2n\alpha \hat{f}_{n-m} \frac{\partial \hat{h}_m}{\partial s} + n^2\alpha^2 \hat{f}_{n-m} \hat{h}_m.$$

Let $\alpha_2 = \frac{pe^{2t}}{q+pe^{2t}}$. Computing derivatives and evaluating at $(s,t) = (t,t)$, we obtain

$$\frac{\partial^2}{\partial s \partial t} \widehat{f_{n-m} * h_m}(t,t) - 2n\alpha \frac{\partial}{\partial s} \widehat{f_{n-m} * h_m}(t,t) + n^2\alpha^2 \widehat{f_{n-m} * h_m}(t,t)$$

$$= b^{-(n-m)} pq (q+pe^{2t})^{n-m-1} (q+pe^t)^{2m} [(1+e^{2t}) + (2 - 4e^t + 3e^{2t})(n-m-1)\alpha_2$$

$$+ (1 - 2e^t + e^{2t})(n-m-1)(n-m-2)\alpha_2^2$$

$$- 2(n-m)\alpha(1 - e^t + e^{2t}) - 2(n-m)\alpha(1 - 2e^t + e^{2t})(n-m-1)\alpha_2$$

$$+ m^2\alpha^2 (1 - 2e^t + e^{2t}) - 2nm\alpha^2 (1 - 2e^t + e^{2t}) + n^2\alpha^2 (1 - 2e^t + e^{2t})]$$

$$+ q(e^{-2t} - 1) b^{-(n-m)} (pe^{2t})^{n-m} (q+pe^t)^{2m} [(n-m)^2 e^{-2t}$$

$$+ 2m\alpha(n-m)e^{-t} + m^2\alpha^2 - 2n\alpha(n-m)e^{-t} - 2nm\alpha^2 + n^2\alpha^2]$$

$$\leq C_1 (n-m)^2 b^{-(n-m)} (q+pe^{2t})^{n-m} (q+pe^t)^{2m}$$

$$+ C_2 (n-m)^2 b^{-(n-m)} (pe^{2t})^{n-m} (q+pe^t)^{2m},$$

where $C_1 := \frac{pq((1+2e^{2t})(1+\alpha_2)^2 + (1-e^t)^2\alpha^2)}{q+pe^{2t}}$ and $C_2 := q(e^{2|t|} - 1)(e^{-t} + \alpha)^2$.

Finally, we combine these bounds with previously obtained expressions to get

$$\|M_n'(t)\|_{L_2(\Omega)}^2$$

$$= \mathbb{E}(M_n'(t))^2$$

$$= \frac{\sum_{m=0}^{n-1}\left(\widehat{\frac{\partial^2}{\partial s\partial t}f_{n-m}\star h_m}(t,t) - 2n\alpha\widehat{\frac{\partial}{\partial s}f_{n-m}\star h_m}(t,t) + n^2\alpha^2\widehat{f_{n-m}\star h_m}(t,t)\right)}{(q+pe^t)^{2n}}$$

$$\leq C_1\sum_{m=0}^{n-1}(n-m)^2\left(\frac{q+pe^{2t}}{b(q+pe^t)^2}\right)^{n-m} + C_2\sum_{m=0}^{n-1}(n-m)^2\left(\frac{pe^{2t}}{b(q+pe^t)^2}\right)^{n-m}$$

$$\leq (C_1+C_2)\sum_{m=1}^{n}\left(\left(m^{\frac{1}{m}}\right)^2 r\right)^m.$$

As shown in the previous theorem, for each p there is a positive integer c such that $r < 1$ whenever $t \in [-2^{-c}, 2^{-c}]$. Moreover, $C := \sup_{t\in[-2^{-c},2^{-c}]}\{C_1 + C_2\} < \infty$. Since $m^{\frac{1}{m}} \to 1$ as $m \to \infty$, there is a fixed positive integer d such that whenever $m \geq d$, we have $m^{\frac{1}{m}}r < 1$. Thus we may define $r' = d^{\frac{1}{d}}r < 1$, so that for any $t \in [-2^{-c}, 2^{-c}]$ and $n > d$ we have

$$\|M_n'(t)\|_{L_2(\Omega)}^2 \leq C\sum_{m=1}^{d-1}\left(\left(m^{\frac{1}{m}}\right)^2 r\right)^m + C\sum_{m=d}^{n}(r')^m < D + C\sum_{m=d}^{\infty}(r')^m < \infty,$$

where $D := C\sum_{m=1}^{d-1}\left(\left(m^{\frac{1}{m}}\right)^2 r\right)^m$. The number $D + C\sum_{m=d}^{\infty}(r')^m$ is independent of n, so for $t \in [-2^{-c}, 2^{-c}]$ we have $\sup_n \|M_n'(t)\|_{L_2(\Omega)} < \infty$. $\qquad\square$

We can now prove the central limit theorem.

Proof of Theorem 3.3.1. For a given $p \in (0,1)$, choose a sufficiently large positive integer c such that, as established in the previous two theorems, we have

$\sup_n \|M_n(t)\|_{L_2(\Omega)} < \infty$ and $\sup_n \|M'_n(t)\|_{L_2(\Omega)} < \infty$ for $t \in [-2^{-c}, 2^{-c}]$. Then for t in this interval, the convergence of $M_n(t)$ to $M_\infty(t)$ is uniform. This implies in particular that M_∞ is P-a.s. continuous at $t = 0$. Thus we obtain

$$e^{-npt} \hat{Q}_n \left(\frac{t}{\sqrt{npq}} \right) = M_n \left(\frac{t}{\sqrt{npq}} \right) e^{-npt} \left(q + pe^{\frac{t}{\sqrt{npq}}} \right)^n \longrightarrow M_\infty(0) e^{\frac{t^2}{2}} = e^{\frac{t^2}{2}}$$

which suffices to prove the claimed CLT. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Returning to the polymer model, we have shown that in a region of weak disorder (see the remark after Theorem 3.2.1), the polymer endpoints do not fracture in the limit. More precisely, the polymer endpoints are a.s. distributed over a random interval $[K_n, n - L_n]$ such that each point of the interval contains at least one particle. Regardless of disorder type, if one suitably centers and scales the distribution of particles then the asymptotic distribution of particles is a.s standard normal.

By an application of a now standard result of [11] [27], [20] for branching random walk, the speed $\lim_{n\to\infty} K_n/n$ of the leftmost particle is zero. As for the rightmost particle, $n - L_n$, its speed is a.s. determined for $\beta \neq 0$ by the almost sure limit

$$\lim_{n\to\infty} \frac{L_n}{n} = \frac{1}{1 + e^{\theta^* - \beta}}, \qquad\qquad\qquad (3.3.3)$$

where $\theta^* \equiv \theta^*(\beta)$ is the unique positive solution to $\theta^* e^{\theta^* - \beta} = (1 + e^{\theta^* - \beta}) \ln(2q_\beta(1 + e^{\theta^* - \beta}))$.

In the case of the simple symmetric random walk with $b > 2$, the rightmost

particle travels at unit speed making $\lim_{n\to\infty} \frac{L_n}{n} = 0$. In fact, as noted at the outset, Grill [19] shows that $L_n \to L < \infty$ a.s., and that the particles are almost surely eventually distributed at each point of an interval $[K, n - L]$ in such cases.

Regarding the role of β_c in such results, it follows by application of results on tree polymers [16], [46], that the respective limits

$$\lim_{n\to\infty} \frac{\ln \sum_{k=0}^n W_{n,k} e^{2\beta k}}{n} \quad \text{and} \quad \lim_{n\to\infty} \frac{\ln \sum_{|s|=n} W_n^{-1} e^{r \sum_{j=0}^n s_j + 2\beta H_n(s)}}{n} \tag{3.3.4}$$

exist almost surely, with a distinctive change in structure at $\beta = \beta_c$. For example, the former a.s. limit coincides with the speed of the right-most particle in the case of strong disorder, illustrating an aspect of *polymer localization*, and the latter coincides with the diffusive limit for symmetric random walk in the weak disorder range of $\beta < \beta_c$. General formulae are available for all values of β above and below criticality for both limits, see [16], and [46], respectively.

Beyond this, such effects of disorder on the support of $W_{n,k}/b^n$ reside in more detailed analysis of the particle counts $W_{n,k}$ at the walker positions $H_n(s), s \in \{-1,1\}^n$. In the case of non-lattice branching random walk, results recently obtained in [30] for the case of critical strong disorder, also referred to as the 'boundary case' in the branching random walk literature, show that the limiting walker locations centered at the left most particle comprise a Poisson cluster (decorated) point process. In this regard, it remains an interesting open problem to determine the existence of an a.s. weak limit point of the sequence $\text{prob}_n(ds)$ under strong disorder. See [24] for related results in the case of *critical* strong disorder.

4 Branching for a Decision: A Site-Selection Model for Honeybee Swarms

4.1 Introduction

It has long been observed that honeybees will swarm in the late spring and early summer. After an existing hive has outgrown itself, the mother queen and about half the worker bees leave the hive and swarm for the purpose of choosing a new site [39], [42], employing a small fraction of so-called scout bees (typically several hundred) to carry out the process [42]. The earliest work on how the swarm goes about making a decision was done by Lindauer [28],[29]. He observed three basic elements of the process. First, bees can communicate the quality and location of a potential site by means of a behavior termed waggle-dancing. Second, bees may cease to dance for a site they currently favor, switching to a different site after observing another bee dancing for it. Third, eventually all the bees appear to be dancing for just a single site. Shortly after this occurs, the swarm flies en masse to the chosen site. This suggests a certain "democratic" aspect to the process. Seeley has informally termed the process "Honeybee Democracy," the title of his excellent popular work on the subject [38].

Field investigations on the mechanisms of this process continued many decades later in [39]. In particular, they ruled out a high mortality rate as the primary

way that bees stop dancing for the sites that are not ultimately chosen, implying that dancing bees either simply stop dancing or switch and dance for another site. The tendency of scout bees to stop dancing over time and the capacity to switch allegiance has also been observed in [37]. Scout bees reduce their waggle-dancing over time until they are no longer dancing at all, and the reduction, on average, is linear [37]. Both these mechanisms will be important features of the model to follow.

The early work of Lindauer seemed to imply that scout bees somehow arrived at a consensus. Later studies cast doubt on this hypothesis. In [41], four swarms were presented with five nest-boxes (sites of high and roughly equal quality) and with only one such box, in paired trials. In the five-box trials, scouts were dispersed among the five boxes, and the swarm's decision was significantly delayed compared to the one-box cases. Also see [40]. Results like these point to an alternative: the quorum-sensing hypothesis. This asserts that once a sufficient number of scouts are dancing for some site, the swarm chooses it, even if other scouts are still dancing for different sites [42]. The model incorporates this hypothesis.

Another well-studied aspect of the site-selection process, although not the focus here, is the mechanism by which the swarm flies to the chosen site. Only a minority of the swarm's population are scouts, yet all the bees in the swarm manage to fly to the new site. In [3], support is provided for the so-called "vision hypothesis," whereby the scouts with information on site-location streak through the swarm in the direction of the site in order to communicate that information to the other bees. Strong support for this hypothesis is also found in [36].

Finally, we mention some of the other modeling approaches that have been taken for the site-selection process. In [15], the spread of site-information among bees is modeled with a system of ODEs in a similar manner to the spread of an infectious disease in a population. Myerscough in [32] developed a Leslie matrix model in which dances age, reproduce themselves, and die. The model predicts that the single best site will be found, although observation has shown that this does not always happen [29], [39]. In [33], an individual-based simulation model is constructed to explore how tuning various behaviors leads to a workable balance between speed and accuracy of the decision-making. Another individual-based model is found in [23], the key feature of which is that bees first attempt to find a dance to follow, but become scouts if they fail to find one. This allows the swarm to potentially increase the number of bees looking for new sites if there is a dearth of high-quality sites known to the swarm. Lastly, in [35], a density-dependent Markov process models the site-selection. Its main conclusion is that a swarm will choose the better of two low-quality sites more consistently than the better of two high-quality sites.

The model presented here is an individual-based model that shares some features with standard models for branching processes and the random walk. It is not a branching random walk, since the usual assumption of i.i.d. displacements from the parent (dancing bee) is not made. Indeed, the capacity of dancing bees to switch to a new site after being influenced by the dances of the other bees is a central part of the model. Crucially, these switching decisions should not depend on the site for which the bee is currently dancing. Moreover, the tendency of bees

to eventually stop dancing makes it is necessary to track the time since a bee first departed at the level of the individual bee.

4.2   General Setup and Assumptions

The following minimal notation will aid in describing the assumptions below:

1. Sites are enumerated $0 < 1 < 2 < \ldots < n$ according to increasing quality, with 0 reserved for bees that are not dancing for a site. The enumeration is merely for convenience and labeling purposes. The only information conveyed by being at site of a given value is its quality relative to the other sites; the value does not indicate its absolute quality nor how much better or worse it is than the other sites.

2. Time is denoted by t and takes the discrete values $0, 1, 2, \ldots$.

3. The number of scouts dancing for site i at time t is denoted by $N_t(i)$.

   The model assumptions are now described in detail:

1. Initial exploration distribution: Scout bees in their initial explorations find a potential site according to a distribution Q_0, fixed for each scout and throughout the process. Q_0 is meant to encode the relative difficulty in finding the various sites. This idealization is probably not too crude, since the site-selection process occurs over at most several days [42].

2. Quorum-sensing: the site-selection process stops at the time step during which some site has garnered support greater than or equal to some fixed

number M. In the event of a tie the higher quality site is selected. This last rule is of course arbitrary, simply made for convenience.

3. Waggle-dance communication: Site-quality and location is communicated to other potential dancers in swarm by waggle-dances performed when a scout returns. Scouts either dance for the found site or not according to its quality and how many time steps have passed since it first left the swarm. The quality of each site is a positive integer.

4. Dances cease: Following [39], it is assumed that dancing bees eventually stop dancing. A scout will dance for a higher-quality site for more time steps than one of lower-quality. A scout that does not dance for a lower-quality site at a given time step may still dance for one of higher-quality at a later time step.

5. Random-observation, switching and branching-recruitment: It was observed in [17] that potential "recruits," i.e. bees observing dances, will follow a dance at random. Moreover, a bee currently dancing for some site may switch allegiance and begin dancing for another [39]. These observations are incorporated into assumptions about the behavior of scouts and potential recruits. Let $p_{\text{follow}} \in [0, 1]$.

   (a) With probability p_{follow}, a returning scout will choose another dancing bee's site to investigate. Denote the total number of bees dancing at a given time step t by N_t, so that $N_t = \sum_{i \neq 0} N_t(i)$. By another assumption below, it will always be the case that $N_t > 0$. At time t, the recruit is

assumed to randomly select the dance to follow according to the random
distribution

$$Q_t := \left\{ p_i : p_i = \frac{N_t(i)}{N_t}, i = 1, \ldots, n \right\}.$$

(b) With probability $1 - p_{\text{follow}}$, the returning scout convinces a random
number of new bees (not from the population of current dancers) to
investigate her site. The number is determined by a fixed offspring
distribution D_{off}.

(c) Note: With $p_{\text{follow}} = 1$, every bee observes another bee at each time step.
With $p_{\text{follow}} = 0$, a bee never switches allegiances. Intermediate values
of p_{follow} allow both mechanisms to operate.

6. Scouts per time step: For simplicity, it is assumed that some fixed number c
of scouts embark at each time step.

7. Time-synchronization: Scouts are assumed to explore, dance, and possibly
convince others in a single step. New followers or initial scouts repeat this
in the next step.

4.3   The General Process

The process evolves in discrete time according to the following general construct.

1. At time $t = 0$, c initial explorers (scouts) independently find sites according
to Q_0. Upon returning to the swarm, they dance for the sites they find,
making their information available to other scouts.

2. At times $t \geq 1$, c new scouts are dispatched, in the same manner as at time $t = 0$. Dancing scouts examine sites communicated to them in the previous step and dance for these sites (or not) according to site-quality and the tendency to cease dancing. Bees determine a site to examine depending on their status:

   (a) Scouts follow the initial observation distribution Q_0,

   (b) new bees following a dancing bee investigate the "parent bee's" site, and

   (c) currently dancing bees that switch allegiance choose a site according to the random distribution Q_t.

3. The process continues in this manner until a quorum is reached, i.e. $N_t(i) \geq M$ for some i and some t, and this t is the smallest t for which this condition is satisfied.

See Figure 4.1 for a schematic diagram of a single time-step.

4.4   Mathematical Formulation

The purpose of this section is to specify the model precisely. Let n be a positive integer. Define a stochastic process, the *bee process*, by

$$\{B_t(r) : r \in \mathbb{N}, t \in \mathbb{N}_0\}$$
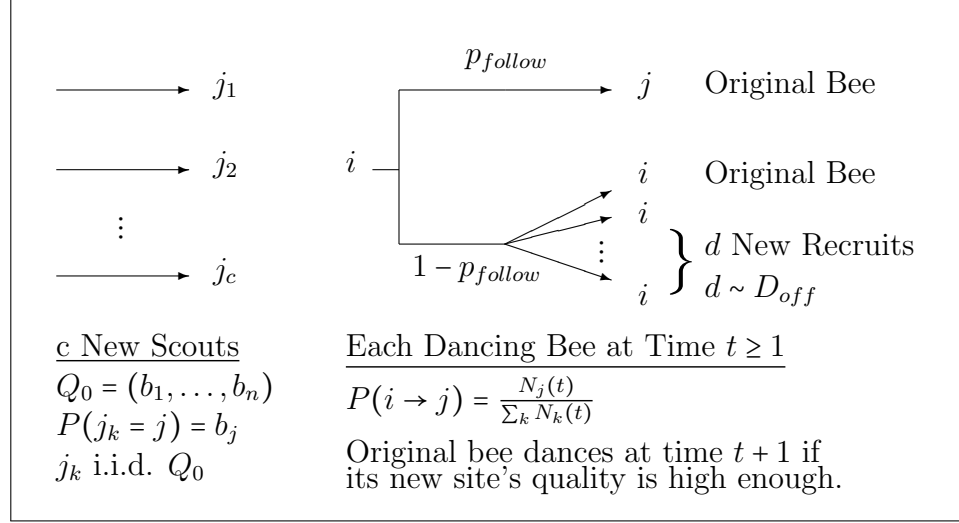
Figure 4.1: Schematic diagram of one time-step

on a probability space $(\Omega, \mathcal{F}, P)$ taking values in the state space $S := \mathbb{N}_0 \times \{0, 1, \ldots, n\}$. For $i \neq 0$, $B_t(r) = (t_0, i)$ conveys that the bee labeled r is dancing for site i at time t, having first embarked at time $t_0 \leq t$. If bee r has not initially embarked at time t, then $B_t(r) = (0,0)$. If $i = 0$ and the scout has embarked by time t, then $B_t(r) = (t_0, 0)$. If a bee is at site 0 at any given time, she has either not yet embarked or will not dance for any site, having first embarked too many steps ago.

Denote the number of scouts that have embarked by time t (inclusive) by $\eta(t)$. Suppose c is some positive integer. Initial scouts embark at time $t = 0$. For $t \geq 1$, the scouts labeled $\eta(t-1) + 1, \ldots, \eta(t-1) + c$ embark at time t.

Let $q = (q_1, \ldots, q_n), q_k \in \mathbb{N}$, be a vector of positive integers coding the quality of the sites. For convenience we require $q_1 \leq \ldots \leq q_n$. $q_k = d$ means any scout that has embarked fewer than d steps ago has the capacity to dance for site k. Any scout that embarked more than $\max\{q_k\}$ steps ago will not dance for any site. Suppose

$Q_0 = (b_1, \ldots, b_n)$ denotes the initial exploration distribution, with $b_i > 0$ for all i.

Define for $i \geq 1$

$$N_t(i) = \sum_{r \in \mathbb{N}} \mathbf{1} \left[B_t(r) = (t_0, i) \text{ and } t_0 + q_i > t \right],$$

i.e. $N_t(i)$ is the number of bees dancing for site i at time t. The right-most condition accounts for the dance-ceasing behavior of individual bees, depending on when they first embarked.

Let $p_{\text{follow}} \in [0, 1]$. Suppose $X_t(r)$ are i.i.d. $0, 1$-valued random variables, $P(X_t(r) = 0) = p_{\text{follow}}$ and $P(X_t(r) = 1) = 1 - p_{\text{follow}}$, which are independent of all other random variables in this model.

1. If $X_t(r) = 1$, a given bee r at time t chooses a new site to explore according to the probabilities $p_{ij}(t, r)$ which give the probability that a bee r dancing for site i at time t dances for site j at time $t + 1$. Define

$$p_{i0}(t, r) = 1 \quad \text{if } B_r(t) = (t_0, i) \text{ and } t_0 + q_i \leq t$$

$$p_{ij}(t, r) = \begin{cases} b_j & \text{if } B_t(r) = (t, 0), \\ \frac{N_t(j)}{\sum_{k=1}^n N_t(k)} & \text{if } \sum_{k=1}^n N_t(k) > 0 \text{ and } j \geq 1 \end{cases}$$

2. If on the other hand $X_t(r) = 0$, then the bee r at time t convinces $d_t(r)$ new scouts to investigate her site. $d_t(r)$ takes values in $\mathbb{N}_0$ according to some common offspring distribution D_{off}.

Suppose $M \in \mathbb{N}$ is the quorum-sensing threshold. Define $T_M = \inf\{t : N_t(j) \geq M \text{ for some } j\}$. Then the process will terminate at time T_M in a decision in favor

of site j if

$$j = \max_k \{k : N_{T_M}(k) \geq M\}.$$

4.5   Reaching a Quorum

This section provides some conditions under which a quorum is guaranteed. Call a quorum threshold $M \geq 0$ *attainable* if $P(N_t(i) \geq M$ for some $t, i) > 0$. The following result partially characterizes the attainable threshold levels:

Theorem 4.5.1. *For the bee process with parameters c, p_{follow}, offspring distribution D_{off}, quality vector $q = (q_1, q_2, \ldots, q_n)$, and initial observation distribution Q_0:*

1. *If $p_{follow} = 1$, then a threshold M is attainable if and only if $M \leq c \max_i \{q_i\}$.*

2. *If $p_{follow} < 1$ and $(1 - p_{follow})\mathbb{E}(D_{off}) \leq 1$, then thresholds satisfying $M \leq c \max_i \{q_i\}$ are attainable.*

3. *If $p_{follow} < 1$ and $(1 - p_{follow}\mathbb{E}(D_{off})) > 1$, then any threshold M is attainable.*

Moreover, in cases one and three,

$$P(N_{t'}(i) \geq M \text{ for some } i \text{ and some } t') = 1. \tag{4.5.1}$$

Proof. For the first statement, suppose k is such that $q_k = \max_i \{q_i\}$, i.e. k is the highest quality site. Denote the set of bees at time t that embarked fewer than q_k steps ago by $R_t(k)$. These are precisely the bees that could dance for site k

at t, and for any $t' \geq q_k$, we clearly have $|R_t(k)| = cq_k$. Thus $M \leq c\max_i\{q_i\}$ is necessary. To see that it is also sufficient, we first show that for some $t' \geq q_k$ we have $p_{ik}(t', r) > 0$ for all bees $r \in R_{t'}(k)$. Initial scouts independently dance for sites according to Q_0, so the probability that no initial explorer dances for site k for *any* time step $t \geq q_k$ is clearly zero. Thus, almost surely, there is a $t' \geq q_k$ such that at least one scout is dancing for site k at t', and therefore $p_{ik}(t', r) > 0$. Since $M \leq cq_k$, we have

$$P(N_{t'+1}(k) \geq M) \geq P(N_{t'+1}(k) = cq_k) \geq p_{ik}^{cq_k} > 0. \qquad (4.5.2)$$

For the second statement, one may repeat the sufficiency argument above with $p_{follow}p_{ik}$ in place of p_{ik}.

To prove the third statement, consider the following branching process obtained from the original process: a single initial scout begins dancing for a site i. With probability $1 - p_{follow}$, this scout convinces some number of other new scouts to dance for site i according to D_{off}, or convinces zero new scouts with probability p_{follow}. Each new scout independently repeats this process. This is an ordinary Galton-Watson branching process which is supercritical by hypothesis, so its extinction probability is less that one. Since each initial scout initiates its own such branching process independently of all the others, and there are infinitely many initial scouts, the overall population of dancers survives almost surely, and thus grows without bound almost surely by standard results on supercritical GW processes. It is clear that the original process grows at least as fast as this collection

of Galton-Watson processes, since all scouts dance for *at least* one time step in the original process. Therefore $\sum_j N_t(j) \to \infty$ as $t \to \infty$, and since the number of sites is assumed to be finite, any M is attainable.

(4.5.1) is obvious in the third case. For the first case, independence and an infinite number of initial scouts guarantee that there are a.s. an infinite number of time points t' for which (4.5.2) holds, and the probabilities $p_{ik}(t', r)$ are always at least $(cq_k)^{-1}$ for $r \in R_t(k)$. $\qquad\square$

4.6   Simulations and Discussion

In this section the factors influencing the swarm decision-making process are explored using computer simulations. The parameter space is large, and all permutations cannot possibly be explored. Nevertheless, the following scenarios will illustrate the most important features of the model. The code for the simulation is recorded in an appendix.

Two basic scenarios and some more specialized simulations are considered. The quality vector q and initial exploration distribution Q_0 are held fixed for each scenario, since these encode environmental information for the most part. The other parameters tune the behavioral mechanisms of individual bees. These will be varied to assess how well the model performs in choosing the best site. The offspring distribution will be assumed Poisson with parameter λ. Most of the analysis below deals with the trade-off between speed and accuracy, and the optimal choice will mostly be achieved by intermediate values of the parameters

p_{follow} and the branching parameter λ. It is still useful however to consider the model for extreme values of these parameters to better understand its behavior.

4.6.1 Scenario 1

The simulated swarm is presented with 4 sites of acceptable quality and a single higher quality site. A quality vector of $q = (10, 10, 10, 10, 15)$ and a uniform initial exploration distribution are used. We first consider the model for the extremes $p_{follow} = 0$ and $p_{follow} = 1$, choosing $\lambda = 1$ in the first case. Both are run with $c = 3$ initial scouts per time step and a quorum threshold of $M = 80$.

For the model with $p_{follow} = 1$, support for the best site quickly builds when bees stop dancing for the lesser quality sites, and settles at a level just below its maximum $cq_5 = 45$ (see Figure 4.2). The quorum of $M = 80$ is not attainable, so a
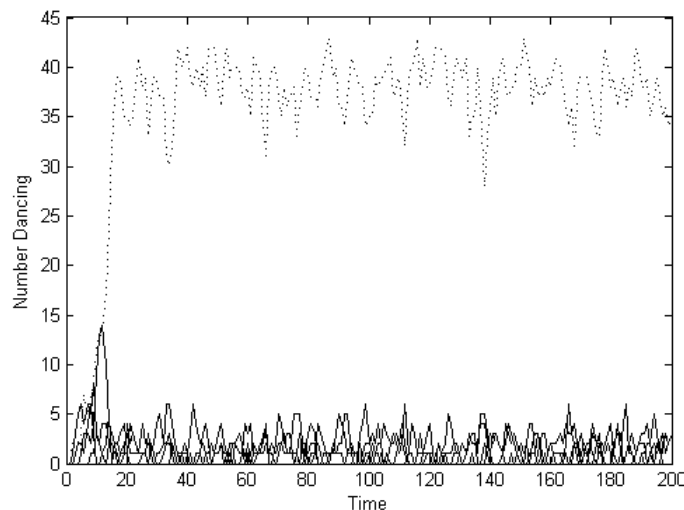
Figure 4.2: Scenario 1 with $p_{follow} = 1$ (best site dotted).

quorum is never reached. This is the main limitation of the $p_{follow} = 1$ case. The process arrives at a quorum for only a relatively narrow range of quorum values.

For the other extreme with $p_{follow} = 0$, bees cannot switch allegiance from one site to another, and support will tend to build exponentially for all sites when the parameter λ is large (see Figure 4.3). The process tends to end in a decision for a
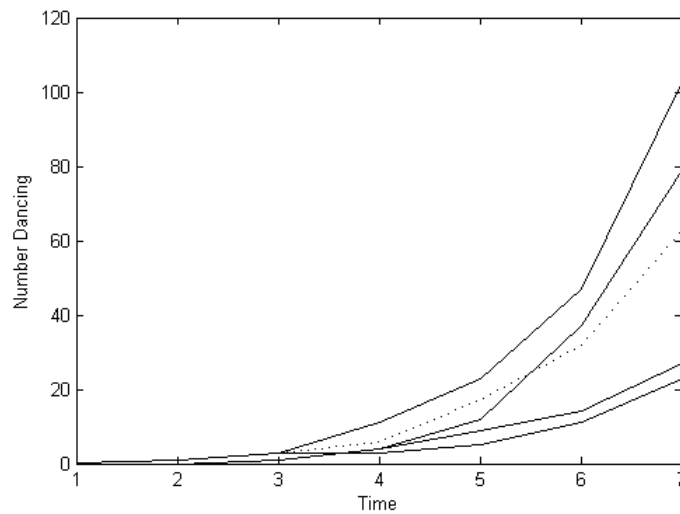
Figure 4.3: Scenario 1 with $p_{follow} = 0$ and $\lambda = 1$ (best site dotted).

random site that gets off to an early lead. Small values of λ that keep a quorum from being reached before bees stop dancing for inferior states do perform better, but not optimally.

For the most part, the general behavior of these extreme cases is similar for all scenarios. It is evident that an effective balance should be sought between the parameters p_{follow} and λ. For any $p_{follow} < 1$, increasing the branching rate λ will speed up the process, but λ is not the only parameter that does. Increasing the

number of scouts c that embark per time step or decreasing the quorum threshold M will also have this effect. To evaluate the speed versus accuracy trade-off, the simulation was run 30 times for each $p_{follow} \in [0.1, 0.9]$ and $\lambda \in [0.1, 1]$, in increments of 0.05. In some instances, where computation time is an issue, smaller intervals were used. Quorum thresholds of $M = 20, 80$ and scouts per time step of $c = 3, 10$ were tested. These choices were made simply to compare the effects of relatively high and low values of each of these parameters in a manageable way, and they will suffice to test the model. Both the mean time to quorum and the proportion of trials for which the best site was selected are considered. Contour plots display the results, (see Figures 4.4 through 4.11).
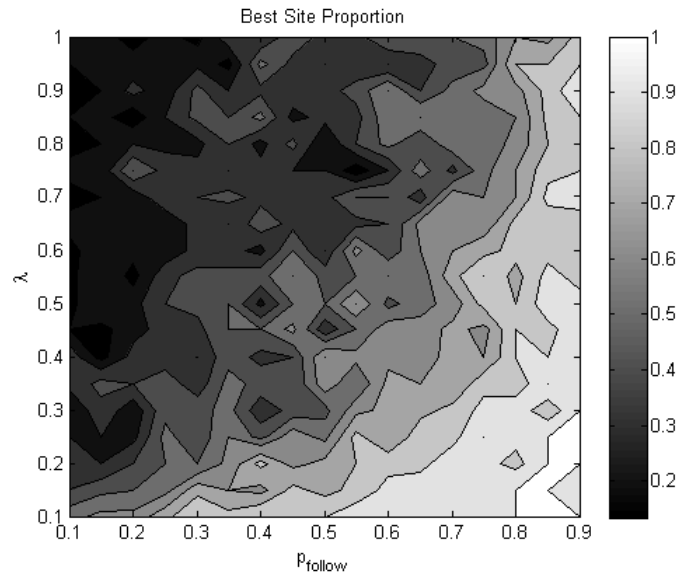
Figure 4.4: Scenario 1 best site proportion with $M = 20$ and $c = 3$.
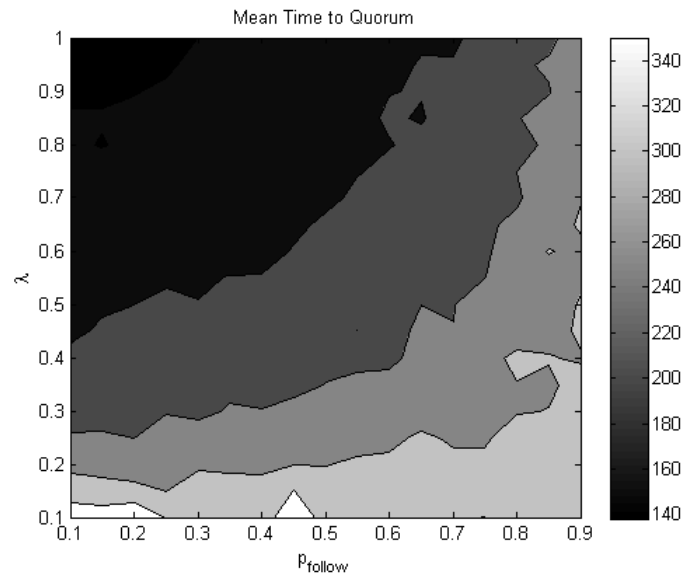
Figure 4.5: Scenario 1 mean time to quorum with $M = 20$ and $c = 3$.
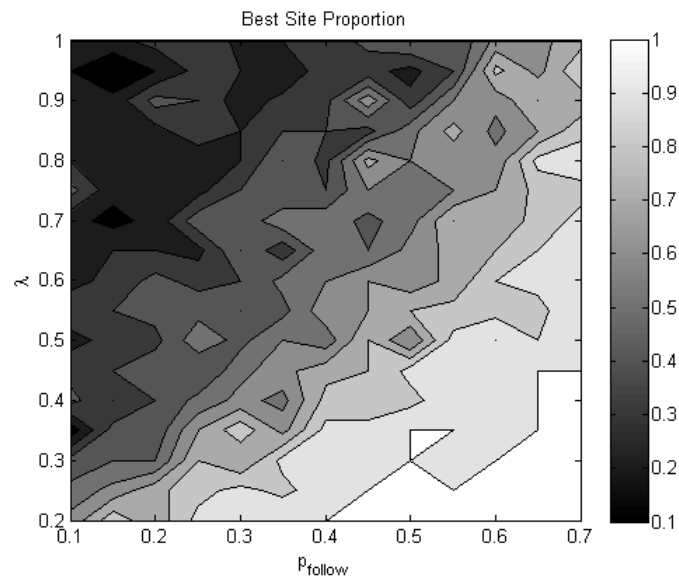
Figure 4.6: Scenario 1 best site proportion with $M = 80$ and $c = 3$.
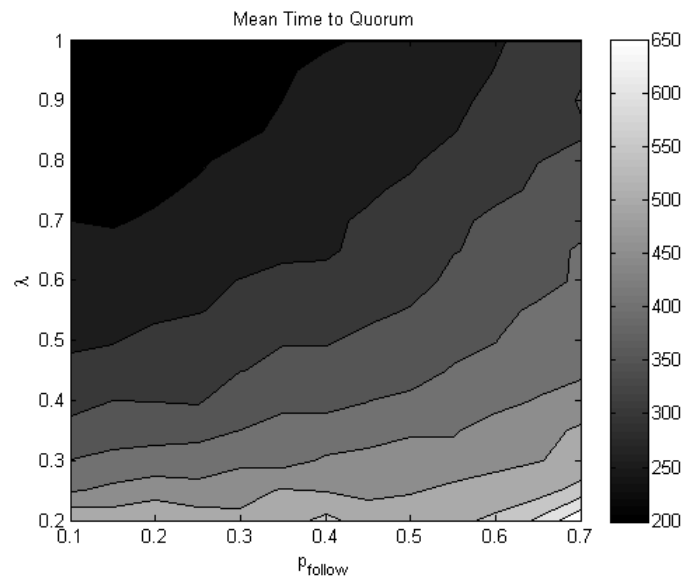
Figure 4.7: Scenario 1 mean time to quorum with $M = 80$ and $c = 3$.
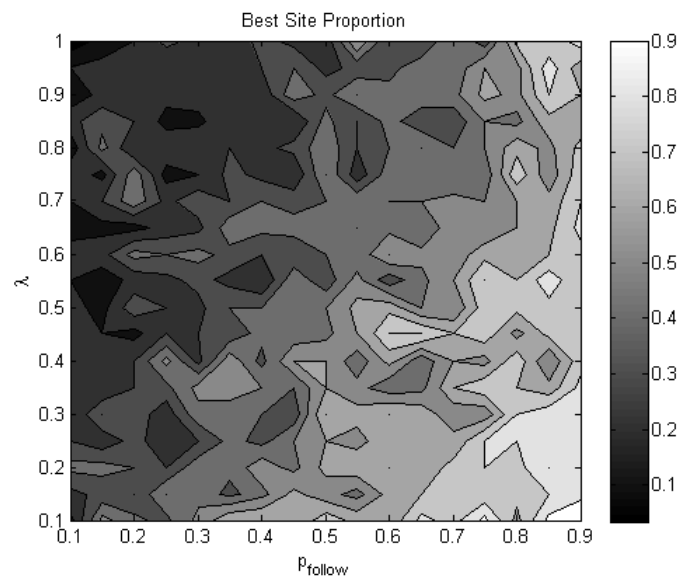
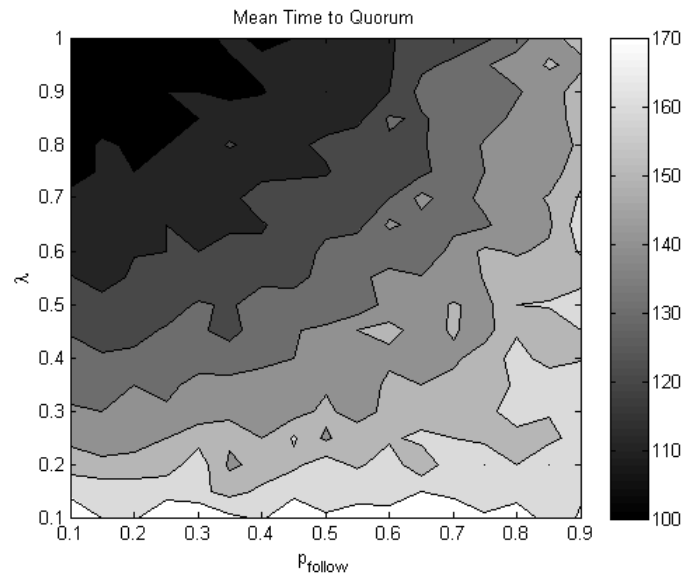Figure 4.8: Scenario 1 best site proportion with $M = 20$ and $c = 10$.

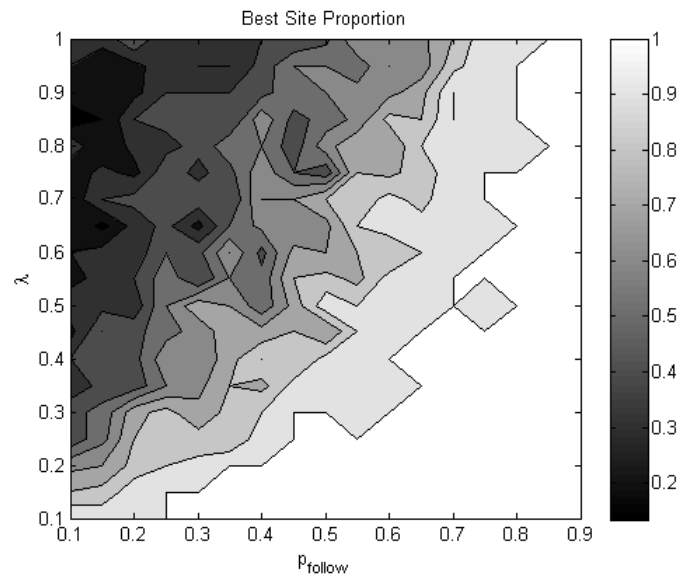Figure 4.9: Scenario 1 mean time to quorum with $M = 20$ and $c = 10$.

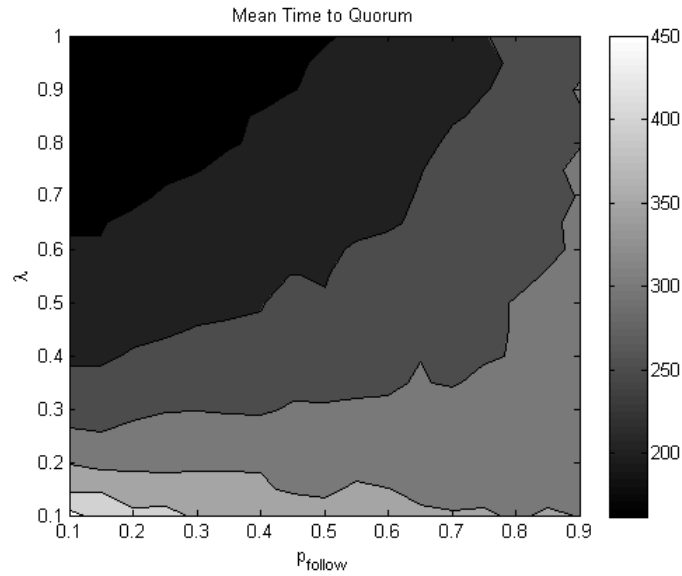Figure 4.10: Scenario 1 best site proportion with $M = 80$ and $c = 10$.

Figure 4.11: Scenario 1 mean time to quorum with $M = 80$ and $c = 10$.

The most prominent feature of these plots is the loss of accuracy in selecting the best site as the speed of the process is increased in any way. The effect due to the difference in scouts per time step is small, but decreasing the quorum threshold has a much larger effect. Smaller values of the branching parameter λ appear to allow for close to optimal decisions across a large range of both p_{follow} values and quorum thresholds. This is important for two reasons. One, the switching behavior governed by p_{follow} may not occur a very large percentage of the time in swarms. Two, the quorum threshold is not necessarily restricted to a narrow range in an actual swarm.

Increasing the quorum threshold appears to result in an improvement in accuracy for λ values that are not too large. Large values magnify early leads and an increase in the threshold will typically just delay a decision in favor of the leading

site, rather than allowing support for a potentially superior site to build before a quorum is reached. This effect does occur for the values of p_{follow} and λ that are closer to optimal. Fixing $c = 10$, $p_{follow} = 0.8$, and $\lambda = 0.2$ and considering quorum thresholds between 0 and 100 demonstrates this effect. See Figure 4.12. Each point is a fraction of the total number of times the model was run (30) for each threshold level.
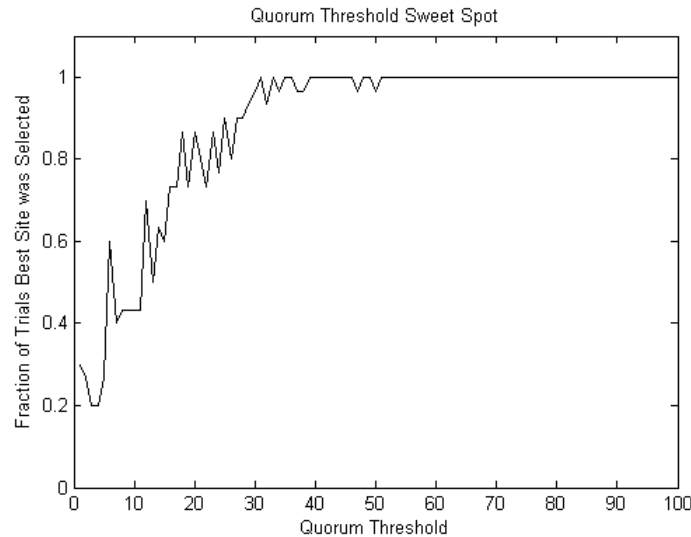

Figure 4.12: Scenario 1 effect of increasing quorum threshold with $c = 10$, $p_{follow} = 0.8$, $\lambda = 0.2$.

At the larger threshold level, the best site is selected a high percentage of the time for a wider range of λ values. Although increasing the threshold increases the mean time to quorum overall, a larger λ will mitigate this effect, and Figures 4.10 and 4.11 show that the increase in mean time to quorum due to an increase in λ will not result in a significant loss in accuracy.

4.6.2 Scenario 2

In this scenario the effect of making the best site harder to find is investigated. As noted in the introduction, field studies have shown that the best site is not always chosen by the swarm, and a potential reason for this is that the best site is found too late in the decision-making process, or not found at all. All parameters from scenario 1 are retained except the initial exploration distribution, which is defined to be $Q_0 = (0.24, 0.24, 0.24, 0.24, 0.04)$. Typical simulation runs for the high threshold, $M = 80$, at both $c = 3$ and $c = 10$ are included in Figures 4.13 and 4.14. See Figures 4.15 through 4.22 for the analysis of p_{follow} versus λ. The corresponding plot showing the effect of increasing the quorum threshold is in Figure 4.23.
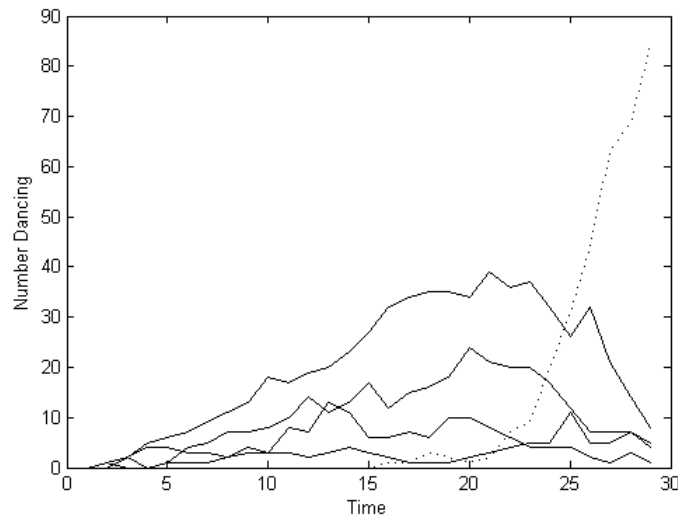

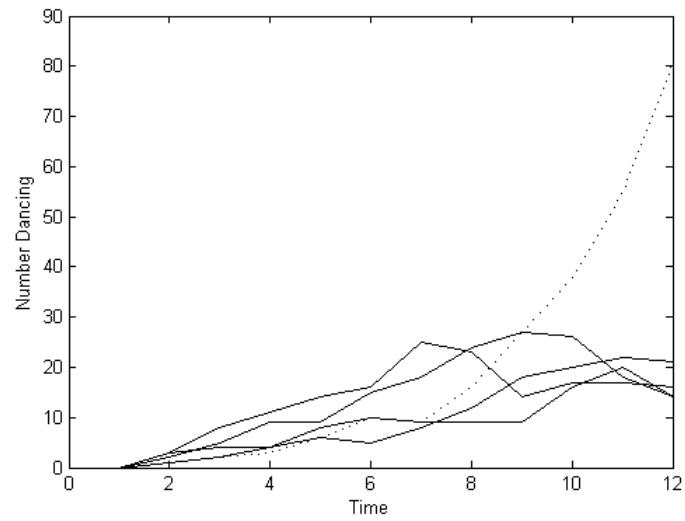Figure 4.13: Scenario 2 typical run with $M = 80$, $c = 3$ (best site dotted).

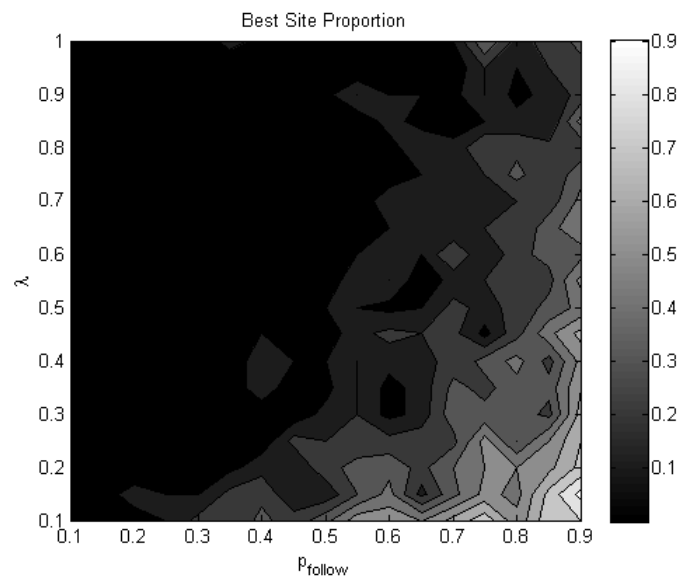Figure 4.14: Scenario 2 typical run with $M = 80$, $c = 10$ (best site dotted).

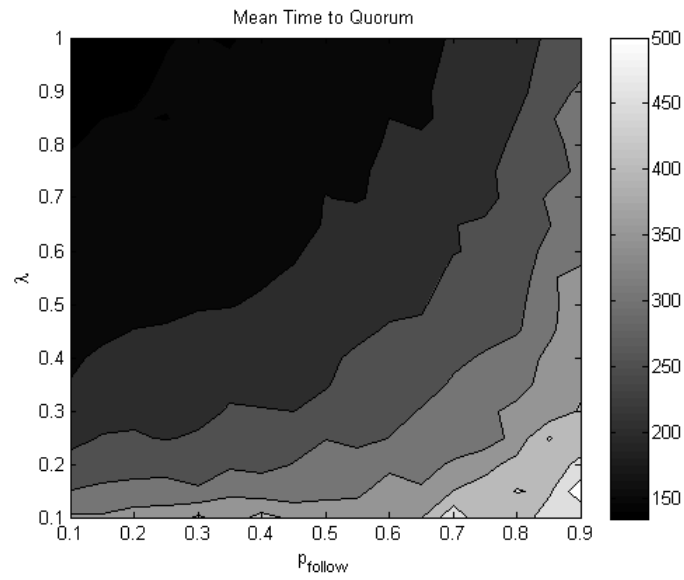Figure 4.15: Scenario 2 best site proportion with $M = 20$ and $c = 3$.

Figure 4.16: Scenario 2 mean time to quorum with $M = 20$ and $c = 3$.

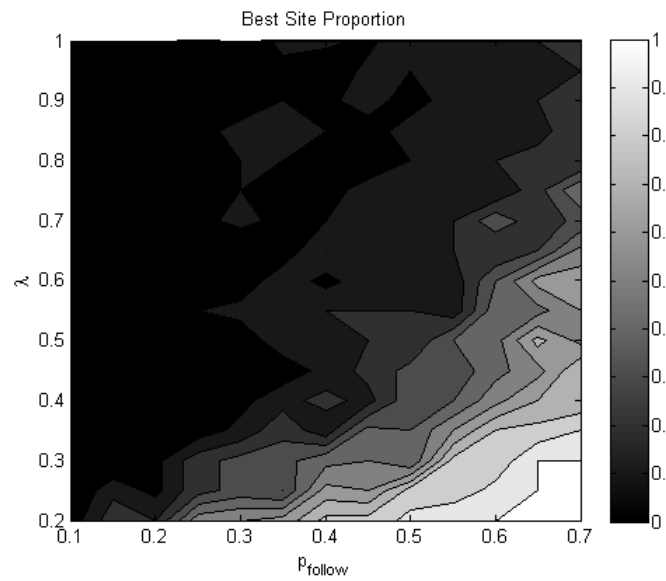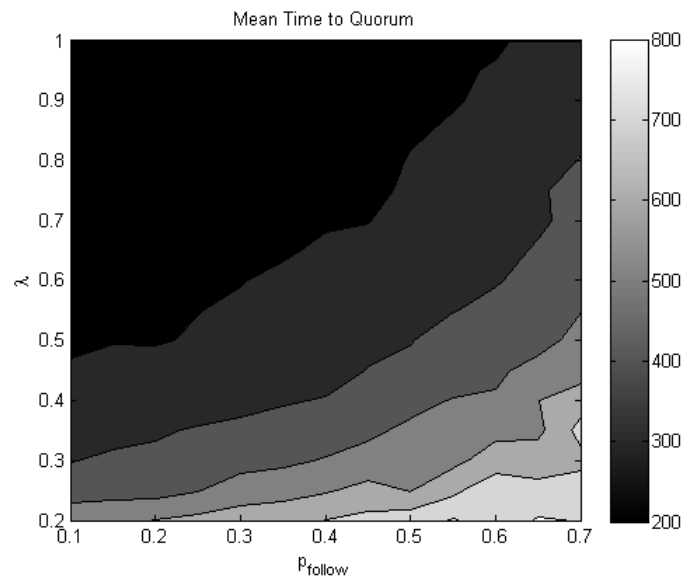Figure 4.17: Scenario 2 best site proportion with $M = 80$ and $c = 3$.

Figure 4.18: Scenario 2 mean time to quorum with $M = 80$ and $c = 3$.
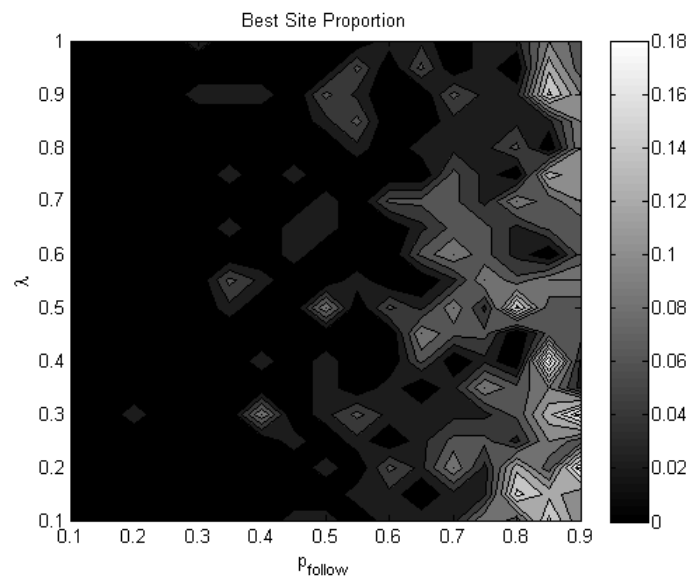

Figure 4.19: Scenario 2 best site proportion with $M = 20$ and $c = 10$.

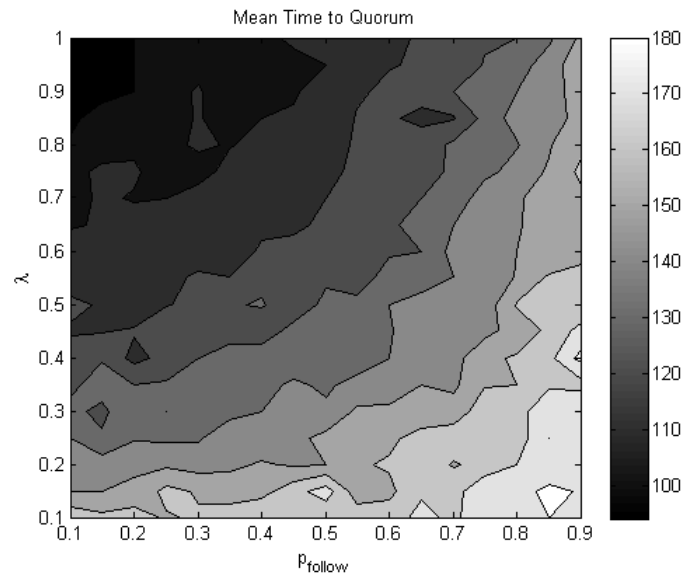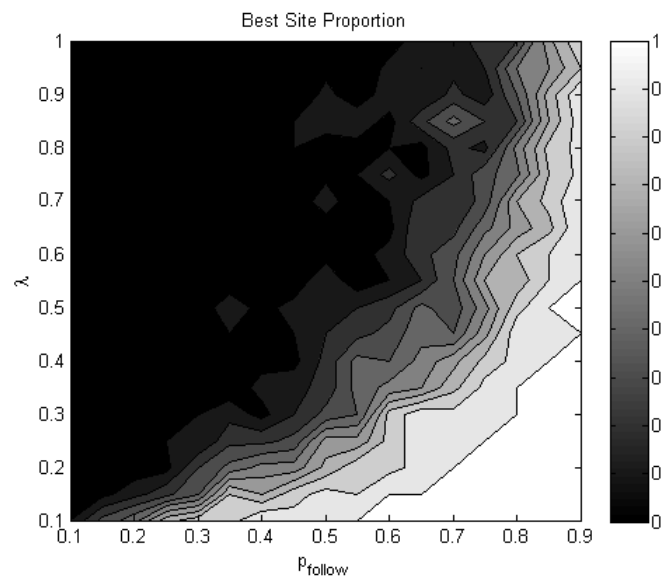Figure 4.20: Scenario 2 mean time to quorum with $M = 20$ and $c = 10$.

Figure 4.21: Scenario 2 best site proportion with $M = 80$ and $c = 10$.

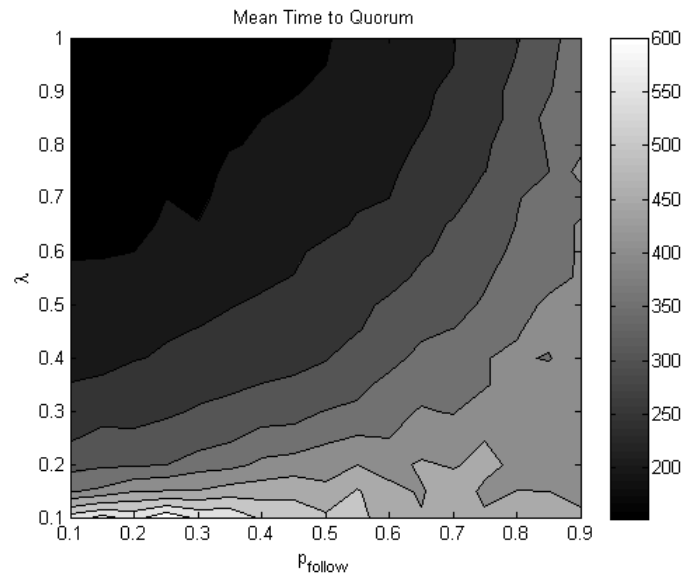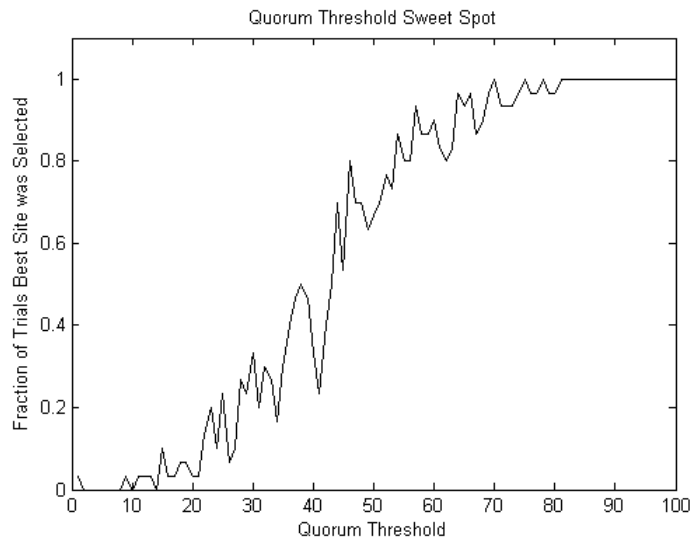Figure 4.22: Scenario 2 mean time to quorum with $M = 80$ and $c = 10$.

Figure 4.23: Scenario 2 effect of increasing quorum threshold with $c = 10$, $p_{follow} = 0.8$, $\lambda = 0.3$.

For the most part the performance in scenario 2 for the low threshold, $M = 20$, is quite poor. Although the best site is selected frequently for values of λ and p_{follow} in the lower right of Figure 4.15, this is mostly due to $M = 20$ being relatively close in value to $cq_5 = 45$, so behavior is similar to the model with $p_{follow} = 1$ (no branching). Increasing the number of scouts per time step drives performance down drastically, in contrast to scenario 1. This is likely due to an excess of early recruitment in favor of the relatively easy to find sites.

In the range of λ and p_{follow} values that lead to a high probability of choosing the best site, a good site that is difficult to find will delay reaching a quorum. This appears to be the main effect. This is also confirmed by Figure 4.23. Increasing the quorum threshold increases the chance that the swarm chooses the best site, but larger thresholds and longer times are required to achieve near perfect selection.

4.6.3 A Possible Quorum Limit

It is easy to speculate on the basis of Figures 4.12 and 4.23 that the probability of selecting the best site should approach 1 as $M \to \infty$. If a result like this is true, it will only be for values of $\mathbb{E}(D_{off})$ that are large enough to ensure that the total number of dancing bees grows without bound. The condition $(1 - p_{follow})\mathbb{E}(D_{off}) > 1$ certainly guarantees this, since the initial scouts at each time step each initiate a process that grows at least as fast as a supercritical GW process. However, the larger the value of $\mathbb{E}(D_{off})$, the more likely it is that an early lead will determine the selection for any threshold. Figure 4.24 shows the effect of increasing the

quorum threshold using the parameters from scenario 1 with $c = 10$, $p_{follow} = 0.8$, $\lambda = 5$, and a maximum threshold of 300. Thirty simulations were run at each
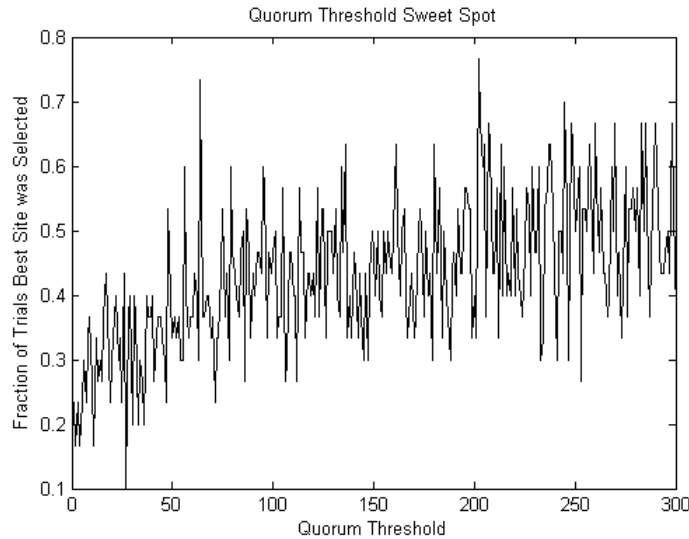
Figure 4.24: Effect of increasing the quorum threshold with $c = 10$, $p_{follow} = 0.8$, $\lambda = 5$.

quorum level. The graph shows a slight improvement in accuracy, on average, for quorum thresholds near the middle of the simulated range, but then it appears to level off. This certainly does not prove that the result cannot hold, but it does suggest that it may only be true for an intermediate range of p_{follow} and $\mathbb{E}(D_{off})$ values.

4.7   Conclusion

This model has largely confirmed the speed versus accuracy trade-off in collective decision-making by the swarm. The single most effective way to improve the

accuracy of swarm decisions is to increase the quorum required to make a decision and thereby increase the time required to make it, all other factors being equal.

5   Conclusion

This thesis studied the branching random walk and tree polymers, and exploited the connections between them to study the implications of disorder type. In chapter two, an infinite volume polymer probability was shown to exist as the weak limit in probability of the sequence of finite volume polymer probabilities in the case of critical strong disorder. This extended results on well-known a.s. infinite volume limits for weak disorder. In chapter three a central limit theorem for the simple branching random walk was shown to hold regardless of the disorder type of its corresponding polymer. A result on the almost sure connectivity of its support was also established, extending known results. The major open problem is to determine the existence of the a.s. weak limit point(s) for the sequence of tree polymer probabilities prob_n under strong disorder.

In chapter four the trade-off between speed and accuracy in optimal decision-making by honeybee swarms was demonstrated with an individual-based stochastic model that incorporates features of standard branching process models. Collective decision-making processes of this kind are by no means limited in either application or general interest to honeybee site selection. The recent work in [34] studies similarites in the key features of cognition in the brains of vertebrates and swarms of honeybees. The work in [31] analyzes several models in this direction, comparing the ability of social insects to make optimal decisions with optimal decision-making

in primate brains. Much work remains to be done in this area.

Bibliography

[1] E. Aidèkon. Convergence in law of the minimum of a branching random walk. `arXiv:1101.1810`, 2011.

[2] E. Aidèkon and Z. Shi. The Seneta-Heyde scaling for the branching random walk. `arXiv:1102.0217v2`, 2012.

[3] M. Beekman, R. L. Fathke, and T. D. Seeley. How does an informed minority of scouts guide a honeybee swarm as it flies to its new home? *Animal Behavior*, 71:161–171, 2006.

[4] S. Bezerra, S. Tindel, and F. Viens. Superdiffusivity for a Brownian polymer in a continuous Gaussian environment. *Ann. Prob.*, 36(5):1642–1675, 2008.

[5] J. D. Biggins. Chernoff's theorem in the branching random walk. *J. Appl. Prob.*, 14:630–636, 1977.

[6] J. D. Biggins. Martingale convergence in the branching random walk. *J. Appl. Probab.*, 14(1):25–37, 1977.

[7] J. D. Biggins. Lindley-type equations in the branching random walk. *Stochastic Process. Appl.*, 75(1):105–133, 1998.

[8] J. D. Biggins. Branching out. In N. H. Bingham and C. M. Goldie, editors, *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*, pages 112–133. Cambridge University Press, 2010.

[9] J. D. Biggins. Spreading speeds in reducible multitype branching walk. `arXiv: 1003:4716v2`, 2011. To appear in Annals of Applied Probability.

[10] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. *Ann. Prob.*, 36:544–581, 2004.

[11] J.D. Biggins. The first-and-last-birth problem for a multitype age-dependent branching process. *Adv. Appl. Probab.*, 8:446–459, 1976.

[12] P. Billingsley. *Convergence of Probability Measures*. Wiley-Interscience, New York, second edition, 1999.

[13] E. Bolthausen. A note on diffusion of directed polymers in a random environment. *Commun. Math. Phys.*, 123:529–534, 1989.

[14] E. Bolthausen. On directed polymers in a random environment. In I.V.Basawa and R.I. Taylor, editors, *Selected Proceedings of the Sheffield Symposium on Applied Probability*, volume 18 of *IMS Lecture Notes Monograph Series*, pages 41–47, 1991.

[15] N.F. Britton et al. Deciding on a new home: how do honeybees agree? *Proc. R. Soc. Lond.*, 269:1383–1388, 2002.

[16] E. Buffet, A. Patrick, and J. V. Pule. Directed polymers on trees: a martingale approach. *J. Phys. A: Math.Gen*, 26:1823–1834, 1993.

[17] S. Camazine, P. K. Visscher, and R. S. Vetter. House-hunting by honey-bee swarms: collective decisions and individual behaviors. *Insectes soc.*, 46:348–360, 1999.

[18] F. M. Dekking and B. Host. Limit distributions for minimal displacement of branching random walks. *Probab. Theory Related Fields*, 90(3):403–426, 1991.

[19] K. Grill. The range of simple branching random walk. *Statistics & Probability Letters*, 26:213–218, 1996.

[20] J. M. Hammersley. Postulates for subadditive processes. *Ann. Prob.*, 2:652–680, 1974.

[21] Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Prob.*, 37(2):742–789, 2009.

[22] P. Jagers. *Branching Processes with Biological Applications.* John Wiley & Sons, 1975.

[23] S. Janson, M. Middendorf, and M. Beekman. Searching for a new home–scouting behavior of honeybee swarms. *Behavioral Ecology*, 18:384–392, 2007.

[24] T. Johnson and E. Waymire. Tree polymers in the infinite volume limit at critical strong disorder. *J. Appl. Prob.*, 48:885–891, 2011.

[25] J-P. Kahane. Random multiplications, random coverings, and multiplicative chaos. In *Proceedings of the Special Year in Modern Analysis*, volume 137 of *London Math. Soc. Lect. Notes*, pages 196–255. Cambridge Univ. Press, London, 1989.

[26] J-P. Kahane and J. Peyrière. Sur certaines martingales de Benoit Mandelbrot. *Adv. in Math.*, 22:131–145, 1976.

[27] J. F. C. Kingman. The first birth problem for age-dependent branching processes. *Ann. Prob.*, 3(5):790–801, 1975.

[28] M. Lindauer. Schwarmbienen auf Wohnungssuche. *Z. Vergl. Physiol.*, 37:263–324, 1955.

[29] M. Lindauer. *Communication among social bees.* Harvard University Press, Cambridge, MA, 1961.

[30] T. Madaule. Convergence in law for the branching random walk seen from its tip. `arXiv:1107:2543`, 2011.

[31] J. Marshall et al. On optimal decision-making in brains and social insect colonies. *J. R. Soc. Interface*, 6:1065–1074, 2009.

[32] M. R. Myerscough. Dancing for a decision: a matrix model for nest-site selection by honeybees. *Proc. R. Soc. Lond.*, 270:577–582, 2003.

[33] K. M. Passino and T. D. Seeley. Modeling and analysis of nest-site selection by honeybee swarms: the speed and accuracy trade-off. *Behav. Ecol. Sociobiol*, 59:427–442, 2006.

[34] K. M. Passino, T. D. Seeley, and P. K. Visscher. Swarm cognition in honey bees. *Behav. Ecol. Sociobiol.*, 62:401–414, 2008.

[35] B. S. Perdriau and M. R. Myerscough. Making good choices with variable information: a stochastic model for nest-site selection by honeybees. *Biol. Lett.*, 3:140–143, 2007.

[36] K. M. Schultz, K. M. Passino, and T. D. Seeley. The mechanism of flight guidance in honeybee swarms: subtle guides or streaker bees? *The Journal of Experimental Biology*, 211:3287–3295, 2008.

[37] T. D. Seeley. Consensus building during nest-site selection in honey bee swarms: the expiration of dissent. *Behav. Ecol. Sociobiol*, 53:417–424, 2003.

[38] T. D. Seeley. *Honeybee Democracy*. Princeton University Press, 2010.

[39] T. D. Seeley and S. C. Buhrman. Group decision making in swarms of honey bees. *Behav. Ecol. Sociobiol*, 45:19–31, 1999.

[40] T. D. Seeley and P. K. Visscher. Choosing a home: how the scouts in a honey bee swarm perceive the completion of their group decision making. *Behav. Ecol. Sociobiol*, 54:511–520, 2003.

[41] T. D. Seeley and P. K. Visscher. Quorum sensing during nest-site selection by honeybee swarms. *Behav. Ecol. Sociobiol*, 56:594–601, 2004.

[42] T. D. Seeley, P. K. Visscher, and K. M. Passino. Group decision making in honey bee swarms. *American Scientist*, 94:220–229, 2006.

[43] E. Waymire and S.C. Williams. A general decomposition theory for random cascades. *Bull. Amer. Math. Soc.*, 31(2):216–222, 1994.

[44] E. Waymire and S.C. Williams. Multiplicative cascades: dimension spectra and dependence. *J. Four. Anal. and Appl., special issue to honor J-P. Kahane*, pages 589–609, 1995.

[45] E. Waymire and S.C. Williams. A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math Soc.*, 348(2):585–632, 1996.

[46] E. Waymire and S.C. Williams. T-martingales, size-biasing and tree polymer cascades. In *Recent Developments in Fractals and Related Fields*, pages 353–380. Birkhauser, Boston, 2010.

APPENDIX

A Code for Swarm Simulation

This is the code for running a single trial of the site-selection model in chapter 4. Other files were also used to generate the results, but they were mainly auxiliary. The simulation was written in MATLAB v7.7.0.

```matlab
function [N,B]=bee_democracy2bp(
n,q,Q_0,M,m,wc,l,cutoff,type,parameter,follow)
% Branching version: bees are either initial explorers, or
% some number of bees, determined by the offspring distribution,
% follow a bee currently dancing for a given site, OR dancing
% bees, rather than branching, may at random choose another
% bee to follow (according to proportional distribution) according
% to the follow parameter (a probability)
% the parameter l is no longer used in this version: set l=0
N=zeros(1,n+1);             % initial support matrix
                           % (dist. of *dancing* bees)
B=[0,0,0,0];               % initial bee data matrix
                           % (initialize first row for while loop)
t_0=[0];                   % initialize first-departure tracker
t=0;                       % initialize time step
c=0;                       % initialize bee counter
```

```
% infinte swarm for this branching version
while (t < cutoff) && (sum(sum([N(:,2:n+1)-M >= 0]))==0)
% calculate random distribtion
N(t+1,1)=sum([B(:,2)==0])-1;
    for i=1:n
      % current population for site i
  N(t+1,i+1)=sum((B(:,2).*(B(:,4) > 0)==i));
    end
    % get random distribution
  DIST=random_dance_distribution(n,B,Q_0);
  % dancing bees branch or follow another bee
% for loop does not run in initial iteration of while loop
[rows,cols]=size(B);
  for k=1:rows-1
siteSum=sum(N(t+1,2:n+1));
if siteSum - 1 > 0
if rand < l
                B(k+1,:)=[k+1,0,t,0];
    % lost bees at site 0, 0 waggles
        % scouts follow a dancing bees with prob follow
      else
        if rand < follow
          i=chooseSite(DIST);
```

```
w=waggle(q,i,m,t,t_0(k+1));

B(k+1,2)=i;

B(k+1,4)=w;
      else
        % with prob 1-follow

        % dancing bee recruits followers

        i=B(k+1,2);

        w=waggle(q,i,m,t,t_0(k+1));

        B(k+1,4)=w;

        if w>0

          b=offspring(type,parameter);
    % number of followers
          for x=1:b

            c=c+1; %increment bee counter

            i=B(k+1,2);

            w=waggle(q,i,m,t,t);

            B(c+1,:)=[c i t w];

            t_0(c+1)=t;

          end

        end

      end

    end

  end
```

```
    end

% dispatch water-cooler scouts

for j=1:wc

c=c+1;                    % increment bee counter

i=chooseSite(Q_0);        % pick a random site according to Q_0

w=waggle(q,i,m,t,t);      % get initial waggle-number for site

B(c+1,:)=[c i t w];       % populate a row with a new bee's data

    t_0(c+1)=t;

  end


t=t+1; % increment time step

end      % end for main while loop

end      % end function

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function w=waggle(q,i,m,t,t_0)

w=max(q(1,i+1)-m*(t-t_0),0);

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function DIST=random_dance_distribution(n,B,Q_0)

  for i=1:n    % current population for site i

  D(i)=sum((B(:,2)==i).*B(:,4));

  end

  danceSum=sum(D);
```

```
    if danceSum>0

      DIST=[0,D]/danceSum;

    else

    DIST=Q_0;

    end

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function s=chooseSite(distribution)

  cdist=cumsum(distribution);

  findindex=cumsum((rand < cdist));

  s=find((findindex==1))-1;

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function b=offspring(type,parameter)

  switch type

    case 'Poisson',

      b=poissrnd(parameter);

    otherwise,

      b=1;

  end

end
```