NUMERICAL ESTIMATES FOR ARM EXPONENTS AND THE ACCEPTANCE PROFILE IN TWO-DIMENSIONAL INVASION PERCOLATION

A Dissertation
Presented to
The Academic Faculty

By

Jiaheng Li

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Mathematics in the
School of Mathematics

Georgia Institute of Technology

May 2020

NUMERICAL ESTIMATES FOR ARM EXPONENTS AND THE ACCEPTANCE PROFILE IN TWO-DIMENSIONAL INVASION PERCOLATION

Approved by:

Dr. Michael Damron, Advisor
School of Mathematics
Georgia Institute of Technology

Dr. Christian Houdré
School of Mathematics
Georgia Institute of Technology

Dr. Konstantin Tikhomirov
School of Mathematics
Georgia Institute of Technology

Date Approved: April 23, 2020

ACKNOWLEDGEMENTS

This thesis and the research behind it would not have been possible without the support of my advisor, Dr. Michael Damron. His enthusiasm, knowledge and exacting attention to detail have been an inspiration and kept my work on track.

This research was supported in part through research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology, Atlanta, Georgia, USA.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

The main object of this thesis is to numerically estimate some conjectured arm exponents when there exist a number of open paths and closed dual paths that extend to the boundary of different sizes of boxes centering at the origin in bond invasion percolation on a plane square lattice by Monte-Carlo simulations. The results turn out to be supportive of the conjectured value in some case. The numerical estimate for the acceptance profile of invasion percolation at the critical point is also obtained, which suggests a neighborhood in which the $\liminf$ and $\limsup$ of the acceptance profile might fall. An efficient algorithm to simulate invasion percolation and to find disjoint paths on most regular 2-dimensional lattices are discussed.

CHAPTER 1

INTRODUCTION

The standard percolation theory was first developed from some physical phenomena in which a fluid spreads randomly through a porous medium [1]. Different from the diffusion theory, percolation theory ascribes the randomness to the medium rather than the fluid and it has been shown to have application to a broad variety of physical and chemical problems [2]. Examples of percolation usually involve fluid transport in random media. Imagine immersing a porous rock in water. The water starts permeating the rock from pores on its surface to the center through narrower throats. The permeating water forms different paths of connected pores and throats in the process of making its way from the surface to the center. What is the probability that the center of the rock is wetted? If we consider an idealized medium in two dimensions, this process can be modelled in a plane square lattice in $\mathbb{Z}^2$. By examining each edge of $\mathbb{Z}^2$, we declare an edge to be *open* with *edge probability* $p \in [0, 1]$ and closed with probability $1 - p$, independent of all other edges. These edges represent the inner throats of the rock and the edge probability p represents the proportion of throats that are broad enough to allow water to pass through them. A vertex x inside the rock is wetted by the water permeating from the surface if and only if there exists a path of connected open edges in $\mathbb{Z}^2$ from x to some vertex on the boundary. Percolation theory is concerned primarily with the existence and properties of such *open paths*.

In 1983, Wilkinson and Willemsen [3] gave birth to invasion percolation, a new form of percolation theory. It was motivated by the study of the displacement process of one fluid by another in a random porous medium. Different from standard percolation, there are invader and defender in invasion percolation. The invading fluid starts from some compact region under the surface, advancing the interface through the throat of least resistance, as opposed to advancing all interfaces up to some fixed threshold in standard percolation. We assume the defender always has an escape route to exit the system. For a plane square lattice in $\mathbb{Z}^2$, each edge is assigned a weight generated from some distribution to represent the resistance encountered by the invader when it tries to replace the defender from a throat. We say an edge e is *incident* to a vertex if this vertex is an endvertex of e. Suppose the invader starts from a single vertex x on the surface and tries to invade into the center. Initially all vertices and edges are set to non-invaded. It first invades the edge with the smallest weight among all edges incident to x. We declare the invaded edge open and mark the non-invaded endvertex of this edge as invaded. Non-invaded edges are considered closed. The invader grows at each time by invading one non-invaded edge with the smallest weight among all edges that are incident to the set of invaded vertices. Such edges that are incident to invaded vertices are said to be *observed*. Paths of open or closed edges will form as the invading fluid invades from the surface to the center.

Consider a subgraph of $\mathbb{Z}^2$ containing the origin, there may exist some disjoint paths of open edges that extend to the complement of the subgraph and intersect only at the origin. As the subgraph increases to infinity, the probability that there exist such disjoint open paths decays polynomially with an exponent. The corresponding exponents are called *arm exponents* and exact values for some of them have not been rigorously derived yet. The main object of this thesis is to numerically estimate conjectured arm exponents under some conditions in invasion percolation by Monte-Carlo simulations and the results turn out to be supportive of some conjectured values. Within the same setting, simulation results also provide us a numerical estimate for the *acceptance profile* of invasion percolation, where no predictions have been made to the best of our knowledge. An efficient algorithm to simulate invasion percolation and to find disjoint paths by Ford Fulkerson Algorithm on most regular 2-dimensional lattices are also discussed.

Chapter 2 establishes the definitions and notations for standard percolation as well as invasion percolation. Some important facts and properties are introduced and discussed. Chapter 3 elaborates how the Monte-Carlo simulation and algorithms work for invasion percolation on 2-dimensional regular lattices. Chapter 4 presents the simulation results and explains how the data should be interpreted. Chapter 5 summarizes the results and concludes the thesis.

CHAPTER 2

BACKGROUND

In this chapter, we shall establish the basic definitions and notations for both standard and invasion percolation models on $\mathbb{Z}^d$.

2.1 Basic Settings

For $x, y \in \mathbb{Z}^d$, we define the distance $\delta(x, y)$ between x and y as

$$\delta(x, y) = \Sigma_{i=1}^{d} |x_i - y_i|,$$

where x_i, y_i are the ith integer coordinate of x and y respectively. We add edges between all pairs $x, y \in \mathbb{Z}^d$ where $\delta(x, y) = 1$ and we say such pair of x and y are adjacent. We represent the edge e from x to y as $\{x, y\}$. $\mathbb{Z}^d$ is the set of vertices and we denote the edges by $\mathbb{E}^d$. Then we write the d-dimensional cubic lattice $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$ and denote the origin of $\mathbb{Z}^d$ by 0.

Next we introduce probability to the percolation model. By examining every edge in $\mathbb{L}^d$, we declare an edge open with $edge\,probability$ $p \in [0, 1]$ and closed with probability $1 - p$, independent by all other edges. the resulting model is called *bond percolation*. An alternative way of introducing randomness is to block vertices rather than edges with probability p, which is referred to as *site percolation*. Without loss of interest, we shall restrict our work to bond percolation.

A $path$ in $\mathbb{L}^d$ is a sequence $x_0, e_0, x_1, e_1, ..., e_{n-1}, x_n$ of distinct vertices x_i and edges $e_i = \{x_i, x_{i+1}\}$. This path has length n and is said to connect x_0 to x_n. A *circuit* of $\mathbb{L}^d$ is a path with $x_n = x_0$. Two paths of $\mathbb{L}^d$ are called disjoint if they have no edges in common. A path is open if all of its edges are open, and closed if all of its edges are closed. If A and B are two sets of vertices in $\mathbb{Z}^d$, we write $A \leftrightarrow B$ if they are connected by an open path from some vertex in A to some vertex in B. The *surface* (or *boundary* in two dimensions) ∂A of A is the set of vertices in A which are adjacent to some vertex not in A. We define by $B(n)$ a *box* with side length $2n$ and center at the origin, i.e.,

$$B(n) = [-n, n]^d.$$

On immersion of the rock in water, it is reasonable to assume that the water passageways inside the rock are of a negligible scale when compared with the overall size of the rock. In the case of a plane square lattice $\mathbb{L}^2$, this indicates the probability that the center of the rock is wetted by water permeating from the surface behaves similarly to the probability that the origin of $\mathbb{Z}^2$ is an endvertex of an infinite open path in $\mathbb{L}^2$.

2.2 Infinite Open Clusters

Consider the lattice $\mathbb{L}^d$ with the vertex set $\mathbb{Z}^d$ and the open edges only. The connected vertices together with the open edges that connect them are called open clusters. We write $C(x)$ for the open cluster containing the vertex x and we call $C(x)$ the open cluster at x. The number of vertices in $C(x)$ is denoted by $|C(x)|$. Typically, we represent the open cluster $C(0)$ at the origin simply as C. Then the large-scale permeation to the center of the rock by water is related to the existence of infinite open clusters at the origin. A fundamental quantity related to the existence of such infinite open clusters is the *percolation probability* $\theta(p)$, which is the probability that a given vertex belongs to an infinite open cluster. By the translation invariance of the lattice and the probability measure given by the edge probability p, we lose no generality by taking this vertex as the origin and define

$$\theta(p) = \mathbb{P}_p(|C| = \infty).$$

The function is a non-decreasing function of p with $\theta(0) = 0$ and $\theta(1) = 1$. A fundamental fact of percolation theory is that there exists a critical value $p_c = p_c(d)$ of the edge probability p such that

$$\theta(p) \begin{cases} = 0 & \text{if } p < p_c \\ > 0 & \text{if } p > p_c. \end{cases}$$

$p_c(d)$ is called the *critical probability* and is formally defined as

$$p_c(d) = \sup\{p : \theta(p) = 0\}.$$

It has been proved that $0 < p_c < 1$ when $d \geq 2$ and the standard percolation at $p = p_c$ is referred to as *critical percolation*. A useful representation of $p_c(d)$ for each value of d has not been found yet but it's not hard to see $p_c(1) = 1$. A nontrivial proof showed that the critical probability of bond percolation on $\mathbb{L}^2$ is $1/2$ [4]. This shall be enough for the purpose of this thesis. With basic features of critical percolation introduced, we now state a theorem about the probability $\psi(p)$ that there exists an infinite open cluster:

$$\psi(p) = \begin{cases} 0 & \text{if } \theta(p) = 0 \\ 1 & \text{if } \theta(p) > 0. \end{cases}$$

2.3 Critical Phenomenon

Studying the sizes and shapes of open clusters are interesting questions of percolation theory and associated quantities $\theta(p)$ and $\psi(p)$ are clearly dependent on p. As p varies from 0 to 1, the standard percolation exhibits a phase transition in terms of the existence of infinite open clusters at or near the critical point p_c. When $p < p_c$, the model is in a subcritical phase and all open clusters are finite almost surely. It has been found that the tail of $|C|$ decreases exponentially as the size of the cluster increases. That is to say, there exists $\alpha(p) > 0$ such that for $p < p_c$,

$$\mathbb{P}_p(|C| = n) = e^{-n\alpha(p)} \qquad \text{as } n \to \infty.$$

When $p > p_c$, the model is in a supercritical phase and there exists a unique infinite open cluster almost surely.

The behavior of percolation away from the critical point is well understood, major open questions occur at and near p_c. When $p = p_c$, it is known that there does not exist infinite open clusters for $d = 2$ or $d \geq 11$, and is generally believed that this is the case for all $d \geq 2$ [5, 6]. Then at what rate does $P_{p_c}(|C| = n)$ decay? For large enough d (currently $d \geq 11$ suffices) or when $d = 2$ on triangular lattice in site percolation, it is found that there exists $\delta = \delta(d) > 0$ such that

$$\mathbb{P}_{p_c}(|C| \geq n) = n^{-1/\delta} \qquad \text{as } n \to \infty.$$

When p is near p_c and approaches p_c from above (or below), $\theta(p)$ behaves as a power of $|p - p_c|$:

$$\theta(p) = |p - p_c|^{\beta + o(1)},$$

where

$$\beta = \lim_{p \downarrow p_c} \frac{\log \theta(p)}{\log(p - p_c)},$$

exists. Above β and δ are called *critical exponents*.

2.4 Arm Events

We now introduce *arm events* for critical percolation on $\mathbb{L}^2$ at $p = p_c = 1/2$. Let $B(n)$ be a box with side length $2n$ and center at the origin. We may observe connections from the origin to the boundary $\partial B(n)$ of $B(n)$ by open and closed paths, or *arms*. We define a dual lattice $\mathbb{L}_*^2 = (\mathbb{Z}_*^2, \mathbb{E}_*^2)$, where

$$\mathbb{Z}_*^2 = \mathbb{Z}^2 + (1/2, 1/2),\ \mathbb{E}_*^2 = \mathbb{E}^2 + (1/2, 1/2).$$

A dual edge e^* bisects exactly one edge $e \in \mathbb{E}^2$ and has the same open-closed state as e. A *closed dual path* is a sequence of closed edges in $\mathbb{E}^2$ whose dual edges form a closed path on the dual lattice $\mathbb{L}_*^2$. We define a color sequence $\sigma = (\sigma_1, ..., \sigma_j)$, where $j \geq 1$ is an integer and each entry represents a symbol O for open paths or C for closed dual paths. A $(j, \sigma)-$ *arm configuration* in $B(n)$ is the data of j disjoint, nonself-intersecting monochromatic arms $\{r_i\}_{1\leq i\leq j}$ connecting the origin and $\partial B(n)$, ordered counterclockwise in a cyclic way. The color of the arm r_i is given by σ_i. We define the corresponding *arm event* by

$$A_{j,\sigma}(n) := \{0 \leftrightarrow \partial B(n) \text{ with arm configuration } (j, \sigma)\}.$$

$A_{j,\sigma}(n)$ only depends on the state of the edges in $B(n)$ and its probability decays polynomially [7]. When $j = 1$,

$$\mathbb{P}(A_{1,\sigma}(n)) = n^{-\frac{5}{48}+o(1)} \qquad \text{as } n \to \infty.$$

Fix $j \geq 2$. For any color sequence σ containing both colors O and C [8],

$$\mathbb{P}(A_{j,\sigma}(n)) = n^{-\alpha_j+o(1)} \qquad \text{as } n \to \infty,$$

where $\alpha_j = \frac{j^2-1}{12}$. The *Arm exponent* α_j describes the probability of observing connections between the origin and the boundary of a large box by a certain number of disjoint open paths and closed dual paths.

Arm events in invasion percolation have also been observed and studied on regular 2-dimensional lattices recently [9]. In the square lattice case with $\mathbb{L}^2 = (\mathbb{Z}^2, \mathbb{E}^2)$, each edge $e \in \mathbb{E}^2$ is assigned a weight w_e independently drawn from a uniform distribution on the interval [0,1]. In the initial configuration, all edges and vertices are non-invaded (closed). Suppose the invader starts from the origin and we mark the origin from non-invaded to invaded. It grows at each time step by invading an edge which has the smallest weight among all that are incident to the set of invaded vertices. This newly invaded edge and its endvertices are then set to be invaded (open). We assume the defender always has an escape route to some sink on the boundary. Paths and open clusters are similarly defined as critical percolation. The color sequence $\sigma = (\sigma_1, ..., \sigma_j)$ for invasion percolation consists of open or closed entries $\{\sigma_i\}_{1 \le i \le j}$. While open entries are open paths in $\mathbb{L}^2$, closed entries are closed dual paths in $\mathbb{L}^2_*$. We write $|\sigma|_O$ and $|\sigma|_C$ for the number of open and closed entries respectively. Then the arm configuration $(j, \sigma)_{ip}$ and arm events $A^{ip}_{j,\sigma}(n)$ for invasion percolation are defined similarly as for critical percolation.

The probability of an arm event is called an *arm probability*. We now state three cases of arm probabilities for invasion percolation with different color sequences. We set $k = |\sigma|_O + |\sigma|_C$.

Case 1: For any σ with $|\sigma|_O \ge 2$ and $|\sigma|_C \ge 1$,

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = n^{-\frac{k^2-1}{12}+o(1)} \qquad \text{as } n \to \infty.$$

Case 2: For any σ with $|\sigma|_O \ge 2$ and $|\sigma|_C = 0$,

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = n^{-\frac{k^2-1}{12}-\frac{5}{48}+o(1)} \qquad \text{as } n \to \infty.$$

Case 3: For any σ with $|\sigma|_O = 1$ and $|\sigma|_C \ge 1$,

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = n^{-\frac{k^2-1}{12}+\frac{5}{48}+o(1)} \qquad \text{as } n \to \infty,$$

when $k = 2, 3$. The exponent in the first case has been proved to exist and its value is known on the triangular lattice; the second exponent is known to exist with a conjectured value. The third exponent has not been proved to exist but a lower bound and an upper bound have been found. If we compare arm probabilities between critical and invasion percolation, they are consistent for the color sequence in case 1. However arm probabilities are different by a power of n in the other two cases. Another difference between the two percolation models can be found in the *acceptance profile* $\alpha_n(p)$ at value p and time n, defined by

$$\alpha_n(p) = \frac{\text{expected number of invaded edges with weight in } [p, p+dp]}{\text{expected number of observed edges with weight in } [p, p+dp]}.$$

As $n \to \infty$, $\alpha_n(p)$ converges to 1 for $p < p_c$ and to 0 for $p > p_c$ in critical percolation on $\mathbb{L}^2$. In invasion percolation, the existence of $\lim_{n\to\infty} \alpha_n(p)$ is not clear. However, the $\liminf_{n\to\infty} \alpha_n(p)$ and $\limsup_{n\to\infty} \alpha_n(p)$ are both in $(0,1)$, away from 0 and 1 when $p = p_c = \frac{1}{2}$ [10].

We simulated the arm exponents of case 2 and case 3 in invasion percolation by expanding the smaller box $B(n)$. The result for case 2 turns out to support the conjectured value while the result for case 3 confirms the existence of the exponent. With the same model, we found that the acceptance profile at $p = \frac{1}{2}$ for invasion percolation falls into a neighborhood of a positive number away from 0 and 1 when $n \to \infty$. The details of simulation will be introduced in chapter 3.

CHAPTER 3

SIMULATION

This chapter discusses the computer experiments and algorithms we used to simulate invasion percolation and to find disjoint paths.

3.1 Invading Algorithms

We use integer coordinates for both $\mathbb{E}^2$ and $\mathbb{Z}^2$ to build the square lattice $\mathbb{L}^2$. Each edge or vertex is assigned a different coordinate and the relative positions among edges and vertices are clear due to the fine structure of the lattice. The invasion percolation process proceeds as described in chapter 2 and stops when a vertex on the boundary of a large box $B(N)$ is invaded. We set the side length of this box as $2N = 2000$. While the invader grows and more edges become observed, finding the smallest weight from a large set of random numbers becomes time-consuming. Min-heap is an efficient data structure to find the smallest element. We push all randomly generated weights into the min-heap and the smallest weight is always sorted to be the first element. Therefore choosing the next target for the invader only costs $O(1)$. In the process of invading, a weight is generated and inserted into the heap once a new edge is observed, and a weight needs to be removed if the corresponding observed edge becomes invaded because no edges will be invaded twice. Since insertion and removal may change the smallest element in the heap, the heap will sort all elements again to find the smallest one after each insertion/removal. Let E be the total number of observed edges after the invader touches the boundary of the large box and stops. Because the heap is a tree-based data structure, inserting into and removing from the heap both cost up to $O(logE)$. As a result, the overall complexity of the whole invasion percolation process is up to $O(E * logE)$. Moreover, If N is the side length of the large box that stops the invasion, $E = N^{\frac{91}{48}+o(1)}$.

3.2 Flow Network

Once the invading process stops, all open paths are formed and we can start counting the number of disjoint open paths within the box $B(n)$ that extend to the complement of $B(n)$. We introduce a flow network to help calculate the number of such disjoint paths. If we consider $\mathbb{L}^2$ as a directed graph and let s be a *source* and t be a *sink* such that s passes $flows$ through the network to t. The capacity for an edge $\{x, y\}$ is denoted by c_{xy} where $x, y \in \mathbb{Z}^2$. A *flow* is a mapping from an edge $\{x, y\}$ to a positive real number $f_{xy} : \mathbb{E}^2 \to \mathbb{R}^+$ satisfying two constraints:

1. For every edge $\{x, y\} \in \mathbb{E}^2,\ f_{xy} \leq c_{xy}$.
2. For each vertex y apart from source s and sink t, the equality

$$\Sigma_{\{x:\{x,y\}\in\mathbb{E}^2} f_{xy} = \Sigma_{\{z:\{y,z\}\in\mathbb{E}^2} f_{yz}$$

holds. The first constraint says the volume flowing through each edge cannot exceed the edge capacity, and the second means the volumes flowing into each vertex equal the volumes flowing out of each vertex except for the source and sink where only outgoing and incoming flows exist. The *value* $|f|$ of the flow in the network is defined by the total flow volumes sending out by the source or the total volumes exiting the network via the sink, i.e., $|f| = \Sigma_{\{x:\{s,x\}\in\mathbb{E}^2\}} f_{sx} = \Sigma_{\{y:\{y,t\}\in\mathbb{E}^2\}} f_{yt}$, where s and t are the source and sink respectively and x, y represent any vertex connected to s or t. The *maximum flow problem* is a problem of maximizing $|f|$, that is to say, it's a problem of sending as much flow as possible from s to t.

Let S and T be two disjoint subsets of $\mathbb{Z}^2$. We define an *s-t cut*, or *cut* $C = (S, T)$ for s and t to be a partition of $\mathbb{L}^2$ such that $s \in S$ and $t \in T$. The s-t cut divides the vertices of $\mathbb{L}^2$ into two parts with the source s in one part and the sink t in the other. The *cut-set* X_C of a cut C is the set of edges that connect the source part and the sink part. If all edges of X_C are removed, the source part and the sink part are disconnected and the value of flow $|f| = 0$. The capacity of a cut C for s and t is the total capacity of all edges in X_C. There may exist different cuts in a network and the cut with the smallest capacity is called the *minimum cut*. A *minimum cut problem* is a problem of finding the minimum cut by selecting different S and T in the network.

Finding cuts with smaller capacity requires careful work and is generally more difficult. However, we could find the minimum cut by calculating the maximum value of flows from s to t.

Minimum-Cut-Max-Flow Theorem. *The maximum value of flow in the network is equal to the capacity of the minimum s-t cut.*

In invasion percolation, if an edge is invaded, we set the capacity of this edge to 1, otherwise 0. Then Menger's theorem helps us find the number of disjoint open paths connecting the origin to the boundary of a box by calculating the max flow.

Menger's Theorem. *Let G be a finite undirected graph and x, y be two distinct vertices. The capacity of the minimum cut for x and y is equal to the maximum number of edge-disjoint paths from x to y.*

In our invasion percolation model, let the origin be the source and the set of invaded vertices on $\partial B(n)$ be sinks. By Menger's theorem, finding the number of disjoint open paths that extend to the complement of $B(n)$ is the same as finding the capacity of the minimum cut for the origin and the set of invaded vertices on $\partial B(n)$ in the network of open edges and invaded vertices within $B(n)$. In order to utilize the minimum-cut-max-flow theorem, we consider all undirected open edges in $B(n)$ as two directed edges, each of which has capacity 1 with an opposite direction. This modification is consistent with invasion percolation and does not affect the number of disjoint open paths we are looking for. Then finding the number of disjoint open paths across $B(n)$ becomes equivalent to finding the maximum flow from the origin to sinks on $\partial B(n)$.

The Ford-Fulkerson method is a greedy method to find the max flow on any flow network. Thanks to the fine structure of the plane square lattice, finding the maximum flow is not as time-consuming as doing that in a completely random graph. We implement this method by the Edmonds-Karp algorithm. With our invading algorithm, there exists at least one open path from the origin to $\partial B(n)$ and Edmonds-Karp algorithm costs $O(E)$ to find one more disjoint path, where E is the number of invaded edges within $B(n)$. Since the invader starts at the origin, there can exist at most four disjoint open paths thus the overall complexity of Edmonds-Karp algorithm finding all possible disjoint open paths is still $O(E)$. Both invading and path finding algorithms would work the same way on other regular lattices including the triangular lattice with a few adjustments.

3.3 Summary

Now that we have successfully counted the number of disjoint open paths across $B(n)$ after the invasion percolation stops; the way of counting closed dual paths is similar but implemented on the dual lattice instead. We summarize the algorithm for simulating invasion percolation and finding the number of such disjoint paths as below:

Step 1. Assign a different coordinate to each edge in $\mathbb{E}^2$ and each vertex in $\mathbb{Z}^2$ based on their positions in $\mathbb{L}^2$. Initially, all edges and vertices are non-invaded and non-observed. If the invader starts from the origin, we mark the origin as invaded.

Step 2. Once a vertex becomes invaded, mark all non-invaded and non-observed edges incident to this vertex as observed. Assign random numbers drawn independently from the uniform distribution on the interval $[0, 1]$ to each of these newly observed edges and insert their weights and associated coordinates into the min-heap.

Step 3. The invader invades the edge with the smallest weight and we change the status of this edge from observed to invaded and mark its non-invaded endvertex as invaded. The weight and coordinate of this newly invaded edge is removed from the min-heap.

Step 4. Since a new invaded vertex appears, repeat step 2 and step 3 till the first vertex on the boundary of the large box $B(N)$ is invaded.

Step 5. Upon stopping of the invasion, build a directed flow network with all open edges and invaded vertices in $B(n)$ based on their positions in $\mathbb{L}^2$ and set the capacity of each edge to 1. Let the origin be the source and vertices on $\partial B(n)$ be sinks.

Step 6. Calculate the maximum flow in this network by Edmonds-Karp algorithm and the result equals the number of disjoint open paths across $B(n)$ by the minimum-cut-max-flow theorem and the Menger's theorem.

3.4 Parameters

The goal of our simulation is to find the arm exponents described in case 2 and 3 in chapter 2. As we are assuming the box $B(n)$ goes to infinity, the larger the size of $B(N)$, the more accurate estimates we are likely to get. Fix the large box $B(N)$ with a side length $2N = 2000$ and we simulate more than 100 data points as the side length n of the box $B(n)$ increases from 10 to 300. Since n is far less than N in some sense, the estimates would be meaningful.

For each data point, we determine the necessary number of samplings by Chernoff bound to control the error rate. Recall that in case 2, the conjectured arm probability is

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = n^{-\frac{k^2-1}{12}-\frac{5}{48}+o(1)} \qquad \text{as } n \to \infty.$$

Let X_i be the ith sampling of the arm probability and M be the total number of samplings for a fixed n. If we denote the arm probability simply by P, we have

$$\mathbb{P}(\Sigma_{i=1}^{M} X_i > (1+\epsilon)PM) < (\frac{e^{\epsilon}}{(1+\epsilon)^{1+\epsilon}})^{PM}$$

, where $\epsilon > 0$ is the ratio by which the sum of samplings may deviate from its expected value. The other direction is similarly given by

$$\mathbb{P}(\Sigma_{i=1}^{M} X_i < (1-\epsilon)PM) < (\frac{e^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}})^{PM}.$$

We take $\epsilon = 0.1$ and set the right side to 0.1 too. As n increases, the arm probability decreases and more samplings are required to keep the same error rate. For $2n = 200$, the enough number of samplings we need to make is 3112 in case 2. Similar arguments can be made to case 3 and the largest number of samplings required is 1031 for $2n = 200$ to keep the same error rate. Overall we make 4000 samplings for each data point through our simulation. The acceptance profile is automatically obtained during the process of invading. We present all simulation results in the next chapter.

CHAPTER 4

RESULTS

We simulated invasion percolation in boxes of different sizes in the square lattice $\mathbb{L}^2$ and counted the number of disjoint arms connecting the origin to the boundary of the box $B(n)$. Each simulated data point is an estimate for the arm probability in a box of side length $2n$. For case 2, we sampled 146 data points where n ranges from 5 to 150, i.e., the side length of $B(n)$ ranges from 10 to 300. For case 3, we sampled 131 data points as n increases from 5 to 135, i.e., the side length of $B(n)$ ranges from 10 to 270. We recorded the values of observed weights and invaded weights during the invading process and obtained 46 data points at different dp of the acceptance profile for invasion percolation as dp approaches 0 when $p = p_c = 1/2$.

4.1 Arm Exponents in Case Two

Recall that the arm exponents in case 2 have been proved to exist and for any color sequence σ with $|\sigma|_O \geq 2$ and $|\sigma|_C = 0$, the value of the exponent has been conjectured as

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = n^{-\frac{k^2-1}{12}-\frac{5}{48}+o(1)} \qquad \text{as } n \to \infty,$$

where $k = 2$ in this case. Once we obtained the raw data, we fit them by a least-square linear regression in the log-log space [11]

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = \frac{c}{n^{\alpha_k}} + \epsilon.$$

The results are shown in figure 1 and indicate a clear polynomial decay with an exponent -0.37, which is closed to the conjectured exponent $\alpha_k = -\frac{k^2-1}{12} - \frac{5}{48} = -0.35$ (here

$k = 2$) when n is small and far from the boundary of the larger box $2N = 2000$. In figure one, the red line is the fitted line of simulated arm probabilities and the blued line is for the conjectured arm probabilities with an exponent -0.35. Since the box is supposed to grow to infinity, we lose accuracy as n increases and gets closer to N. One solution to improve the accuracy when $B(n)$ becomes large is to increase N, but this increases the simulation time significantly. If we fix N, the larger box $B(N)$ stops the invasion earlier, and fewer edges within the smaller box $B(n)$ get invaded, which possibly reduces the number of additional disjoint open paths. The arm probability then should receive a positive correction for large n, which would be an increasing function of n. However, despite many efforts, we were unable to construct an effective function for all n. Though the simulation result is not exactly the same with the conjectured value, it still suggests that the conjectured exponent when $k = 2$ is true with high probability.

4.2 Arm Exponents in Case Three

In case 3, the arm probability for any σ with $|\sigma|_O = 1$ and $|\sigma|_C \geq 1$ has been conjectured as

$$\mathbb{P}(A^{ip}_{k,\sigma}(n)) = n^{-\frac{k^2-1}{12}+\frac{5}{48}+o(1)} \qquad \text{as } n \to \infty,$$

when $k = 2, 3$. We fit the data with the same regression model in case 2 in the log-log space. It's clear to see the existence of such a polynomial decay in figure two though, the estimate for this exponent given by the simulated data points and the fitted line (red line) is -0.11, indicating a slower arm probability decay than the conjectured speed $-\frac{k^2-1}{12} + \frac{5}{48} = -0.14$ ($k = 2$, blue line) as the smaller box $B(n)$ expands. As opposed to the situation in case 2, the simulated arm probabilities in case 3 are higher than the conjectured ones, even when n is small. It's not surprising and probably because closed dual paths are more likely to exist since edges are open only when they are assigned a small weight; therefore open paths are comparatively hard to form and the earlier stop of the invading process further

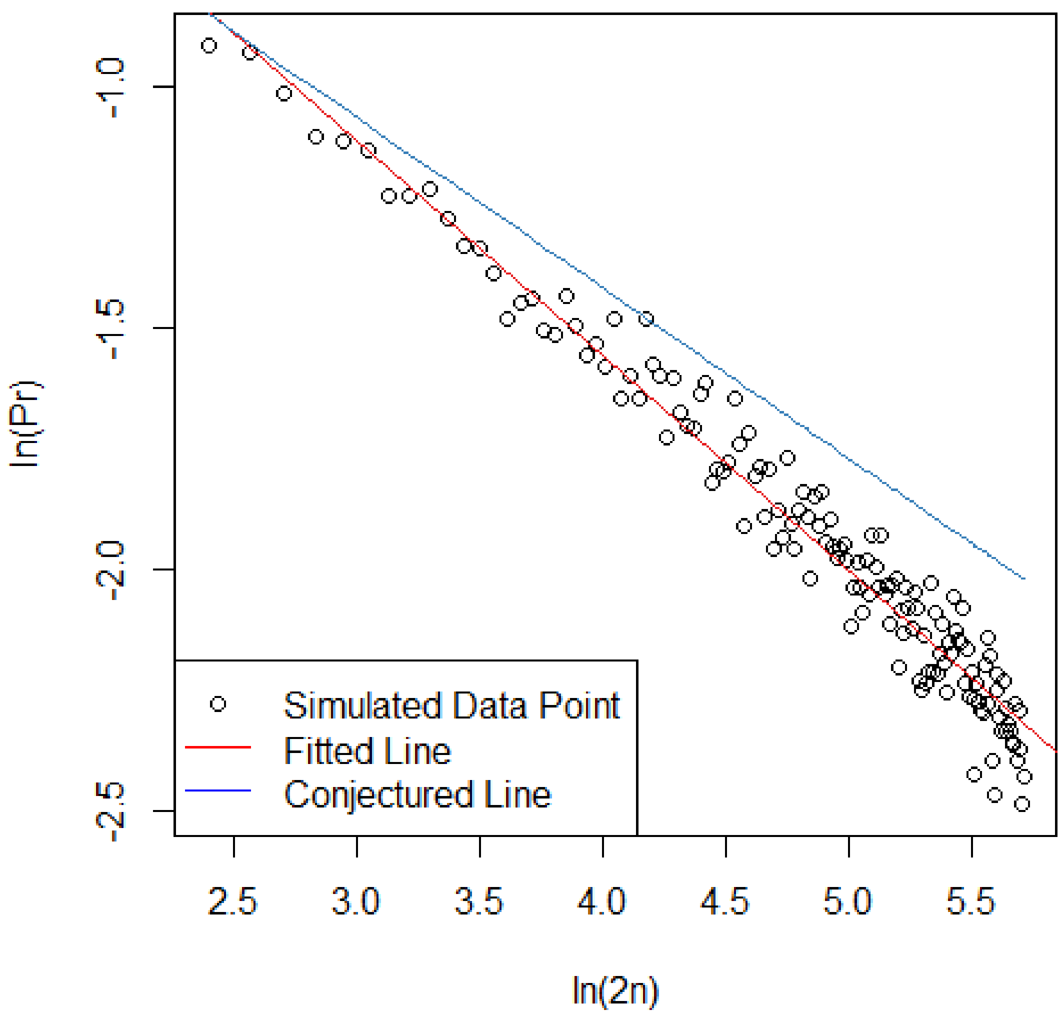

Figure 4.1: Arm Probability in Case 2 against Side Length of $B(n)$ in Log-log Space.

increases the number of closed edges by making fewer edges open (invaded). Instead of finding a negative correction term in the regression, increasing the size of $B(N)$ seems more important in this case.

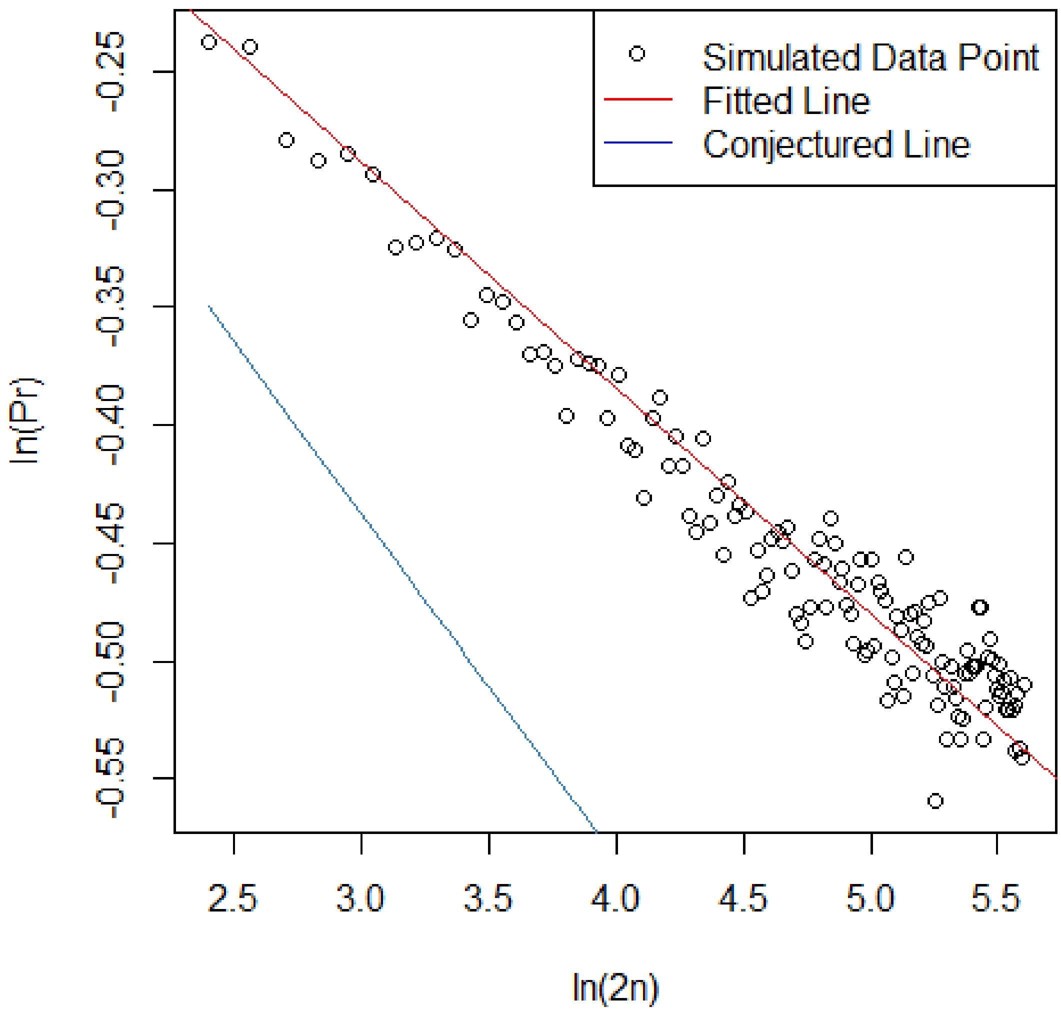

Figure 4.2: Arm Probability in Case 3 against Side Length of $B(n)$ in Log-log Space.

4.3 The Acceptance Profile

The acceptance profile at value p and time n is defined by

$$\alpha_n(p) = \frac{\text{expected number of invaded edges with weight in } [p, p + dp]}{\text{expected number of observed edges with weight in } [p, p + dp]}.$$

In the simulation, we assume $n \to \infty$ and take dp from 10^{-1} to 10^{-6} and equally divide each $[10^{-i+1}, 10^{-i}]$ into 10 intervals for $i = 2, ..., 6$. As a result, 46 data points were simulated and each data point was obtained by taking the average of 4000 samplings at a value of dp. The difference between two neighboring dp's depends on the interval $[10^{-i+1}, 10^{-i}]$ and we let the x-axis be the negative logarithmic value of dp. Then all data points are put in the graph with better distance on x-axis to help observe the trend.

The graph exhibits a jump from 0 to a neighborhood of 0.08 as dp approaches 0. We observe fluctuations within a range and cannot conclude the existence of the limit by the graph, but it can be a reasonable numerical estimate for the $\limsup$ and $\liminf$ of $\alpha_n(p)$ as $n \to \infty$.

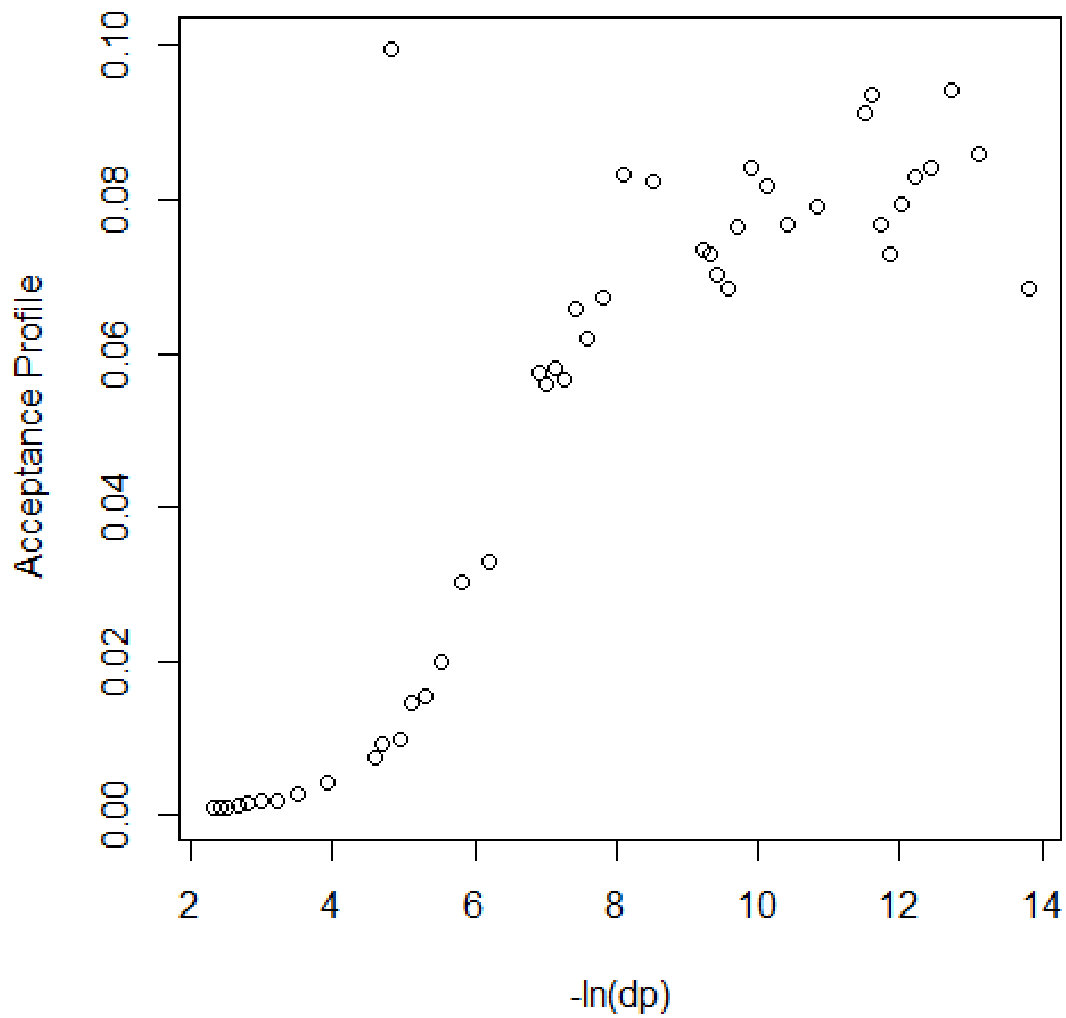

Figure 4.3: The Acceptance Profile against -Ln(dp).

CHAPTER 5

CONCLUSIONS

As simulation results suggest, the conjecture for arm exponents in case 2 has a high probability to be true. The arm probability in case 3 exhibits a clear polynomial decay, but the estimate for the exponent need to be verified by more data points from a larger $B(N)$ before any conclusion can be drawn. Simulated values for the acceptance profile are not strong enough to conclude the existence of $\lim_{n\to\infty}\alpha(p)$, however, they fluctuate in a neighborhood of 0.08 thus provide numerical estimates for the $\limsup_{n\to\infty}\alpha(p)$ and $\liminf_{n\to\infty}\alpha(p)$.

The simulation algorithms discussed in chapter 3 would be easily extended to most regular two-dimensional or three-dimensional lattices with a few modifications and the fine structure of regular lattices guarantees a $O(N * logN)$ complexity for the invading process. It's important to note that all results are subject to the limitation of the simulation scale, including the size of the larger box $B(N)$ and the number of samplings for each data point. Another possible improvement could be made through a regression model with a proper correction term. While Increasing the scale of simulation can be done by utilizing more computing resources, we would gladly hear from any ideas on the data fitting problems.

REFERENCES

[1] S. Broadbent and J. Hammersley, "Percolation process: Crystals and mazes," *Proc. Cambridge Philos. Soc.*, vol. 53, pp. 629–641, 1957.

[2] H. Frisch and J. Hammersley, "Percolation processes and related topics," *J. Soc. INDUST. APPY. MATH.*, vol. 11, pp. 894–918, 1963.

[3] D. Wilkinson and F. Willemsen, "Invasion percolation: A new form of percolation theory," *J. Phys. A: Math. and Gen.*, vol. 16, pp. 3365–3376, 1983.

[4] G. Grimmett, *Percolation*. Springer, 1999.

[5] A. Jarai, "Invasion percolation and the incipient infinite cluster," *Comm. Math. Phsy.*, vol. 236, pp. 311–334, 2003.

[6] H. Kestern, "The incipient infinite cluster in two-dimensional percolation," *Probab. Theory Related Fields*, vol. 73, pp. 369–394, 1986.

[7] S. Smirnov and W. Werner, "Critical exponents for two-dimensional percolation," *Mathematical Research Letters*, vol. 8, pp. 729–744, 2001.

[8] M. Aizenman, B. Duplantier, and A. Aharony, "Path crossing exponents and the external perimeter in two-dimensional percolation," *Comm. Math. Phsy.*, vol. 83, p. 1359, 1999.

[9] M. Damron, J. Hanson, and P Sosoe, "Arm events in two-dimensional invasion percolation," *Journal of Statistical Physics*, vol. 173, pp. 1321–1352, 2018.

[10] B. Bock and M. Damron, *The acceptance profile of invasion percolation at p_c in two dimensions*, 2019. arXiv: `1904.08893 [math.PR]`.

[11] V. Beffara and P. Nolin, "Numerical estimates for monochromatic percolation exponents," 2009.