Fluctuations of Passage Time around Time Constant in First Passage Percolation and Its Simulation

Kaile Ding

October 2020

Abstract

First Passage Percolation is a study on shortest path problems on graphical models. It was first introduced by Hammersley and Welsh in 1965 as a model of fluid flow through a porous media[10]. In this paper, we first introduce the first passage percolation model defined on the square lattice $\mathbb{Z}^2$ and then describe some basic results about the *time constant* and *limit shape*. Then, we focus on simulating the time constant and limit shapes, and discuss the graph algorithm used and theoretical error bounds. Finally, we present the results of our simulation.

1 Introduction

In 1965, Hammersley and Welsh introduced a formulation of first passage percolation as a model of fluid flow through a porous media [10]. This model has a rather simple definition yet incredible connections with several fascinating conjectures. In this paper, we will focus on the square lattice $\mathbb{Z}^2$ for simplicity and mathematical tractability. However, it's worth noting that a number of results are applicable to the general lattice $\mathbb{Z}^d$ too.

Let $E(\mathbb{Z}^2)$ be the collection of nearest-neighbor edge in $\mathbb{Z}^2$. Then, for each edge e in $E(\mathbb{Z}^2)$, assign a non-negative random variable τ_e to it. Note, in this paper, we assume the collection (τ_e) is independent and identically distributed with common distribution F and measure μ. There are some variants of the model that do not assume i.i.d weights, but they are beyond the scope of this paper.

A *path* Γ is a sequence of edges $(e_1, e_2, e_3, ...)$ such that for each $n \geq 1$, e_n and e_{n+1} share exactly one endpoint and all edges (and therefore all vertices) are distinct. For any finite path Γ, its *passage time* or *first passage time* $T(\Gamma)$ is defined as follow:

Definition 1 (Passage Time)**.**

$$T(\Gamma) = \sum_{e \in \Gamma} \tau_e$$

and for two given points x, $y \in \mathbb{Z}^2$

$$T(x, y) = \inf_{\Gamma} T(\Gamma)$$

where the infimum is over all nearest-neighbor lattice paths Γ *from* x *to* y

Then, for each $t \geq 1$, let

$$B(t) = \{y \in \mathbb{Z}^2 : T(0, y) \leq t\}$$

where B could be viewed as the set of all points that have a shortest path with distance smaller or equal to t from the origin. In the case that $F(0) = 0$, we have that $(\mathbb{Z}^2, T(\cdot, \cdot))$ is a metric space and $B(t) \cap \mathbb{Z}^d$ is a random ball with radius t around the origin.

A typical question in first passage percolation is about the behavior of $B(t)$ as $t \to \infty$. This question is closely related to the asymptotic behavior of the passage time $T(0, x)$ as $x \to \infty$ which is the subject of this paper. We first introduce the result observed by Hammersley and Welsh [10] that the expected passage time was *sub-additive* and, therefore, showed the existence of the *time constant*. This result was later strengthened by Kingman [14] by applying his *sub-additive ergodic theorem*.

Later, building on Richardson's *shape theorem* and *the sub-additive ergodic theorem*, Cox and Durrett [5] and Kesten [12] gave an analogue of the law of large numbers for the random ball $B(t)$. It roughly says that there exist a deterministic *limit shape* B_0 such that $\frac{B(t)}{t} \to B_0$. Then, we discuss the fluctuation of this limit shape, which includes the variance and the order of fluctuations. Though there are many open questions related to the fluctuation of the passage time, we first introduce the predictions from physicists' simulation in two dimensions (there are conflicting predictions in higher dimensions), and then demonstrate the rigorous result of the sub-linear variance from Benjamini-Kalai-Schramm [3] and Alexander's contribution in finding the upper bounds on non-random fluctuations [1].

With these theories, we would then demonstrate a simulation to verify the results in limiting shapes and it related fluctuation. Several keys result in graph algorithms that are going to be helpful for the implementation will be introduced, including the Bellman-Ford algorithm and Dijkstra's algorithm.

2 Limit Shape

2.1 Sub-additivity and the time constant

The first step to study the limiting behavior of the passage time was taken by Hammersley and Welsh [10] that they observed the passage time is *sub-additive* by construction.

Definition 2 (Sub-additivity)**.** *A function $h : \mathbb{Z}^2 \to \mathbb{R}$ is sub-additive if*

$$h(x) + h(y) \geq h(x + y) \quad \textit{for all } x, y \in \mathbb{Z}^2$$

Then, we can make the following proposition:

Proposition. *$h(x) = \mathbb{E}T(z, x)$ is sub-additive for any fixed z*

The proof of this proposition follows from part (a) of the proof of Theorem 1 from [2] as we will discuss in this paper. Then, by Fekete's lemma:

Lemma 1 (Fekete's Sub-additive Lemma)**.** *For every sub-additive sequence $\{a_n\}_{n=1}^{\infty}$, the limit $\lim_{n\to\infty} \frac{a_n}{n}$ exosts and is equal to the infimum $\inf \frac{a_n}{n}$. (The limit may be $\pm\infty$)*

we have that $\frac{h(nx)}{n}$ always has a limit and, thus, $\frac{t(0,ne_1)}{n} \to \mu(e_1)$ in probability where $e_1 \in \mathbb{Z}^2$ and $\mu(e_1) = \inf \mathbb{E}t(0, ne_1)/n$.

This result was later strengthened by Kingman to almost sure (a.s) and L^1 convergence by applying his *sub-additive ergodic theorem* [15], and we are describing the time constant via the following theorem:

Theorem 1 (Theorem 2.18 in [11])**.** *Assume that*

$$\mathbb{E}min[t_1, ..., t_{2d}] < \infty$$

where t_i are i.i.d copies of τ_e. Then there exists an constant $\mu(e_1) \in [0, \infty)$ (called the time constant) such that

$$\lim_{n\to\infty} \frac{T(0, ne_1)}{n} = \mu(e_1) = \inf_n \mathbb{E}\frac{T(0, ne_1)}{n} \quad \textit{a.s and in } L^1$$

In plain words, this result shows that for e_1 in the lattice, the passage time $T(0, ne_1)$ grows linearly in n and the growth rate is defined as the *time constant*. It's important to note that e_1 may be replaced by any $x \in \mathbb{R}^d$ by appropriately extending $T(0, x)$ to $\mathbb{R}^d$ (will be discussed on Page 5).

The proof of the existence of the time constant is a classic application of the *sub-additive ergodic theorem*[15] defined as follow:

Theorem 2 (Sub-additive Ergodic Theorem [15])**.** *Let $(X_{m,n})_{0\leq m<n}$ be a family of random variables that satisfies:*

(a) $X_{0,n} \leq X_{0,m} + X_{m,n}$, for all $0 < m < n$.

(b) The distribution of sequence $(X_{m,m+k})_{k\geq 1}$ and $(X_{m+1.m+k+1})_{k\geq 1}$ is the same for all $m \geq 0$.

(c) For each $k \geq 1$, the sequence $(X_{nk,(n+1)k})_{n\geq 0}$ is stationary.

(d) for some finite constant c

Then

$$\lim_{n\to\infty} \frac{X_{0,n}}{n} \text{ exists a.s. and in } L^1$$

Furthermore, if the stationary sequence in (c) is also ergodic, then the limit is constant almost surely and equal to

$$\lim_{n\to\infty} \frac{\mathbb{E}X_{0,n}}{n} = \inf_n \frac{1}{n}\mathbb{E}X_{0,n}$$

We show how Theorem 2 can be used to prove Theorem 1. Note, the proof of sub-additive ergodic theorem can be found in [14], which is beyond the scope of this paper.

Proof of Theorem 1 from [2]. Let $X_{m,n} = T(me_1, ne_1)$ and we show that this $X_{m,n}$ satisfies conditions (a) to (d) in Theorem 2:

For (a), note that a path from 0 to ne_1 does not necessarily go through me_1, while a concatenation of paths from 0 to me_1 and from me_1 to ne_1 gives us a path from 0 to ne_1. Therefore, by the defintion of T, we have

$$T(0, ne_1) \leq T(0, me_1) + T(me_1, ne_1)$$

which satisfies (a) (triangle inequality).

For (b), (c), and ergodicity, since the collection of τ_e is i.i.d., it is invariant under horizontal shifts of $\mathbb{Z}^2$. $\mathbb{E}X_{0,n} > -cn$ holds as τ_e is non-negative for all edges e in $\mathbb{Z}^2$.

As for (d), we have assumed that

$$\mathbb{E}min[t_1, ..., t_{2d}] < \infty$$

in Theorem 1. Since there are $2d$ disjoint deterministic paths Γ_1, ..., Γ_{2d} in $\mathbb{Z}^d$ ($d = 2$ in the square lattice) joining 0 to e_1, we have

$$T(0, e_1) \leq \min\{T(\Gamma_1), ..., T(\Gamma_{2d})\}$$

Then, we can reorder them so that Γ_1 is the path with the largest number of edges and let B be the number of edges in Γ_1. Then

$$\mathbb{P}(T(0, e_1) > s) \leq \prod_{i=1}^{2d} \mathbb{P}(T(\Gamma_i) > s) \leq \mathbb{P}(T(\Gamma_1) > s)^{2d}$$

and

$$\mathbb{P}(T(\Gamma_1) > s) \leq B\mathbb{P}(\tau_e > s/B)$$

Then, by setting $Y = \min\{t_1, ..., t_{2d}\}$ and combining the previous inequalities, we have

$$\mathbb{P}(T(0, e_1) > s) \leq B^{2d}\mathbb{P}(\tau_e > s/B)^{2d} = B^{2d}\mathbb{P}(Y > s/B)$$

which is the desired result. □

Similarly, we can extend the result above to the rational coordinates and define a homogeneous function $\mu : \mathbb{Q}^2 \to \mathbb{R}$ such that, for any $x \in \mathbb{Q}^2$,

$$\lim_{n\to\infty} \frac{T(0, nx)}{n} = \mu(x) \quad \text{a.s. and in } L^1$$

It's easy to see the following properties for μ for $x, y \in \mathbb{Q}^d$ and $c \in \mathbb{Q}$

1. $\mu(x + y) \leq \mu(x) + \mu(y)$
2. $\mu(cx) = |c|\mu(x)$
3. μ is invariant under symmetries of $\mathbb{Z}^d$ that fix the origin
4. μ is uniformly continuous and Lipschitz on bounded subsets of $\mathbb{Q}^d$, so it has a unique continuous extension to $\mathbb{R}^d$

The last term can be done by using the standard real analysis approach: let $(x_n)_n$ be a sequence in $\mathbb{Q}^2$ such that $x_n \to x$ as $n \to \infty$. Then, $\mu(x_n)$ is Cauchy and $\mu(x)$ could be defined as its limit, which extends the time constant to $\mathbb{R}^2$.

It might be noticeable that the result above is an analogue of the central limit theorem and one may presume that

$$\frac{T(0,nx) - n\mu(x)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1)$$

However, as we will discuss in the fluctuation section, a central limit theorem with Gaussian fluctuations does not hold.

2.2 Limit shape B_0

For each $x \in \mathbb{Z}^2$, we now know the average time to go in the direction of $(0,x)$ converges to $\mu(x)$. Then, a natural question was raised by Hammersley and Welsh: "What does a ball of large radius look like?" [10]. In this subsection, we would discuss how the time constant μ describes the random ball $B(t)$ when $t \to \infty$.

Let p_c be the threshold for Bernoulli bond percolation, which is the critical value that large clusters and long-range connectivity first appears (we know that $p_c = 1/2$ for the square lattice from [13]). Then, let $\mathcal{M}$ be the set of Borel probability measures on $[0,\infty)$ satisfying

$$\mathbb{E}\min\{t_1^d, .., t_{2d}^d\} < \infty$$

where t_i are independent copy of τ_e and with

$$F(0) < p_c(d)$$

These two restrictions ensure that: 1. holes do not form inside the growing cluster and 2. there is no infinite cluster through the origin (i.e. $\mu(x) \neq 0$) (see Remark).

Then, we have this important shape theorem proposed by Cox and Durrett in 1981:

Theorem 3 ([5]). *For each $v \in \mathcal{M}$, there exists a deterministic convex, compact set B_v in $\mathbb{R}^d$ such that for each $\epsilon > 0$,*

$$\mathbb{P}\left((1-\epsilon)B_v \subseteq \frac{B(t)}{t} \subseteq (1+\epsilon)B_v \text{ for all large } t\right) = 1$$

Furthermore, B_0 has a non-empty interior and is symmetric about the axes of $\mathbb{R}^d$.

Remark. In the case that $F(0) < p_c(d)$ fails. Kesten proved that

$$\text{For FPP on } \mathbb{Z}^d,\ \mu(e_1) > 0 \text{ if and only if } F(0) < p_c(d)$$

as Theorem 6.1 in [11]. This theorem indicates that, in our case, the time constant $\mu(e_1) = 0$. Therefore, we have the limit shape $B_v = \mathbb{R}^d$ as shown in Theorem 1.10 in [11], which is essentially an infinite cluster through the origin.

To prove Theorem 3, we first need to show that the growth of $B(t)$ in a fixed rational direction is linear. From Theorem 1, we have that, for $x \in \mathbb{Q}$

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{T(0,nx)}{n} \text{ exists}\right) = 1$$

and, therefore,

$$\mathbb{P}\left(\bigcap_{x\in\mathbb{Q}} \left(\lim_{n\to\infty} \frac{T(0,nx)}{n} \text{ exists}\right)\right) = 1$$

This implies that for all ϵ and x, we have

$$\left|\frac{T(0,nx)}{x} - \mu(x)\right| < \epsilon \text{ for large enough } n$$

Then, we make the following claim:

Claim. *Theorem 3 is equivalent as saying*

$$\overline{\lim_{||x||_1\to\infty}} \frac{T(0,x) - \mu(x)}{||x||_1} = 0 \tag{1}$$

Proof. For the forward direction of this claim, first observe that μ is bounded. Since μ is a nondegenerate norm of $\mathbb{R}^d$ and any two norms on a finite-dimensional space are equivalent, there exist constants C_1, C_2 such that

$$C_1||x||_1 \le \mu(x) \le C_2||x||_1 \text{ for all } x \in \mathbb{R}^d \tag{2}$$

Take $B_v = \{x : \mu(x) \le 1\}$.

Suppose (1) is true. Then from (1), we have that for all $\epsilon > 0$, there exist k such that for all $y \in B(t)$ with $||y||_1 > k$.

$$\mu(y) - \epsilon||y||_1 \leq T(0,y) \leq \mu(y) + \epsilon||y||_1$$

Let

$$\sup_{||x||_1 \leq k} |T(0,x) - \mu(x)| = C_3$$

We first show that

$$\frac{B(n)}{n} \subseteq (1+\epsilon)B_v$$

Take $y \in B(n), T(0,y) \leq n$. If $||y||_1 > k$, we have $\mu(y) - \epsilon||y||_1 \leq n$. Then, by (2), we have $||y||_1 \leq \frac{n}{C_1 - \epsilon}$. Thus, by the properties mentioned in Section 2.1,

$$\mu(y/n) = \frac{\mu(y)}{n} \leq 1 + \frac{\epsilon||y||_1}{n} \leq 1 + \epsilon C$$

for some constant C. The last step is true since $||y||_1 \leq \max k, \frac{n}{C_1 - \epsilon}$. Therefore, we get the desired result as

$$\frac{B(n)}{n} \subseteq (1+\epsilon C)B_v = (1+\epsilon')B_v \text{ by taking the first } \epsilon \text{ small}$$

Then, for the other part:

$$(1-\epsilon')B_v \subseteq \frac{B(n)}{n}$$

we take $y \in \mu(y) \leq 1 - \epsilon'$. By (2), we have $||y||_1 \leq \frac{1-\epsilon'}{C_2}$. Then, by (1),

$$|T(0,ny) - n\mu(y)| \leq \epsilon n||y||_1$$

which implies

$$\begin{aligned} T(0,ny) &\leq n(1-\epsilon') + n\mu(y) \\ &= n(1-\epsilon') + \frac{n\epsilon(1-\epsilon')}{C_2} \\ &\leq n - \epsilon' + \frac{\epsilon - \epsilon'\epsilon}{C} \\ &\leq n \end{aligned}$$

Therefore, $ny \in B(n)$, so $y \in \frac{B(n)}{n}$, which is the desired result.

Conversely, we prove the counter-positive of the backward direction. Suppose there exist sequence $(x_n)_n$ in such that

$$\overline{\lim_{||x_n||_1 \to \infty}} \frac{T(0, x_n) - \mu(x_n)}{||x_n||_1} = c_0 > 0$$

Without loss of generality, we choose $(x_n)_n$ so that

$$\frac{x_n}{n} \in (1-\epsilon)B_v \text{ and } T(0, x_n) > n$$

, for all n. Then, we want to show that $(1-\epsilon)B_v \not\subseteq \frac{B(n)}{n}$. Rearrange $(x_n)_n$ in a descending order and find a sub-sequence $(x_{ak})_k$ such that there is a sequence $(b_k)_k$ where $b_k = \frac{1-\epsilon}{\mu(x_{ak})}$ that makes

$$\mu(\frac{x_{ak}}{b_k}) = 1 - \epsilon$$

Then, we have

$$\begin{aligned} T(0, x_{ak}) &\geq \mu(x_{ak}) + \frac{c_0||x_{ak}||_1}{2} \\ &\geq (1+C)\mu(x_{ak}) \text{ for constant } C \\ &= (1+C)(1-\epsilon)b_k \geq b_k \text{ for small enough } \epsilon \end{aligned}$$

with a positive probability for small enough ϵ (this is possible when ak is large). Therefore, for infinitely many k,

$$x_{ak} \notin B(b_k) \Leftrightarrow \frac{x_{ak}}{b_k} \notin \frac{B(b_k)}{b_k}$$

so,

$$(1+\epsilon)B_v \not\subseteq \frac{B(n)}{n}$$

□

Lemma 2 (Difference estimate)**.** *Let $v \in \mathcal{M}$. Then there exists a constant $\kappa < \infty$ such that, for any $x \in \mathbb{Z}^d$*

$$\mathbb{P}\left(\sup_{z \in \mathbb{Z}^d, z \neq x} \frac{T(x,z)}{||x-z||_1} < \kappa \right) > 0 \tag{3}$$

The proof of this Lemma is skipped here and a sketch of it could be found in [2].

Proof for Theorem 3. We call the vertex $x \in \mathbb{Z}^d$ that appears in (3) as a "good" vertex. Then, to take advantage of the Lemma, we first prove the following claim

Claim. *Let $\zeta \in \mathbb{Z}^d \setminus \{0\}$. For a given realization of edge-weights, denote by $(n_k)_k$ the sequence of natural numbers such that $n_k\zeta$ is a good vertex. Then, with probability one, the sequence $(n_k)_k$ is infinite and*

$$\lim_{k\to\infty} \frac{n_{k+1}}{n_k} = 1$$

To show this claim holds, we use the fact that the passage time is ergodic (as previously proven). The ergodic theorem implies that the sequence $(n_k)_k$ is infinite almost surely. Let A_m denote the event that $m\zeta$ is a good vertex, then

$$\frac{k}{n_k} = \frac{1}{n_k} \sum_{i=1}^{n_k} \mathbb{1}_{A_i}$$

where $\mathbb{1}$ stands for the indicator function. Then, by the ergodic theorem, we have that the right-hand side converges to the probability in (3) as n gets large. Suppose the probability it converges to is p. Therefore, by

$$\frac{n_{k+1}}{n_k} = \frac{n_{k+1}}{k+1}\frac{k+1}{k}\frac{k}{n_k} \xrightarrow{n\to\infty} p\frac{1}{p} = 1$$

we get the desired result.

Then, to prove Theorem 3, we construct a contradiction. Suppose Theorem 3 is not true. Then, by our claim above, since the probability is one, there are instance ω in the collection of edge-weight configuration such that

1) for all $\zeta \in \mathbb{Z}^d$, the condition in the claim above holds

and

2) there exist sequence $(x_i)_i$ such that $\lim_{n\to\infty} \dfrac{T(0,x_i) - \mu(x_i)}{||x_i||_1} = \delta > 0$ (4)

Assume that $x_i/||x_i||_1$ converges to some y with $||y||_1 = 1$. This holds because the finite dimensional unit sphere is compact. Then, we have

$$\left| \frac{\mu(x_n)}{||x_n||_1} - \mu(y) \right| < \delta/2 \text{ for large n} \tag{5}$$

Therefore, by (4),

$$|T(0,x_n) - ||x_n||_1 \mu(y)| > \frac{\delta ||x_n||_1}{2}$$

We show the contradiction by showing that the left hand side of (5) is small. We find some $z \in \mathbb{R}^d$ with $||z||_1 = 1$ such that $||z-y||_1 < \delta'$ and $z = x/M$ for some $x \in \mathbb{Z}^d$ and $M \in \mathbb{Z}+$. Then, there must exist a sequence (n_k) such that $n_k M z$ is a good vertex and $\frac{n_{k+1}}{n_k}$ goes to one. For any n, there is a $k\mathbb{Z}+$ such that

$$n_{k+1} M \geq ||x_n||_1 \geq n_k M$$

Let is k be $k(n)$. Fix $K > 0$ such that $n_{k+1} < (1+\delta')n_k$ and

$$\left| \frac{T(0, n_k M z)}{n_k M} - \mu(z) \right| < \delta' \text{ for all } k > K$$

Let n be large enough so that $k(n) > K$. Then, we have

$$\begin{aligned} \left| \frac{T(0,x_n)}{||x_n||_1} - \mu(y) \right| \leq & \left| \frac{T(0,x_n) - T(0,n_k M z)}{||x_n||_1} \right| \\ & + \left| \frac{T(0,n_k M z)}{||x_n||_1} - \frac{T(0,n_k M z)}{n_k M} \right| \\ & + \left| \frac{T(0,n_k M z)}{n_k M} - \mu(z) \right| + |\mu(z) - \mu(y)| \end{aligned}$$

For the first term, because $k(n) > K$, we have

$$\begin{aligned} & n_k M \leq ||x_n||_1 \leq (1+\delta') n_k M \\ & ||n_k M y - n_k M z||_1 \leq \delta' n_k M \\ & ||\frac{x_n}{||x_n||_1} - y||_1 < \delta' \end{aligned} \tag{6}$$

Therefore, $||x_n - n_k M z||_1 \leq 2\delta' ||x_n||_1$. Since $n_k M z$ is a good vertex,

$$|T(0,x_n) - T(0,n_k M z)| \leq \kappa ||x_n - x_k M z||_1 \leq 2\kappa\delta' ||x_n||_1$$

so this term is small.

For the second term, rewrite it as

$$\left| \frac{T(0,n_k M z)}{n_k M} \left(\frac{n_k M}{||x_n||_1} - 1 \right) \right|$$

Then, by (6) we have that the second factor is small. Also, since $k > K$, we have that the first factor is around $\mu(z)$. Therefore, their product is small.

For the third term, it's bounded above by δ' as a direct result from $k > K$.

For the fourth term, if μ is identically zero, then this term is trivially zero. If μ is not identically zero, it's a norm on $\mathbb{R}^d$. Because $||z - y||_1 < \delta'$ and norms on a finite-dimensional vector space are equivalent, the forth term is bounded above by $C\delta'$ where C is a constant, so it's small.

Therefore, we have bounded the left-hand-side of (5) and it's going to zero as δ' goes to zero. Since δ' is chosen arbitrarily, this is a contradiction with our assumption and, therefore, proves the theorem. □

3 Fluctuation

From the previous discussion on time constant, we have that

$$\lim_{n\to\infty} \frac{T(0, nx)}{n} = \mu(x) \quad \text{a.s. and in } L^1$$

which is equivalent as

$$\frac{T(0, x)}{||x||_1} \to \mu(\frac{x}{||x||_1})$$

From this, we can approximate the passage time between 0 and a vertex $x \in \mathbb{Z}^2$ can be almost surely as

$$\begin{aligned} T(0, x) &= \mu(x) + o(||x||_1) \\ &= \mu(x) + o(||x||_1) + \mathbb{E}T(0, x) - \mathbb{E}T(0, x) \end{aligned}$$

which implies

$$o(||x||_1) = T(0, x) - \mathbb{E}T(0, x) + \mathbb{E}T(0, x) - \mu(x)$$

Traditionally, the error term is split into two parts. The first part $(T(0, x) - \mathbb{E}T(0, x))$ is a random fluctuation that can be treated using concentration of measure, and the second part $(\mathbb{E}T(0, x) - \mu(x))$ is a non-random fluctuation that is analyzed using the bound of the first. We will discuss them separately in this paper.

In the physics literature [17, 18], it is expected from simulation that, under some mild moment conditions, both the random and non-random

fluctuations have order x^χ where χ is independent of of the direction of $x \in \mathbb{Z}^d$.

$$\mathrm{Var}(T(0,x)) \sim ||x||_1^{2\chi} \text{ and } \mathbb{E}(T(0,x)) - \mu(x) \sim ||x||_1^{\chi}$$

We call this χ the *fluctuation exponent*. It's also expected that the exponent χ depends only on the dimension d not the distribution F, and the following dependence on d is predicted for χ from simulations:

d	χ
1	1/2
2	1/3
3	?
.	.

For $d = 1$, the passage time $T(0,x)$ is a sum of $||x||_1$, which are i.i.d. random variables. By central limit theorem, we have that $\chi(1) = 1/2$. However, for $d \geq 2$, it's only predicted that $\chi < 1/2$ and $\chi(2) = 1/3$ (there are conflicting predictions for higher dimensions). Though there are many open questions and comments regarding variance bounds and the fluctuation exponent, this section will be devoted to the rigorous results we have in the field.

3.1 Upper bound for random fluctuation

The first important work regarding the upper bound for the variance was done by Kesten in 1993. He introduced the "method of bounded differences" to the first passage percolation model and improved the bounds for χ to:

$$0 \leq \chi(d) \leq 1/2 \text{ for all } d \geq 1 \tag{7}$$

Theorem 4 ([12]). *Assume $\mathbb{E}\tau_e^2 < \infty$, $\mathbb{P}(\tau_e = 0) < p_c(d)$ and that the distribution of τ_e is not concentrated at one point. There exist $C_1, C_2 > 0$ such that for all non-zero $x \in \mathbb{Z}^d$,*

$$C_1 \leq \mathrm{Var}(T(0,x)) < C_2||x||_1$$

Note that we are using the notation $x_+ = max\{0, x\}$. To prove Theorem 4, the Efron-Stein inequality is a very handy tool and its proof can be found in [9].

Theorem 5 (Efron-Stein's inequality [9]). *Let $X_1, X_2, ...$ be independent and let X_i' be an independent copy of X_i for $i \geq 1$. If f is an L^2 function of $(X_1, X_2, ...)$, then we have*

$$\mathrm{Var}(f) \leq \sum_{i=1}^{\infty} \mathbb{E}[(Z_i - Z)_+]^2,$$

where $Z = f(X_1, X_2, ...)$ and

$$Z_i = f(X_1, ..., X_{i-1}, X_i', X_{i+1}, ...)$$

Then, by applying Theorem 5 to the passage time, we have that

$$\mathrm{Var}(T(0,x)) \leq \sum_{i=1}^{\infty} \mathbb{E}[(T_i(0,x) - T(0,x))_+]^2$$

and the rest of the proof for Theorem 4 could be found in [2].

While (7) is the best result we have for the general lattice $\mathbb{Z}^d$, significant improvement has been obtained for the square lattice ($d = 2$). The first proof of sublinear variance for $T(0,x)$ was done by Benjamini et al. in [3] for Bernoulli τ_e.

Theorem 6 (Theorem 1 in [3]). *For τ_e that are Bernoulli: there exist $0 < a < b < \infty$ such that τ_e takes values a or b with probability 1/2, there is a constant $C = C(d, a, b)$ such that for all $x \in \mathbb{Z}^d$, $||x||_1 \geq 2$,*

$$\mathrm{Var}(T(0,x)) \leq C \frac{||x||_1}{log||x||_1}$$

The most recent proof was done by Damron et al. in [6] that applies to all distribution with $2 + log$ moments.

Theorem 7 ([6]). *For $d \geq 2$, suppose $\mathbb{P}(\tau_e = 0) < p_c$ and $\mathbb{E}\tau_e^2(log\tau_e)_+ < \infty$. There exists $C > 0$ such that for all $x \in \mathbb{Z}^d$ with $||x||_1 > 1$,*

$$\mathrm{Var}(T(0,x)) \leq C \frac{||x||_1}{log||x||_1}$$

3.2 Alexander's Theorem for non-random fluctuation

As for the non-random fluctuation, the most significant contribution was contributed by Alexander in his paper [1], and this result concerns the rate

of convergence of a deterministic non-negative sub-additive function h on the lattice $\mathbb{Z}^d$ to it limit g where

$$g(x) := \lim_n \frac{h(nx)}{n}$$

It's easy to see that g is approached from above:

$$g(x) \leq h(x) \text{ for all } x \in \mathbb{Z}^d$$

Let Φ be the set of all positive non-decreasing functions on $(1, \infty)$. Then, fix $a < 1$ and, for $\phi \in \Phi$, let $Q_x(a, \phi)$ be

$$Q_x(a, \phi) := \{y \in \mathbb{Z}^d : h(y) < g_x(y) + |x|^a \phi(|x|), g_x(y) \leq g(x)\}$$

Definition 3 (CHAP: convex hull approximation property)**.** *A sub-additive function h satisfies the convex hall approximate property (CHAP) with exponent a and correction ϕ if there exist $L > 1$ such that*

$$x/\alpha \in Co(Q_x(a, \phi)) \text{ for some } \alpha \in [1, L], \text{for all } x \in \mathbb{Q}^d$$

where $Co(*)$ denotes the convex hull.

Definition 4 (GAP: general approximation property)**.** *For $a > 0$ and $\phi \in \Phi$, we say a sub-additive function h satisfies general approximation property (GAP) with exponent a and correction factor ϕ if there is $M > 1$ and $C > 0$ such that*

$$g(x) \leq h(x) \leq g(x) + C|x|^a \phi(|x|)$$

for all $x \in \mathbb{Z}^d$ and $|x| \geq M$.

With these definition, Alexander's main theorem is the following:

Theorem 8 (Theorem 1.8 in [1])**.** *Suppose h is a non-negative sub-additive function on $\mathbb{Z}^d$ which has sub-linear growth. If h satisfies CHAP with exponent a and correction factor ϕ for some $a > 0$, then h satisfies GAP with exponent a and correction factor ϕ*

The proof of this theorem can be found in Section 2 of [1]. Then, by applying Theorem 8 to the first passage percolation model, the upper bound of the non-random fluctuation could be limited to:

$$\mathbb{E}T(0, x) - \mu(x) \leq C||x||^{1/2} log(||x||_1),$$

and the proof is provided in Section 3 of [1].

Also, Alexander improved the the error bounds in the shape theorem

Theorem 9 (Theorem 3.1 in [1]). *If $F(0) < p_c$ and $\int e^{\lambda x} dF(x) < \infty$ for some $\lambda > 0$, then for constants C_1 and C_2,*

$$\mathbb{P}\left((1 - C_1 t^{-1/2} logt)B_0 \subset t^{-1}B(t) \subset (1 + C_2 t^{-1/2} logt)B_0 \mid t \text{ large}\right) = 1$$

Note, this result will be used to get a lose estimator in our simulation.

The current state of art was proven in 2014 that for the low moment case along with a concentration inequality for the lower tail of $T(0, x)$ in [7] by using Alexander's method.

Theorem 10 (Proposition 1.1 in [7]). *Assume that $P(\tau_e = 0) < p_c$ and $\mathbb{E}Y^2 < \infty$ where Y is the minimum of d i.i.d. copies of τ_e. There exists C such that for all $x \in \mathbb{Z}^d$ with $||x||_1 > 1$,*

$$\mu(x) < \mathbb{E}T(0, x) \leq \mu(x) + C\sqrt{||x||_1 log(||x||_1)}$$

4 Practical Algorithms for the Shortest Path Problem

In this section, we would introduce several key algorithms in graphical optimization theory that are handy to find the shortest path in the graphical model. Instead of limiting our discussion in the square lattice $\mathbb{Z}^2$ as we did in the previous sections, these algorithms will be discussed on the general graph $G = (V, E)$.

Similar to the definitions in the square lattice, in graph $G = (V, E)$, E represents the collection of all edges and V represents the collection of all vertices. To solve the shortest path problem, we are provided with a direct, weighted graph $G = (V, E)$ with weight function $w : E \to \mathbb{R}$ (we sometimes use w_e for the weight of e, which is equivalent to $w(e)$). A path p is an ordered sequence of vertices $(v_0, v_1, v_2, ..., v_k)$ such that there exist edge $e = (v_i, v_{i+1})$ for each $i \in \{0, 1, 2, ..., k\}$ in E. The weight of path p is the sum of its edges:

$$w(p) = \sum_{i=0}^{k-1} w(v_i, v_{i+1})$$

The *shortest path weight* $\delta(u, v)$ and *shortest path* p from vertex u to vertex v is defined as:

Definition 5 (Shortest Path and Shortest Path Weight)**.**

$$\delta(u,v) = \begin{cases} min\{w(p)\}, & \textit{if there is a path } p \textit{ from } u \textit{ to } v \\ \infty, & \textit{otherwise} \end{cases}$$

and a shortest path from vertex u to vertex v is any path p with weight $w(p) = \delta(u,v)$

It's clear that this shortest path weight has the same definition as the passage time between two vertices (as we introduced in section 1) and that's why we are introducing these algorithms for practical use of the first passage percolation model. Note, there are other variants of the shortest path problem, but we are only focusing on the single-source shortest-paths problem in this section.

One of the key properties of the shortest path is its optimal substructure, which many algorithms rely on. The following lemma states this property more precisely.

Lemma 3 (Subpaths of shortest paths are shortest paths)**.** *For weighted, direct graph* $G = (V, E)$ *with weight function* $w : E \to \mathbb{R}$*, let path* $p =< v_0, v_1, ..., v_k >$ *be a shortest path from vertex* v_0 *to vertex* v_k *and, for any* i *and* j *such that* $0 \le i \le j \le k$*, let* $p_{ij} =< v_i, v_{i+1}, ..., v_j >$ *be the subpath of* p *from* v_i *to* v_j*. Then,* p_{ij} *is a shortest path from* v_i *to* v_j*.*

Proof. Let's prove by contradiction. It's obvious that $w(p) = w(p_{0i}) + w(p_{ij}) + w(p_{jk})$ if we decompose path p into p_{0i}, p_{ij}, and p_{jk}. Suppose that there exist a path p'_{ij} from vertex v_i to vertex v_j with weight $w(p'_{ij}) < w(p_{ij})$. Then, the combination of paths p_{0i}, p'_{ij}, and p_{jk} is a path from v_0 to v_k whose weight $w(p_{0i}) + w(p'_{ij}) + w(p_{jk}) < w(p)$. This is a contradiction with the assumption that p is a shortest path from v_0 to v_k. □

4.1 Relaxation

The key technique used in the shortest path algorithm is called *relaxation*. To introduce this process, we would first need to give an estimated shortest path weight (or called *shortest-path estimate*) $v.d$ to each vertex v in V. To initialize the shortest-path estimate, $v.d$ is set to ∞ for every vertex v in V except the source node s, and $s.d$ is set to 0 as the shortest path of getting back to itself is always 0.

Meanwhile, in order to keep track of the shortest path, another attribute $v.\pi$ is added to each v in V in order to record the vertex's predecessor in the shortest path. In the initialization, $v.\pi$ is set to NIL for each vertex v in V.

Algorithm 1 Initialize_Single_Source(G, s)

for each vertex $v \in G.V$ **do**
  $v.d = \infty$
  $v.\pi = NIL$
end for
$s.d = 0$

Then, to *relax* an edge $e = (u, v)$, we need to check whether the shortest path to v could be improved by going through u. If so, update v's estimated weight $v.d$ and predecessor $v.\pi$ by the following pseudo-code.

Algorithm 2 Relax(u, v, w)

if $v.d > u.d + w(u, v)$ **then**
  $v.d = u.d + w(u, v)$
  $v.\pi = u$
end if

This process has many interesting properties that are essential for discussions in this section. Here are several of them with corresponding proofs [4].

Lemma 4 (Triangle inequality)**.** *For any edge $(u, v) \in E$, we have $\delta(s, v) \leq \delta(s, u) + w(u, v)$.*

Proof. We prove this lemma by contradiction: Assume that there exist an edge $(u, v) \in E$ such that $\delta(s, v) > \delta(s, u) + w(u, v)$.

Then, this contradicts the definition of shortest path weight since combining the shortest path from s to u with the edge (u, v) would yield a path that has a lower weight than the shortest path weight from s to v.

Therefore, $\delta(s, v) \leq \delta(s, u) + w(u, v)$ for all edges (u, v) in E. □

Lemma 5 (Upper-bound property)**.** *For all vertices $v \in V$, $v.d \geq \delta(s, v)$ and once $v.d$ achieves the value $\delta(s, v)$. it never changes.*

Proof. We prove this lemma by induction over the number of relaxation steps.

For the base case, as we have initialized $v.d = \infty$ for all vertices except the source s, we have that $v.d \geq \delta(s, v)$ for all $v \in V - \{s\}$. Also, since $\delta(s, s) = -\infty$ if s is on a negative-weight cycle and 0 otherwise, we have $s.d \geq \delta(s, s)$. Therefore, the statement holds for the base case. For the inductive step, consider relaxing edge (u_0, v_0). By the inductive hypothesis, we have that $v.d \geq \delta(s, v)$ for all $v \in V$ before the relaxation. Then, by this relaxation, only $v_0.d$'s value might be changed due to the definition of relaxation. If it doesn't change, the statement holds automatically. If it changes, we have

$$v.d = u.d + w(u, v) \geq \delta(s, u) + w(u, v) \geq \delta(s, v)$$

by applying the triangle inequality and inductive hypothesis. Therefore, the statement holds.

To show that $v.d$ doesn't change any more once $v.d = \delta(s, v)$. We have shown above that $\delta(s, v)$ is a lower bound for $v.d$ and, by the definition of relaxation, the estimated weights couldn't be increased. Therefore, the value of $v.d$ is monotonically decreasing and doesn't change any more once it reaches its lower bound. □

Lemma 6 (Convergence property)**.** *If path $p =< s, v_1, ..., u, v >$ is a shortest path in G for some $u, v \in V$, and if $u.d = \delta(s, u)$ before relaxing the edge (u, v), then $v.d = \delta(s, v)$ after relaxing (u, v).*

Proof. After relaxing (u, v), we have

$$v.d \leq u.d + w(u, v) = \delta(s, u) + w(u, v) = \delta(s, v)$$

by the definition of relaxation and Lemma 3. Then, by the upper-bound property, we have $v.d \geq \delta(s, v)$, which indicates the equality that we are looking for. □

Lemma 7 (Path-relaxation property)**.** *If path $p =< s, v_1, ..., v_k >$ is a shortest path from s to v_k, and we relax the edges of p in the order $(s, v_1), (v_1, v_2)$, ..., (v_{k-1}, v_k), then $v_k.d = \delta(s, v_k)$. This property holds regardless the occurrence of other edge relaxations.*

Proof. We prove this lemma by induction: This lemma is equivalent as showing that after the i-th edge of path p is relaxed, we have $v_i.d = \delta(s, v_i)$.

For the base case $i = 0$, this is the situation before any edges of p have been relaxed and we have, as initialized, $s.d = 0 = \delta(s, s)$. By the upper-bound property proven above, we know that $s.d$ always equal to 0 after initialization no matter how the edges are relaxed.

For the inductive step, we assume $v_{i-1}.d = \delta(s, v_{i-1})$. Then, by the convergence property proven above, after relaxing this edge, we have $v_i.d = \delta(s, v_i)$, which was to be shown. □

Lemma 8 (No-path Property)**.** *If there is no path from s to v, then we always have $v.d = \delta(s, v) = \infty$*

Proof. By the upper-bound property (Lemma 5), we have that $\infty = \delta(s, v) \leq v.d$. Therefore, $v.d = \infty = \delta(s, v)$ □

4.2 Bellman-Ford algorithm

It's obvious from the aforementioned properties that by repetitively relaxing every edge in the graph, the estimated shortest path weights are monotonically decreasing until reaching the lowest possible value, which is the actual shortest path weights.

The *Bellman-Ford algorithm* uses this approach to solve the single-source shortest-path problem. However, this strategy will fail when there is a negative-weight cycle as the path could traverse the negative-weight cycle for arbitrarily many times. Therefore, the *Bellman-Ford algorithm* employs a check that runs through all the edges after all relaxations and, if there is a pair of vertices that can still be optimized, it would return $False$ to indicate that there is a negative-weight cycle (it returns $True$ if and only if there is no reachable negative-weight cycle).

The pseudo-code for the Bellman-Ford algorithm is shown in Algorithm 3, and its run-time complexity is $O(|V||E|)$ where $|V|$ stands for the number of vertices in V and $|E|$ stands for the number of edges in E.

It might be noticeable that the Bellman-Ford algorithm relaxes all edges by exactly $|V| - 1$ times. We would prove why $|V| - 1$ times is sufficient in the following lemma, and use it to prove the correctness of the Bellman-Ford algorithm.

Lemma 9. *For weighted, directed graph $G = (V, E)$ with source vertex $s \in V$ and weight function $w : E \to \mathbb{R}$. If there is not negative-weight cycles reachable from s in G, then, after relaxing all edges by $|V| - 1$ times, we have $v.d = \delta(s, v)$ for all v reachable from s.*

Algorithm 3 Bellman_Ford(G, w, s)

```
Initialize_Single_Source(G, s)
for i = 1 to |G.V| − 1 do
  for each edge (u, v) ∈ G.E do
    Relax(u, v, w)
  end for
end for
for each edge (u, v) ∈ G.E do
  if v.d > u.d + w(u, v) then
    return False
  end if
end for
return True
```

Proof. Let v be any vertex reachable from s and let $p =< v_0, v_1, ..., v_k >$ be any shortest path from s to v (i.e. $v_0 = s$ and $v_k = v$). A shortest path doesn't contain duplicated edges, so p has at most $|V| - 1$ edges, which means $k \leq |V| - 1$. Then, as the i-th iteration relaxes the edge v_{i-1}, v_i for $i = 1, 2, ..., k$, we get $v.d = v_k.d = \delta(s, v_k) = \delta(s, v)$ by the path-relaxation property (Lemma 7). □

Theorem 11 (Correctness of the Bellman-Ford algorithm [4])**.** *Consider a weighted, direct graph $G = (V, E)$ with source s and weight function $w : E \to \mathbb{R}$. The Bellman-Ford algorithm would return True if G doesn't contain negative-weight cycles reachable from s, then for all v in$V$ we have $v.d = \delta(s, v)$, and the predecessor subgraph G_π is a shortest-paths tree rooted at s. If G contains a negative-weight cycle, then the algorithm would return False.*

Proof. Suppose that G doesn't contain any negative-weight cycles reachable from s. Lemma 9 above proves that at termination, $v.d = \delta(s, v)$ for all vertices $v \in V$ that's reachable from s. As for the vertices that's not reachable from s, $v.d = \delta(s, v)$ follows from the no-path property (Lemma 8). Therefore, we have that at termination, $v.d = \delta(s, v)$ for all vertex $v \in V$.

Then, we show that the algorithm would return True if there are no reachable negative-weight cycles. From the result above, we have

$$v.d = \delta(s, v) \leq \delta(s, u) + w(u, v) = u.d + w(u, v)$$

by triangle inequality (Lemma 4). Therefore, none of the if tests in the Bellman-Ford algorithm will return False, and it would therefore return

True.

As for the case that G contains reachable a negative-weight cycle from source s. Let's say that the negative-weight cycle is $c =< v_0, v_1, ..., v_k >$ where $v_0 = v_k$, and

$$\sum_{i=1}^{k} w(v_{i-1}, v_i) < 0$$

Let's prove that the algorithm will return False by contradiction. Assume that the Bellman-Ford algorithm would return True, which means $v_i.d \leq v_{i-1}.d + w(v_{i-1}, v_i)$ for $i = 1, 2, ..., k$. This implies

$$\begin{aligned}\sum_{i=1}^{k} v_i.d &\leq \sum_{i=1}^{k} (v_{i-1}.d + w(v_{i-1}, v_i)) \\ &= \sum_{i-1}^{k} v_{i-1}.d + \sum_{i=1}^{k} w(v_{i-1}, v_i)\end{aligned}$$

Since $v_o = v_k$, so we have

$$\sum_{i=1}^{k} v_i.d = v_{i-1}.d$$

Also, because the negative-weight cycle c is reachable from s, so each $v_i.d$ is finite for $i = 1, 2, ..., k$. Therefore,

$$0 \leq \sum_{i=1}^{k} w(v_{i-1}, v_i)$$

which contradicts with the assumption that c is a negative-weight cycle. Therefore, the Bellman-Ford would return False, which completes the proof. □

4.3 Dijkstra's algorithm

Another commonly used algorithm in solving the single-source shortest path problem is the *Dijkstra's algorithm*. Though Dijkstra initially proposed the algorithm to find the shortest path between two vertices [8], its variant is used to find the shortest paths from a fixed source vertex to all other

vertices and produce a shortest-path tree [4]. Different from the Bellman-Ford algorithm, Dijkstra's algorithm solves the single-source shortest path problem on a weighted, directed graph $G = (V, E)$ where all the edges have a non-negative weight (i.e. the weight function is $w : E \to [0, \infty)$). As the random variable τ_e assigned to each edge in our first passage percolation model is non-negative, we can take advantage of this non-negative weights restriction of Dijkstra's algorithm to get a better running time compared to the Bellman-Ford algorithm, which will be shown later.

From our previous discussion in the properties of relaxation and repeating relaxation on all edges, it might be noticeable that, roughly speaking, the vertices closer to the source would achieve their shortest-path weight earlier than the farther ones (and we took advantage of this fact to prove the correctness of the Bellmen-Ford algorithm (Theorem 4)). The key idea in the Dijkstra's algorithm is to maintain a set S of vertices whose shortest-path weights from source vertex s are already determined. The algorithm repeatedly chooses the vertex $u \in V - S$ with the minimum estimated shortest path weight $u.d$, adds u to S, then relaxes all outgoing edges leaving u.

In the following algorithm (Algorithm 4), we use a min-priority queue Q to keep all the vertices in $V - S$, keyed by their $v.d$ values. The Extract_Min method is the standard method that removes the smallest object in the min-priority queue.

Algorithm 4 Dijkstra(G, w, s)

```
Initialize_Single_Source(G, s)
S = ∅
Q = G.V
while Q ≠ ∅ do
  u = Extract_Min(Q)
  S = S ∪ {u}
  for each vertex v ∈ G.adj[u] do
    Relax(u, v, w)
  end for
end while
```

It's important to notice that Dijkstra's algorithm always takes the vertex with the smallest estimated shortest-path weight, which is a greedy strategy. However, as a fundamental fact in search algorithms, we know that greedy strategies do not always yield the optimal solution. Therefore, in order to show the correctness of Dijkstra's algorithm, it's essential to show that for

each u extracted from Q, we have $u.d = \delta(s, u)$.

Theorem 12 (Correctness of Dijkstra's algorithm (Theorem 24.6 in [4]))**.** *Dijkstra's algorithm, run on a weighted, directed graph $G = (V, E)$ with non-negative weight function w and source s, terminates with $u.d = \delta(s, u)$ for all vertices $u \in V$.*

Proof. To show the correctness of Dijkstra's algorithm, we prove that at the start of each iteration of the **while** loop, $v.d = \delta(s, v)$ for each vertex $v \in S$. Then, since all vertices in $G.V$ are added to S, we then get the desired result.

Then, it is sufficient to show that we have $u.d = \delta(s, u)$ when u is added to set S. Once we have that $u.d = \delta(s, u)$, the upper-bound property (Lemma 5) indicates that the equality would hold thereafter, which is the property that we wanted for the set S (the vertices' shortest-path weight from s is determined).

For the initialization phase, we have $S = \emptyset$ so it's trivially true that all vertex v in S has $v.d = \delta(s, v)$.

In the maintenance phase, we prove that in each iteration $u.d = \delta(s, u)$ when u is added to set S by contradiction: let u be the first vertex such that $u.d \neq \delta(s, u)$ when it's added to S.

Trivially, we have that $u \neq s$ since s is the source vertex and it's the first vertex added to S with $s.d = delta(s, s) = 0$. Then, $u \neq s$ indicates $S \neq \emptyset$ before u is added.

Also, if there is no path from s to u, we have $u.d = \delta(s, u) = \infty$ by initialization and the no-path property (Lemma 8), which is a contradiction with our assumption that $u.d \neq \delta(s, u)$. Therefore, we only need to consider the case that there are paths from s to u.

Trivially, since there is at least one path, there is a shortest path from s to u. Let's call this shortest path p. Note that prior adding u to S, p connects a vertex in S (which is s) and a vertex in $V - S$ (which is u). Assume vertex y is the first vertex in p such that y is in $V - S$, and let x be y's predecessor along p, which is in S by the assumption of y. Then, we can decompose p into to paths: p_1 from s to x, from x to y, and p_2 from y to u.

We claim that $y.d = \delta(s, y)$ when u is added to S. To show this claim, we make use of the fact that x (y's predecessor along p) is in S. Since we assumed that u is the first vertex added to S such that $u.d \neq \delta(s, u)$, we have $x.d = \delta(s, x)$ when x is added to S. Then, since y is adjacent to x, so edge (x, y) is relaxed and the claim follows from the convergence property (Lemma 6).

Now, we can get a contradiction. Because y appears before u along the shortest path from s to u and all edges have non-negative weights, we have $\delta(s, y) \leq \delta(s, u)$. Therefore

$$\begin{aligned} y.d &= \delta(s, y) \\ &\leq \delta(s, u) \\ &\leq u.d, \quad \text{by the upper-bound property} \end{aligned}$$

However, u and y were both in $V - S$ when the algorithm chooses u in *Extract_min(Q)*, so we have $u.d \leq y.d$. Therefore, by double inclusion, we have

$$y.d = \delta(s, y) = \delta(s, u) = u.d$$

which is a contradiction with our choice of u ($u.d \neq \delta(s, u)$). Therefore, we have that, when u is added to S, $u.d = \delta(s, u)$ and the equality holds thereafter.

At termination, we have $Q = \emptyset$ (which is the termination condition of the while loop in the algorithm), which implies $S = V$ since $Q = V - S$. Therefore, we have $v.d = \delta(s, v)$ for all vertex v in V, which is the desired result. □

As we mentioned previously, Dijkstra's algorithm runs faster than the Bellman-Ford algorithm and we can take advantage of it since the random variables τ_e for each edge e in our first passage percolation are non-negative. In the following part, we are going to analyze the running time of Dijkstra's algorithm. Note, the following analysis is based on specific implementations of the min-priority queue, and the result varies if other implementations are used.

In Dijkstra's algorithm, it maintains a min-priority queue Q, which is equal to $V - S$, and there are three operations: INSERT, EXTRACT_MIN, and DECREASE_KEY (which is implicit in RELAX since the queue may need to be updated when the estimated weights change during relaxation). For each vertex, the operation INSERT and EXTRACT_MIN are called exactly once because vertex are added to the queue at initialization ($Q = G.V$) and each vertex is extracted exactly once. Also, because each vertex is added to S exactly once, each edge in $G.E$ is relaxed once in the **for** loop that iterates through $G.adj[u]$ for each u added to S (it's easier to see by considering an

adjacency list implementation of the edges). Therefore, the algorithm calls DECREASE_KEY at most $|E|$ time.

Then, the question comes down to be how the min-priority queue should be implemented. An intuitive approach can advantage of the fact that the size of the queue is known, which is $|V|$. Then, we can just keep an array of size $|V|$ and store $v.d$ in the v-th entry of the array. Each INSERT and DECREASE_KEY would take $O(1)$ time, and each EXTRACT_MIN would take $O(V)$ time (by searching through the entire array). Therefore, the overall run time of the algorithm is $O(V^2 + E) = O(V^2)$.

In our first passage percolation model on the square lattice $\mathbb{Z}^2$, the graph is very sparse, and we can take advantage of it. In a more general case, we consider a graph is sufficiently sparse if $E = o(V^2/\log V)$, and we can improve the run time of Dijkstra's algorithm by implementing the min-priority queue with a binary min-heap. Then, the time to build this binary min-heap is $O(V)$. Each EXTRACT_MIN and DECREASE_KEY operation would take $O(\log V)$ time. There are $|V|$ EXTRACT_MIN operations and $|E|$ DECREASE_KEY operations, which makes the running time $O((V + E)\log V)$. If all vertices are reachable from the source, then the run time is $O(E \log V)$ since every vertex has at least one edge connected to it. This is an improvement from the aforementioned straightforward array implementation if $E = o(V^2/\log V)$ as mentioned.

5 Simulation

Our simulation for fluctuation simulates the first passage percolation model with exponentially distributed weights on $\mathbb{Z}^2$ and was performed on a MacBook Pro with 3.1 GHz Intel Core i5. The reason we use this distribution is due to the memoryless property of the exponential distribution makes the generation of the model much faster than the others. Also, we limited the number of verteices to 9,000,000 (3,000 $\times$ 3,000) due to the limitation in computational power.

Once the weights are assigned to each edge. We used Dijkstra's algorithm integrated in the igraph package (library with Python bindings) to compute the passage time for each vertex. Then, by using a greedy algorithm to find the smallest passage time among the vertices on the side of the box, we find the time for the model to hit the sides of the box and denote this time as t_0. The blue plot in Figure 1 is the scatter plot of all the vertices in $\{x : T(0, x) < t_0\}$, which is an approximation of the limit shape. It's easy

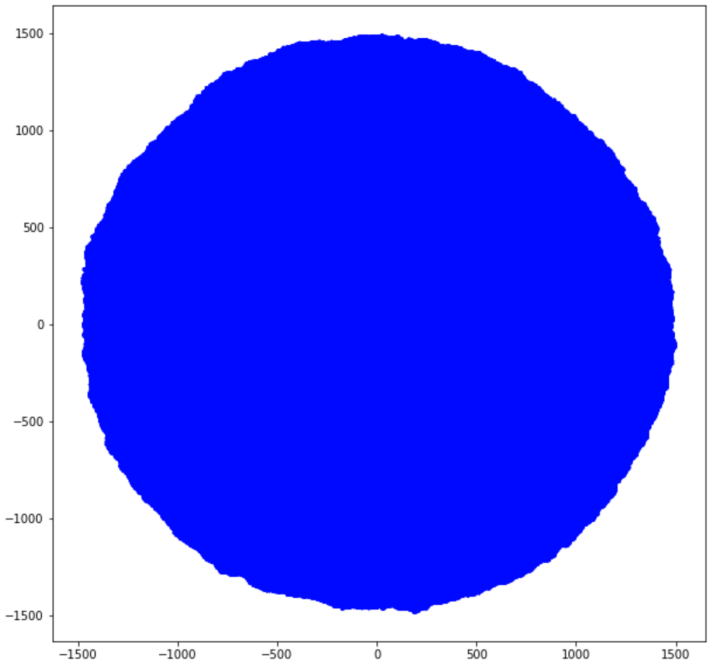

Figure 1: Approximation of Limit Shape

to see that the limit shape is curved and resemble a circle, which is coherent with Richardson's result in [16].

Also, we adopted the results from studies in the fluctuation in passage time to limit the error in our estimation of the time constant. We used

$$\hat{\theta} = \frac{1}{k}\sum_{i=1}^{k} T_i(0,x)$$

as an unbiased estimator for $\mathbb{E}T(0,x)$. Then, because

$$\mathrm{Var}(T(0,x)) = \mathbb{E}[T(0,x) - \mathbb{E}T(0,x)]^2$$

and this variance is bounded up by $C\frac{||x||_1}{log||x||_1}$ as in Theorem 7, we have that

$$\sqrt{\mathrm{Var}\left(\frac{T(0,x)}{||x||_1}\right)} \leq \sqrt{\frac{C_1}{(1+log||x||_1)k||x||_1}}$$

by choosing appropriate constant C_1.

Also, from Alexander's theorem, we can limit the non-random fluctuation as following (choosing $C_2 = 1$)

$$\left|\frac{\mathbb{E}T(0,x)}{||x||_1} - \mu(\frac{x}{||x||_1})\right| \leq \frac{C_2\sqrt{||x||_1}log||x||_1}{||x||_1} = \frac{log||x||_1}{\sqrt{||x||_1}}$$

Therefore, by finding the appropriate constants C_1 and C_2, we can find a threshold for $||x||_1$ to ensure a small error ϵ. As far as we can find, there are not explicit computation of C_1 and C_2 in the literature. To obtain explicit error bounds, we would need to determine these constants in Theorem 7 and Alexander's Theorem.

References

[1] Kenneth S. Alexander. "Approximation of Subadditive Functions and Convergence Rates in Limiting- Shape Results". In: *The Annals of Probability* 25.1 (1997). Publisher: Institute of Mathematical Statistics, pp. 30–55. ISSN: 0091-1798. URL: `http://www.jstor.org/stable/2959528` (visited on 11/16/2020).

[2] Antonio Auffinger, Michael Damron, and Jack Hanson. "50 years of first passage percolation". In: *arXiv:1511.03262 [math-ph]* (Sept. 2016). arXiv: 1511.03262. URL: `http://arxiv.org/abs/1511.03262` (visited on 08/28/2020).

[3] Itai Benjamini, Gil Kalai, and Oded Schramm. "First Passage Percolation Has Sublinear Distance Variance". In: *arXiv:math/0203262* (May 2007). arXiv: math/0203262. URL: http://arxiv.org/abs/math/0203262 (visited on 11/27/2020).

[4] Thomas H. Cormen, ed. *Introduction to algorithms*. 3rd ed. OCLC: ocn311310321. Cambridge, Mass: MIT Press, 2009. ISBN: 978-0-262-03384-8.

[5] J Theodore Cox and Richard Durrett. "Some limit theorems for percolation processes with necessary and sufficient conditions". In: *The Annals of Probability* (1981). Publisher: JSTOR, pp. 583–603.

[6] Michael Damron, Jack Hanson, and Philippe Sosoe. "Sublinear variance in first-passage percolation for general distributions". In: *Probability Theory and Related Fields* 163.1-2 (2015). Publisher: Springer, pp. 223–258.

[7] MICHAEL Damron and NAOKI Kubota. "Gaussian concentration for the lower tail in first-passage percolation under low moments". In: *arXiv preprint arXiv:1406.3105* (2013). Publisher: Citeseer.

[8] Edsger W Dijkstra et al. "A note on two problems in connexion with graphs". In: *Numerische mathematik* 1.1 (1959), pp. 269–271.

[9] Bradley Efron and Charles Stein. "The jackknife estimate of variance". In: *The Annals of Statistics* (1981). Publisher: JSTOR, pp. 586–596.

[10] J. M. Hammersley and D. J. A. Welsh. "First-Passage Percolation, Subadditive Processes, Stochastic Networks, and Generalized Renewal Theory". en. In: *Bernoulli 1713 Bayes 1763 Laplace 1813: Anniversary Volume*. Ed. by Jerzy Neyman and Lucien M. Le Cam. Berlin, Heidelberg: Springer, 1965, pp. 61–110. ISBN: 978-3-642-99884-3. DOI: 10.1007/978-3-642-99884-3_7. URL: https://doi.org/10.1007/978-3-642-99884-3_7 (visited on 10/06/2020).

[11] Harry Kesten. "Aspects of first passage percolation". In: *École d'été de probabilités de Saint Flour XIV-1984*. Springer, 1986, pp. 125–264.

[12] Harry Kesten. "On the speed of convergence in first-passage percolation". In: *The Annals of Applied Probability* (1993). Publisher: JSTOR, pp. 296–338.

[13] Harry Kesten et al. "The critical probability of bond percolation on the square lattice equals 1/2". In: *Communications in mathematical physics* 74.1 (1980), pp. 41–59.

[14] John FC Kingman. "The ergodic theory of subadditive stochastic processes". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 30.3 (1968). Publisher: Wiley Online Library, pp. 499–510.

[15] Thomas M Liggett. "An improved subadditive ergodic theorem". In: *The Annals of Probability* 13.4 (1985). Publisher: Institute of Mathematical Statistics, pp. 1279–1285.

[16] Daniel Richardson. "Random growth in a tessellation". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 74. Issue: 3. Cambridge University Press, 1973, pp. 515–528.

[17] DE Wolf and János Kertész. "Noise reduction in Eden models. I". In: *Journal of Physics A: Mathematical and General* 20.4 (1987). Publisher: IOP Publishing, p. L257.

[18] JG Zabolitzky and D Stauffer. "Simulation of large Eden clusters". In: *Physical Review A* 34.2 (1986). Publisher: APS, p. 1523.