SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Fluctuations in First Passage Percolation : an overview from the perspective of Boolean functions

av

Damian Cid

2023 - M5

Fluctuations in First Passage Percolation : an overview from the perspective of Boolean functions

Damian Cid

Självständigt arbete i matematik 30 högskolepoäng, Advanced nivå

Handledare: Daniel Ahlberg

2023

Abstract

The study of fluctuations is one of the central open problems in growths models. We focus on one of the most popular growth models : first passage percolation, and we restrain to distributions of shape $(1-p)F+pG$ with G dominates F in analogy to the Bernoulli distribution. In this context, the study of fluctuations is deeply connected to the study of influences which in turn is a central concept in the study of threshold phenomena for Boolean function. We give a first overview to Boolean theory and the notions of influence and threshold. Finally, we show that this threshold phenomena also appears in the context of first passage percolation for the functions $p \mapsto \mathbb{P}_p(Tn \geq \theta n)$.

Contents

1 Introduction **7**

2 First Passage Percolation **8**
- 2.1 Definition . . . 8
- 2.2 First properties . . . 8
- 2.3 Geodesics . . . 9

3 Boolean functions **9**
- 3.1 Pivotal bit and influence . . . 9
- 3.2 The study of monotone functions . . . 10
- 3.3 The sharp threshold phenomena . . . 11
- 3.4 Links between Variance and Influence . . . 14

4 Influences in first passage percolation **15**
- 4.1 Graph settings . . . 15
  - 4.1.1 The torus . . . 15
  - 4.1.2 The box . . . 16
  - 4.1.3 Point to point . . . 16
- 4.2 Estimating Influences in the different settings . . . 16
  - 4.2.1 Bounding influence on the Torus . . . 16
  - 4.2.2 Bounding the influence on the square and point to point . . . 17

5 Proving sharp threshold from bounds on the variance **17**
- 5.1 A first example : Tribes . . . 17
- 5.2 Existence of sharp threshold in first passage percolation . . . 19
- 5.3 Some finer control : The study of $N(\varepsilon)$ and the sharpness of the threshold. . . . 20

6 Bound of the variance in FFP **21**

7 Strict monotonicity of the time constant[van den Berg and Kesten, 1993] **23**
- 7.1 Definition of an order over the distributions . . . 24
- 7.2 Proof of increasing time . . . 25
- 7.3 Back to the limiting function . . . 29

8 Conclusion **30**

A Proofs **32**
- A.1 Russo theorem . . . 32
- A.2 Limiting behavior of tribes . . . 32
- A.3 Characterization of second order stochastic domination . . . 34

B Fourier Theory **35**
- B.1 Monotone functions and their spectra . . . 38
- B.2 Non-uniform probability . . . 38
- B.3 Another proof of the Margulis-Russo . . . 39

C The Margulis-Russo Formula **40**
- C.1 What regularity do we have for $\mu(p)$? . . . 40

Notation

- $\mathbb{G} = (V, E)$ denotes a graph
- $\mathbb{P}$ is a measure of probability and $\mathbb{E}()$ is the expectation operator, $\mathrm{Var}\,()$ is the variance operator.
- γ denotes a path in a graph.
- π, π_n denotes an optimal path and will often be called a geodesic.
- $\mathbb{1}_A$ is the indicator function of the event A.
- $\Omega = (\Omega, \mathcal{F}, \mathbb{P})$ is a probabilistic space.
- ω is an element of Ω.
- T is the optimal time of passage.
- $\tau = (\tau_e)_e$ is a collection of random variables that indicate the time of passage through an edge.
- F, G are cumulative distribution function.
- ν is the distribution law. We will often mix up a distribution, and it's cumulative function.
- X, Y are random variables.
- $\|.\|_p$ are the $p-$norm operators defined in any measured $\mathbb{R}$-vectorial space (E, μ) as $(\int |\,.\,|^p d\mu)^{\frac{1}{p}}$.
- $\|.\|_\infty$ is the infinite norm defined on $\mathbb{R}^I$ as $\sup_{i\in I} |x_i|$
- $f = O(g)$ or $f \leq O(g)$ if there exists some C such that $f \leq Cg$
- $f = \Omega(g)$ if $g = O(f)$
- $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$
- $f \sim g$ if $f = g + o(g)$

1 Introduction

First passage percolation was first introduced by Hammersley and Welsh [Hammersley and Welsh, 1965] in 1965 to model a fluid flowing through a random medium. The definition of the model it's simple and beautiful (we will focus on the definition over the lattice $\mathbb{Z}^d$, other variants will be introduced later) and goes as follows : we connect each point of $\mathbb{Z}^d$ to its nearest neighbors (for the ℓ_1 metric); we place a random variable τ_e on each of these connections (edges). The collection $(\tau_e)_e$ is supposed to be non-negative, independent identically distributed with common distribution F. The random variable τ_e is interpreted as the time needed to traverse e. For any path γ, the time to cross this path is the sum of the τ_e. Given two points on $x, y \in \mathbb{Z}^d$ we define

$$T(x,y) = \inf_{\gamma \in \Gamma_{x,y}} \sum_{e \in \gamma} \tau_e$$

where the infimum is taken over all the finite paths from x to y. This defines a random metric over $\mathbb{Z}^d$.

One of the main results is known as the shape theorem that states that under mild conditions, there exist a norm N such that $\frac{T(x,y)}{N(x,y)} \xrightarrow{\mathbb{P}} 1$ when $\|x-y\|$ goes to $+\infty$. Despite existing for almost 60 years, there are still many unsolved questions concerning this model and growth models in general; one of them is the estimation of fluctuations, i.e. the $T(x,y) - N(x,y)$ residue. Several efforts have been made to have a better understanding of the fluctuations around the time constant and the study of influences plays a central role in current state of the art [Benjamini et al., 2011] where it was shown (for a Bernoulli distribution) that

$$\mathrm{Var}\left(T(0, ne_1)\right) = O\left(\frac{n}{\log n}\right)$$

thanks to Talagrand's inequality that gives a bound on the variance that depends on the influences. Finally, equation (12) from [Ahlberg et al., 2023] gives us a relationship between the variance and influence, showing that the understanding of the behavior of the time passage under the perturbation of only one edges leads the way to a global comprehension of fluctuations.

The study of influences is tightly connected to the study of sharp thresholds in boolean theory : that is, for indicator functions defined over $\{0,1\}^n$, we see it as a random variable $f(X_1, \dots, X_n)$ where X_i are independent Bernoulli variables of parameter p. Because f is an indicator function, also follows a Bernoulli law, of parameter $q(p)$. When $q(p)$ is an increasing function that approaches a step function (flat except over a very small interval where it is very steep) we say that the function presents a sharp threshold.

To try to leverage the theory of boolean functions (which are these functions defined over the hypercube $\Omega_n = \{0,1\}$) we will take interest on the case where the edges follow a distribution of type $(1-p)F + pG$ in analogy to what happens in the Bernoulli setting. We will seek to study the time constant (and other random variables) as a function of p. In this context, we show that the functions $f_n = \mathbb{1}_{T(0,ne_1) > \theta n}$ present a sharp threshold. The proof is obtained using bounds on fluctuations and the fact that $\mathbb{E}_p[f_n]$ is strictly increasing as a function of p.

The main motivation for this work has been to investigate the connection between influence and variance. We use known bound on the variance to derive the existence of a sharp threshold (which is indicator of small influences). We would have liked to obtain and improvement on the bounds on the variance unfortunately this seems quite hard.

2 First Passage Percolation

2.1 Definition

Definition 2.1 (First passage percolation model (general)).
Let $\mathbb{G} = (V, E)$ be a (countable) graph. We consider a collection of non-negative independent, identically distributed random variables indexed by E : $(\tau_e)_{e\in E}$. We note ν their common law i.e. $\forall e \in E, \tau_e \sim \nu$.
Consider $\mathcal{P}_{x,y}$ to be the collection of finite paths from x to y.
For $\gamma \in \mathcal{P}_{x,y}$ we define the time of passage through γ as

$$\gamma \cdot \tau = \sum_{e\in\gamma} \tau_e.$$

We can thus define the distance between two vertices of the graph as the best time of passage between x and y.

$$T_\omega(x, y) = \inf_{\gamma\in\mathcal{P}_{x,y}} \gamma \cdot \tau. \tag{1}$$

We note T_ω this metric when we want to emphasize the fact that the metric is random (depends on ω), but we will generally denote it simply $T(x, y)$.

Because the time of passage through an edge is non-negative, we restrain ourselves without loss of generality to paths without loops which allows us to represent paths by their indicator function

$$\mathbb{1}\gamma : E \to \{0, 1\}, e \mapsto e \in \gamma.$$

By abuse of notation, we will often confuse a path, and it's indicator function.

This justifies the scalar product notation for the time of passage through γ. Indeed, this corresponds to the natural dot product in $\mathbb{R}^E$ between the random function τ and the indicator function of γ.

2.2 First properties

Proposition.
If τ is almost surely positive, then, T_ω is almost surely a distance.

Proof. Indeed, if $\mathbb{P}(\tau_e = 0) = 0$ for every $e \in V$, as V is a countable set, we have $\mathbb{P}\left(\bigcup_{e\in V}(\tau_e = 0)\right) = 0$.

- (Symmetry) We have symmetry as $\mathcal{P}_{x,y} \simeq \mathcal{P}_{y,x}$ via the path inversion.
- (Non-negative) For any path γ, $\sum_{e\in\gamma} \tau_e$ is almost surely well-defined and positive (as the sum of almost surely positive random variables).
- (Triangular inequality) Let $x, y, z \in \mathbb{G}$. Let $\varepsilon > 0$. By definition of the inf there exists a path $\gamma_x^y \in \mathcal{P}_{x,y}$ such that $\gamma_x^y \cdot \tau \le T(x, y) + \varepsilon$. We also have $\gamma_y^z \cdot \tau \le T(y, z) + \varepsilon$.

  Then $\gamma_x^z := \gamma_y^z + \gamma_x^y \in \mathcal{P}_{x,z}$ where $+$ is the concatenation operation $+ : \mathcal{P}_{y,z} \times \mathcal{P}_{x,y} \to \mathcal{P}_{x,z}$. We thus have $T(x, z) \le \gamma_x^z \cdot \tau = (\gamma_x^y + \gamma_y^z) \cdot \tau = \gamma_x^y \cdot \tau + \gamma_y^z \cdot \tau \le T(x, y) + T(y, z) + 2\varepsilon$. We can then make ε tend to 0 to obtain the desired result.
- (Defined) We have for every $x \quad T(x, x) = 0$ (this comes from the convention of the existence of the empty path $\emptyset_x \in \mathcal{P}_{x,x}$). We also have for every $x \ne y, T(x, y) \ge \min_{e\in\mathcal{N}(x)} \tau_e$ almost surely greater than 0. $\mathbb{G}$ being countable, we have $\mathbb{P}\left(\bigcup_{x\ne y}(T(x, y) = 0)\right) = 0$

∎

In [Hammersley and Welsh, 1965], the model introduced was with $V = \mathbb{Z}^d$ with $E = \{(x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d \mid \|x - y\|_1 = 1\}$, this has been the most common setting in the literature as well. We will as well take interest mainly in this setting, but we will introduce other graphs settings on section 4.1.

When looking at the model over $\mathbb{Z}^d$, one can extend the distance T_ω to $\mathbb{R}^d$ by adding a time of travel from any real vector to its closest whole vector (this time is often taken as 0, which makes us lose the property of separability).

We note $B_\omega(t) = \{x \in \mathbb{R}^d \mid T_\omega(0,x) \leq t\}$. Let $\mathcal{M} \subset \{\nu : \mathcal{B}([0,+\infty)) \to [0,1]\}$ be the set of non-negative, Borel probability measures satisfying that $\min(t_1,\ldots,t_{2d}) \in L^1$ where t_i are independent and of law ν and with $\nu(\{0\}) < p_c(d)$ where $p_c(d)$ is the threshold for bond percolation in $\mathbb{Z}^d$. The main result in this setting is the Shape Theorem due to Cox and Durett [Cox and Durrett, 1981]

Theorem 2.2 (Shape Theorem)**.**
For each $\nu \in \mathcal{M}$, there exist a deterministic, convex, compact set $\mathcal{B}_\nu$ in $\mathbb{R}^d$ such that for every $\varepsilon > 0, \exists T > 0$ such that :

$$\mathbb{P}\left((1-\varepsilon)\mathcal{B}_\nu \subset \frac{1}{t}B(t) \subset (1+\varepsilon)\mathcal{B}_\nu \quad \textit{for all } t \geq T\right) = 1$$

Furthermore $\mathcal{B}_\nu$ has non-empty interior and is symmetric with respect to the axis of $\mathbb{R}^d$.

Because of the characterization of norms over $\mathbb{R}^d$, this is equivalent to saying that there exist a norm N over $\mathbb{R}^d$ such that

$$\text{for all } x,y \in \mathbb{Z}^d \ : T(x,y) = N(x-y) + o(N(x-y)).$$

2.3 Geodesics

In the case the infimum over all paths presented in 1 is reached, then there exists a path π such that $T(x,y) = \pi \cdot \tau$. In such cases, π is called a geodesic. The existence and uniqueness of geodesics is an important question. In later proof we will often fix a geodesic, without necessary verifying that such object exists, this is done for the sake of simplicity, as carefully approximating the infimum would just add an unnecessary layer of complexity. To reassure ourselves that this is legitimate, we can use the result (9.23) from [Kesten et al., 1986] that states the existence of geodesics for the model over $\mathbb{Z}^d$ when $F(0) < p_c$.

3 Boolean functions

We will try to leverage the knowledge in boolean theory to study first passage percolation. As we will see, boolean theory is very important in classical percolation and has already been used for the study of first passage percolation, notably for the study of fluctuations [Benjamini et al., 2011].

If we restrict ourselves to the case where $\tau_e \sim a + (b-a)\mathcal{B}(p)$ where $\mathcal{B}(p)$ is the Bernoulli law of parameter p. (We can also write $\tau_e \sim (1-p)\delta_a + p\delta_b$) we can choose $\Omega = \{0,1\}^E$ or $\Omega = \{-1,1\}^E$. In this setup, indicator functions will be functions from $\{0,1\}^E \to \{0,1\}$. The functions from the hypercube $\Omega_n = \{0,1\}^n$ to $\{0,1\}$ are called boolean functions. It is a very rich theory that arises quite naturally in problems of combinatorics and computer science.

This section has been extracted from [Garban and Steif, 2012] chapters I and III. Refer to it to go deeper on the study of boolean functions.

3.1 Pivotal bit and influence

Some key concepts from the theory of boolean function are the notion of influence (and pivotal bit).

Define $\begin{array}{cccc} \sigma_i : & \Omega_n & \longrightarrow & \Omega_n \\ & \omega & \longmapsto & \omega^I \end{array}$ with ω^I equal to ω except in the ith coordinate where the bit has been flipped.

We say a coordinate is pivotal for a function if changing its value changes the outcome of the function. More in detail :

Definition 3.1 (Pivotal bit and pivotal set [Garban and Steif, 2012] I.7-I.8).
Let $f : \Omega_n \to \{0, 1\}$ be a boolean function. Let $\omega \in \Omega_n$.
We say $i \in [\![1, n]\!]$ is pivotal for f at ω if $f(\omega) \neq f(\sigma_i(\omega))$

The set of bits that are pivotal (for a fixed ω) is called the pivotal set, we can then define a random variable Π_f with values in $\mathcal{P}([\![1, n]\!])$ as follows:

$$\Pi_f(\omega) := \{i \in [\![1, n]\!] \mid i \text{ is pivotal for } f \text{ for } \omega\}$$

Note that the event $f(\omega) \neq f(\omega^I)$ is independent with the projection along the ith coordinate ω_i. If we identify Ω_n to $(\mathbb{Z}/2\mathbb{Z})^n$, σ_i is just the translation by $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ where the only non-zero component is the ith one; and if we identify it to $\mathbb{U}_2^n$ then it is the multiplication by $\chi_i = (1, \dots, 1, -1, 1, \dots, 1)$

If we push this idea a little more, one can see some links with classical analysis:

The discrete directional derivative of f along the ith component can be written as

$$\delta_i f(\omega) := f(\omega + e_i) - f(\omega) = f(\sigma_i \omega) - f(\omega).$$

We can establish the link between i is pivotal for f and ω and the derivative along the ith component of f at ω is non-null.

It can be seen as the non-zero coordinates of the gradient of f at ω.

Definition 3.2 (Influence of a bit [Garban and Steif, 2012] I.9).
*The **influence** of the ith bit $\mathbb{I}_i(f)$ is defined by :*

$$\mathbb{I}_i(f) = \mathbb{P}\left(i \text{ is pivotal}\right) = \mathbb{P}\left(i \in \Pi_f\right) = \mathbb{E}[|\delta_i f|]$$

*We can introduce quite naturally the **influence vector** which is simply*

$$Inf(f) = (\mathbb{I}_i(f))_{i \in [\![1,n]\!]}$$

The study of influence originally arose in the study of political science to measure the power of different voters to flip the election.[1]

Definition 3.3 (Total influence [Garban and Steif, 2012] I.10).
*Finally, the sum of all the influences is called the **total influence** and defined as*

$$\mathbb{I}(f) = \sum_i \mathbb{I}_i(f) = \|Inf(f)\|_1 = \mathbb{E}[|\Pi_f|]$$

In the case the distribution is not uniform in Ω_n, we will add p to the notation :

Definition 3.4 (Influence of a bit at level p [Garban and Steif, 2012] I.11).
*The **influence** of the ith bit $\mathbb{I}_i^p(f)$ is defined by :*

$$\mathbb{I}_i^p(f) = \mathbb{P}_p\left(i \text{ is pivotal}\right) = \mathbb{P}_p\left(i \in \Pi_f\right)$$

*The **total influence** at level p is defined as*

$$\mathbb{I}^p(f) = \sum_i \mathbb{I}_i^p(f) = \|Inf_p(f)\|_1 = \mathbb{E}_p[|\Pi_f|]$$

3.2 The study of monotone functions

The monotone functions are an interesting object of study in boolean theory, on one hand because they have nice properties, on the other hand because many of the functions of interest happen to be monotone functions.

[1] In CS this can be used to measure how robust is an algorithm to bit flipping and which are the bits that would be worth copying

For instance the existence of an infinite cluster is a non-decreasing function, we will see later that the time of passage is also a non-decreasing function. On of the main results in the study of monotone functions is the Margulis-Russo formula which allows us to deduce interesting results. Finally, the study of sharp threshold makes essentially sense for monotone functions.

To talk of monotony, we need an ordering, we identify the hypercube Ω_n to the natural lattice $(\mathcal{P}(\llbracket 1, n\rrbracket), \subset)$ which gives us the following ordering in the hypercube : $x \leq y$ if $\forall i \in \llbracket 1, n\rrbracket x_i \leq y_i$.

Definition 3.5.
A function f is monotone if $x \leq y$ implies $f(x) \leq f(y)$. An event is monotone if its indicator function is monotone.

We can now present the Margulis-Russo formula (other proof can be found on the appendix B.3 and C.1).

Theorem 3.6 (Margulis-Russo Formula)**.**
Let $\mathbb{1}_A : \Omega_n \to \{0,1\}$ be a monotone function.
Then

$$\frac{d}{dp}\mathbb{E}_p[\mathbb{1}_A] = \sum_i \mathbb{I}_i^p(\mathbb{1}_A)$$

Proof. Suppose each variable x_i has its own parameter p_i. We can thus rewrite this expression as

$$\sum_i \frac{\partial}{\partial p_i}\mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A] = \sum_i \mathbb{P}_{p_1,\dots,p_n}(i \text{ is pivotal for } \mathbb{1}_A)$$

It thus suffices to prove that

$$\forall_i \quad \frac{\partial}{\partial p_i}\mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A] = \mathbb{P}_{p_1,\dots,p_n}(i \text{ is pivotal for } \mathbb{1}_A)$$

Let's fix $i \in \llbracket 1, n\rrbracket$. We then have

$$\begin{aligned}\mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A] &= \mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A(\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}} + \mathbb{1}_{i\notin\Pi_{\mathbb{1}_A}})] \\ &= \mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}}] + \mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A\mathbb{1}_{i\notin\Pi_{\mathbb{1}_A}}]\end{aligned}$$

We have $\mathbb{1}_A\mathbb{1}_{i\notin\Pi_{\mathbb{1}_A}}$ that does not depend on the value of p_i because it is independent of the value of x_i and thus the partial derivative vanishes.

Likewise, we have $\omega \in A$ and i pivotal for ω and $\mathbb{1}_A$ if and only if $x_i = 1$ (because we assume 1_A to be increasing) and i pivotal for ω and $\mathbb{1}_A$.

Furthermore, we have $\mathbb{1}_A\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}} = \mathbb{1}_{x_1=1}\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}}$ and by independence of $\mathbb{1}_{x_1=1}$ and $\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}}$ we have

$$\mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}}] = p_i\mathbb{E}_{p_1,\dots,p_{i-1},p_{i+1},\dots,p_n}[\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}}]$$

Thus,

$$\begin{aligned}\frac{\partial}{\partial p_i}\mathbb{E}_{p_1,\dots,p_n}[\mathbb{1}_A] &= \mathbb{E}_{p_1,\dots,p_{i-1},p_{i+1},\dots,p_n}[\mathbb{1}_{i\in\Pi_{\mathbb{1}_A}}] \\ &= \mathbb{P}_{p_1,\dots,p_n}(i \text{ is pivotal for } \mathbb{1}_A)\end{aligned}$$

∎

3.3 The sharp threshold phenomena

The notion of threshold functions appears already in [Erdős et al., 1960], the authors here look at random graph where N edges are uniformly sampled to connect n different vertices (because law of large numbers, this can be connected to independently adding each edge with probability $p = \frac{N}{\binom{n}{2}}$). They consider some properties of the graph (for instance, connexity, existence of cycles of certain size, ...). For each of these properties, they show that there exist a threshold function $A(n)$ such that if $N(n) = o(A(n))$ then the probability of the graphs fulfilling this property goes to 0 and symmetrically if $A(n) = o(N(n))$ then the probability of the graphs fulfilling this property goes to 1.

For example, for the probability of the graph being connected, we have

$$\lim_{n\to+\infty} \mathbb{P}(G(n,N) \text{ is connected}) = e^{-e^{-2y}} \qquad \text{where} \quad y = \lim_{n\to\infty} \frac{N(n) - 0.5n\log n}{n}$$

Closer to the setting of first passage percolation is the "classical" percolation setting where one starts with a graph, such as the square lattice (defined as $(Z^2, E(\mathbb{Z}^2))$ where $E(\mathbb{Z}^2)$ is the collection of closest neighbor edges, that means the edges with the shape $(x, x+e_i), i = 1, 2, x \in \mathbb{Z}^2$); and then erase each edge of the graph with probability $(1-p)$. One is then interested in the remaining subgraph and again in the properties of this random subgraph: notably how these properties depend on p.

One property that's very often studied is the existence of an infinite connected component. This event belongs to the tail σ-algebra, thus, by the Kolmogorov zero-one law, for a fixed p, we will have that this event happens either almost surely or is negligible. Because this event is increasing (we can not make the infinite cluster disappear by adding edges), there exist $p_c \in [0,1]$ such that for all $p < p_c$, the probability of being an infinite cluster is zero and for $p > p_c$ the probability is one. If we see the probability of this event as a function of p, is simply the step function $\mathbb{1}_{[p_c,1]}$. p_c acts as the threshold value, separating two kinds of regimes that are the complete opposite. The good understanding of this type of discontinuity can be very important ; for instance a good understanding of the critical point for social graphs helps define vaccination policies.

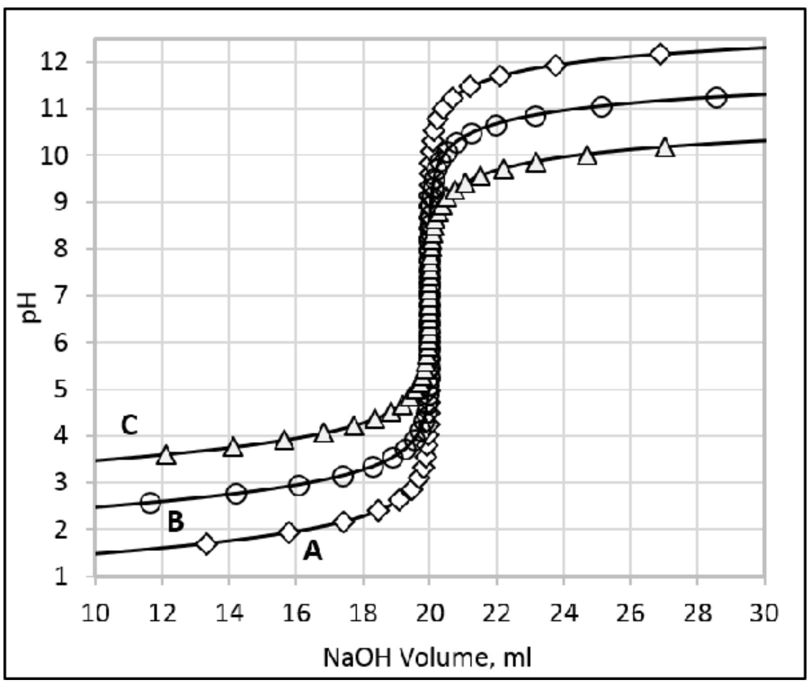


Figure 1: A real life example (from chemistry) of sharp threshold phenomena. The stronger the acid-base couple, the sharper the threshold.

To obtain a better understanding of these thresholds [Russo, 1982] proposes to approximate this tail event by a series of events that is adapted to the (some) filtration of finitely generated σ-algebra. In Russo's paper, the main idea is to argue that the influence of tail events is 0, thus, if the influence of a series of increasing events goes to zero, one can expect that the probability of these events (as a function of p) converges pointwise to a step function. As a consequence of this, we have an approximate zero-one law : the probability of these events as a function of p can be separated into three regimes : one where the probability is very low, one where the probability is very high and finally a transition window where the probability goes from close to zero to close to one; the thinner the window, the sharpest the threshold.

To have a visualization of what these thresholds look like, we can use an example from real-life, in chemistry, in an acid-base titration (see Figure 1), we can see a sharp threshold phenomena in the pH of the solution as a function of the volume of titrant added. In case the function is derivable, the derivative can be seen as an approximation of the Dirac delta function.

Definition 3.7 (Sharp threshold)**.**
*Let $f : \Omega_n \to \{0,1\}$ be a boolean function. We say that f has an ε-**sharp threshold** if there exists $p_0 \in (0,1)$ such that for all $p \leq p_0 - \varepsilon : \mathbb{E}_p[f] \leq \varepsilon$ and for all $p \geq p_0 + \varepsilon : \mathbb{E}_p[f] \geq 1 - \varepsilon$.*
*A series of functions $(f_n)_n$ is said to have **sharp threshold** if there exists a series $(\varepsilon_n)_n$ going to 0 such that for all n, f_n has ε_n-sharp threshold. (Notice that p_0 may depend on n).*

We now have all the elements to prove sharp threshold in the case of evenly distributed influences.

Assume $f : \Omega_n \to \{0,1\}$ is an increasing (non-constant) boolean function such that $\mathbb{I}_i^p(f) = \mathbb{I}_j^p(f)$ for all i, j. We thus have by 3.6 and the fact that all the influences are the same :

$$\mathbb{I}^p(f) \geq c\text{Var}_p\left(f\right)\log(n)$$

By writing $\text{Var}_p\left(f\right) = \mathbb{P}_p\left(f = 1\right)\left(1 - \mathbb{P}_p\left(f = 1\right)\right) \geq \frac{1}{2}\min(\mathbb{P}_p\left(f = 1\right), \left(1 - \mathbb{P}_p\left(f = 1\right)\right))$ and using 3.6 we get :

$$\frac{d}{dp}\mathbb{E}_p[f] \geq \frac{c}{2}\log(n)\min(\mathbb{P}_p\left(f = 1\right), \left(1 - \mathbb{P}_p\left(f = 1\right)\right))$$

Because we assumed f to be non-constant, we have $f(0_n) = 0$ and $f(1_n) = 1$ thus $\mathbb{P}_0(f = 1) = 0$ and $\mathbb{P}_1(f = 1) = 1$.

We also have $p \mapsto \mathbb{P}_p\left(f = 1\right)$ is increasing and continuous and the derivative is strictly positive around $\mathbb{P}_p\left(f = 1\right) = 0.5$ thus there exist a unique p^* such that $\mathbb{P}_{p^*}(f = 1) = 0.5$.

Let p_1 be such that $\mathbb{P}_{p_1}(f = 1) \geq \varepsilon$. Then for $p \in (p_1, p^*)$ we have $\frac{d}{dp}\mathbb{P}_p\left(f = 1\right) \geq \frac{c}{2}\log(n)\mathbb{P}_p\left(f = 1\right)$. Thus, $\frac{d}{dp}\log(\mathbb{P}_p\left(f = 1\right)) \geq \frac{c}{2}\log(n)$. Integrating this inequality between p_1 and p^* we get :

$$\log(\frac{1}{2}) - \log(\varepsilon) \geq \frac{c}{2}\log(n)(p^* - p_1)$$

.

This shows that $(p^* - p_1) \leq \frac{2}{c\log(n)}\log(\frac{1}{2\varepsilon})$ which will go to 0 as n goes to ∞ (for fixed ε).

Let p_2 be such that $\mathbb{P}_{p_2}(f = 1) \leq 1 - \varepsilon$. (This case is symmetric). We have for $p \in (p^*, p_2)$, $\frac{d}{dp}\mathbb{P}_p\left(f = 1\right) \geq \frac{c}{2}\log(n)(1 - \mathbb{P}_p\left(f = 1\right))$ which becomes $-\frac{d}{dp}\log(1 - \mathbb{P}_p\left(f = 1\right)) \geq \frac{c}{2}\log(n)$. We integrate between p^* and p_2 which yields :

$$\log(\frac{1}{2\varepsilon}) \geq \frac{c}{2}\log(n)(p_2 - p^*)$$

Finally, putting everything together, we get

$$p_2 - p_1 \leq \frac{4}{c}\frac{\log(1/2\varepsilon)}{\log(n)}$$

This yields for any p_1 and p_2 such that $\mathbb{P}_{p_1}(f = 1) \geq \varepsilon$ and $\mathbb{P}_{p_2}(f = 1) \leq 1 - \varepsilon$. We deduce that we pass from a probability ε to a probability $1 - \varepsilon$ in the span of $c_1\frac{\log(1/2\varepsilon)}{\log(n)}$ which means we have sharp threshold!

Proposition.
If $f : \Omega_n \to \{0,1\}$ is an increasing function for which the influence is equally distributed. Then f has a sharp threshold. This threshold is at least of size $O(\frac{1}{\ln n})$.

Finally, let's present the theorem due to Russo that we outline on section 3.3. The proof that we present on the annex A.1 is not the original proof from Russo as it uses the BBKL theorem that was not yet proved when Russo published his theorem.

Theorem 3.8 ([Russo, 1982]).
For all $\varepsilon > 0$, there exist $\delta > 0$ such that if A is an increasing event satisfying $\mathbb{I}_i^p(A) < \delta$ for all p and all i, then there exists p_0 such that:

$$\mathbb{P}_p(A) \leq \varepsilon \quad \text{for all } p \leq p_0 - \varepsilon$$

$$\mathbb{P}_p(A) \geq 1 - \varepsilon \quad \text{for all } p \geq p_0 + \varepsilon$$

3.4 Links between Variance and Influence

The reason we are interested in the concept of influence, is because it can be linked to the one of variance. The proof of Russo theorem can be deduced quite swiftly from the BKKL theorem.

The theorem (due to Kahn Kalai Linial) gives us the best bound (up to a constant) of the variance by the sum of the influences :

Theorem 3.9 (Theorem 3.1 [Kahn et al., 1989]).
There exist a universal $c > 0$ such that for any boolean function $f : \Omega_n \to \{0, 1\}$, then there exist $i \in [\![1, n]\!]$ such that

$$c\,Var(f)\,\frac{\log(n)}{n} \leq \mathbb{I}_i(f)$$

The bound obtained in the theorem cannot be improved (without making further hypothesis about f), the case of tribes being an example where the bound is sharp.

Example (Tribes).
Partition $[\![1, n]\!]$ into disjoints blocks of length $\log_2(n) - \log_2(\log_2(n))$: $[\![1, n]\!] = \sqcup_{k=1}^K B_k$ with $|B_k| = \lfloor \log_2(n) - \log_2(\log_2(n)) \rfloor$ and $|B_{K+1}| \leq \lfloor \log_2(n) - \log_2(\log_2(n)) \rfloor$. Define f_n to be $\max_{1 \leq k \leq K} \min_{i \in B_k} x_i$. In other words, f_n is equal to 1 if at least one of the blocks if filled with ones.

We can easily approximate the probability of $(f_n = 0)$ because

$$\begin{aligned}
(f_n = 0) &= \bigcap_{k=1}^{K} (\min B_k = 0) \\
\mathbb{P}(f_n = 0) &= \prod_{k=1}^{K} (1 - 2^{-|B_k|}) \\
&\approx (1 - \frac{\log_2(n)}{n})^K \\
&= e^{-\frac{n}{\log_2(n)} \frac{\log_2(n)}{n} + o(1)} \\
&\to e^{-1}
\end{aligned}$$

This shows that the function is non degenerated, i.e. the probability of the function being equal to 1 will be away from 0 or 1, thus the variance will stayed bounded away from 0.

Now, $i \in B_k$ is pivotal only if all the other $x_j \in B_k$ in the same block are equal to 1 and the output from the others blocks is equal to 0.

So

$$\begin{aligned}
\mathbb{I}_i(f_n) &= \mathbb{P}\left(\max_{l\neq k}\min_{x\in B_l} x = 0 \;\cap \min_{j\in B_k, j\neq i} x_j = 1\right) \\
&= \mathbb{P}\left(\max_{l\neq k}\min_{x\in B_l} x = 0\right)\mathbb{P}\left(\min_{j\in B_k, j\neq i} x_j = 1\right) \\
&= \left(1-2^{-|B_j|}\right)^{K-1} 2^{-|B_k|+1} \\
&= \mathbb{P}\left(f_n = 0\right)\frac{2^{-|B_k|+1}}{(1-2^{-|B_j|})} \\
&= \mathbb{P}\left(f_n = 0\right)\frac{2\log_2(n)}{n-\log_2(n)}
\end{aligned}$$

We thus have in this case :

$$C\mathrm{Var}\left(f\right)\frac{\log(n)}{n} \geq \mathbb{I}_i(f)$$

The theorem was extended to all p by Bourgain, Kahn, Kalai, Katznelson and Linial. We give here a second form of the theorem.

Theorem 3.10 (Theorem I [Bourgain et al., 1992]).
There exist $c > 0$ such that if f is a boolean function $f : \Omega_n \to \{0,1\}$ then for any $p \in (0,1)$:

$$\|Inf_p(f)\|_1 \geq c\,Var_p\left(f\right)\log\frac{1}{\|Inf_p(f)\|_\infty}$$

4 Influences in first passage percolation

As we've seen in the previous chapter, boolean theory has been very useful to prove results in classical percolation theory. We wish to leverage the boolean theory but instead for the study of first passage percolation. We will thus restrict ourselves to the boolean setting and take a particular interest on the existence of sharp thresholds and to a lesser extent on bounds on the variance and influences. For this, assume for the moment, the time of passage $(\tau_e)_e$ are issued from a distribution ν of the form $(1-p)\delta_a + p\delta_b$ where $0 < a < b < +\infty$.

4.1 Graph settings

We will restrain ourselves to three different types of graphs :

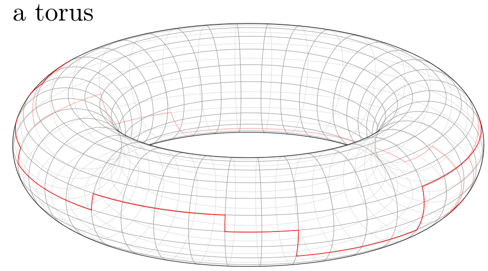
- Torus
- Box
- Point to point

They will all be presented in the d-dimensional case, with each dimension having its own size n_i. Yet we will often restrain ourselves to the 2 dimensional case, with same size n in both dimensions.

4.1.1 The torus

The d-dimensional torus is defined by $V = \prod_{i=1}^{d}(\mathbb{Z}/n\mathbb{Z})$ and edges being $E = \{(x, x+e_i) \mid x \in V, 1 \leq i \leq d\}$ where (e_i) is the canonical base of V seen as a $\mathbb{Z}-module$. It is a very interesting setting, because the group of automorphism of this graph is very large (the group of automorphism is isomorphic to V).

Figure 2: Circumference of a torus



In the torus the time variable of interested will be the circumference across the first coordinate : i.e.

$$T_n = \inf_{\gamma\in Circ_1} \gamma\cdot\tau$$

where $Circ_1$ is the collection of paths whose projection onto the first dimension is surjective and whose starting and ending vertices are equal. Because we are looking for the paths that minimizes the distance, this is equivalent to consider only path that do one loop along the first coordinate.

4.1.2 The box

The d-dimensional box is defined as $V = \prod_{i=1}^{d} [\![1,n]\!]^d$ and the edges being $E = \{(x, x+e_i) | x \in V, 1 \leq i \leq d\}$ where (e_i) is the canonical base of $\mathbb{Z}^d$ seen as a $\mathbb{Z}-module$.

The time variable of interested will be the time of crossing for the first coordinate : i.e.

$$T_n = \inf_{\gamma \in Cross_1} \gamma \cdot \tau$$

where $Croos_1$ is the collection of finite paths such that they start at x and finish at y with $x_1 = 0$ and $y_1 = n$.

4.1.3 Point to point

This is the most classical setting. We consider $V = \mathbb{Z}^d$ and $E = \{(x, x+e_i) | x \in V, 1 \leq i \leq d\}$. The value of interest will be $T_n = T_\omega(0, nx)$ as n grows large. We note the particular case where $x = e_1$ where the normalized limit is called the time constant.

4.2 Estimating Influences in the different settings

The proof of sharp threshold is often done by bounds on the influence. Let's first see the case with more symmetry the torus which should be easier.

4.2.1 Bounding influence on the Torus

We consider the torus $(\mathbb{Z}/n\mathbb{Z})^2$ and T_n the shortest circumference along the first coordinate.

We restrain ourselves to the study the Bernoulli setting in $\mathbb{Z}^2$. Set $\Omega = \Omega_{2n^2}$ and fix an enumeration of the edges that is coherent with $(\mathbb{Z}/n\mathbb{Z})^2 \subset (\mathbb{Z}/m\mathbb{Z})^2$ for $n \leq m$.

Finally, we study the indicator function

$$1_{T_n \geq \gamma n} \qquad \text{with } \gamma \in (a, b).$$

What comes next follows the lines of Chapter VII [Garban and Steif, 2012]. There are two key elements : the first is that, as the distribution is upper and lower bounded by b and a respectively, we obtain very quickly a bound on the length of the geodesic : $|\pi_n| \leq \frac{b}{a} n$.

We can notice that for a given configuration ω, we can first notice that $\mathbb{P}(\nabla_e T_n \neq 0) = 2\mathbb{P}(\nabla_e T_n < 0)$. An edge has a negative gradient only if the edge belongs to the intersection of all geodesic when we set its value to a.

We thus have $\mathbb{P}(\nabla_e T_n \neq 0) \leq 2\mathbb{P}(e \in \pi_n)$. Thus,

$$\sum_e \mathbb{P}(\nabla_e T_n \neq 0) \leq 2\mathbb{P}(e \in \pi_n) = 2\mathbb{E}[|\pi_n|] \leq \frac{2bn}{a}.$$

Because

$$(e \; is \; influencial \; for \; \mathbb{1}_{T_n \geq \gamma n}) \subset (\nabla_e T_n \neq 0)$$

we deduce that

$$\sum_e \mathbb{I}_e(\mathbb{1}_{T_n \geq \gamma n}) \leq \frac{2bn}{a}$$

By an argument of symmetry, all vertical edges and all horizontal edges have the same influence; the torus is invariant by vertical and horizontal translation: the group of isometries being equal to $(\mathbb{Z}/n\mathbb{Z})^2$. As there are $O(n^2)$ vertical and horizontal edges, we have that

$$\mathbb{I}_e(\mathbb{1}_{T_n \geq \gamma n}) = O\left(\frac{1}{n}\right)$$

But to use A.1, one needs to bound the influences for all p. This method does not work, as we have $\mathbb{P}_p(\nabla_e T_n \neq 0) = \frac{1}{1-p}\mathbb{P}(\nabla_e T_n < 0)$. As p goes to 1, all the horizontal edges become very influential for T_n.

To control this, we need a finer bound :

$$(e\ is\ influencial\ for\ \mathbb{1}_{T_n \geq \gamma n}) \subset (\nabla_e T_n \neq 0) \cap (T_n \in [\gamma n - (b-a), \gamma n + (b-a)])$$

Because $\frac{T_n}{n}$ is known to converge in L^2 [Kesten, 1993] and thus, we deduce that $\sigma(T_n) = o(n)$. By continuity (in p) we expect for p large enough, the expectation of T_n to be close enough of bn so that $\mathbb{E}_p[T_n] - k\sigma(T_n) \geq \gamma n + (b-a)$. Making the probability of $(T_n \in [\gamma n - (b-a), \gamma n + (b-a)])$ very small by Tchebichev inequality.

In fact, this approach with the Tchebichev inequality is enough and allows us to prove the existence of sharp threshold even in the case where we have more difficulties to estimate bounds on the influence.

4.2.2 Bounding the influence on the square and point to point

For the case of the square or point to point, we lose this natural symmetry making the bounding of the influences harder (one expects the influence of edges close to the starting or ending point to be quite large), to solve this problem, [Benjamini et al., 2011] use an averaging trick : they introduce a random variable that introduces a random translation of the problem : we replace $d(x,y)$ by $d(x+z, y+z)$ where z is a bounded and centered random variable in $\mathbb{Z}^d$, the influences for this new variable are $\mathbb{I}_e(d(x+z, y+z)) = \mathbb{E}_z[\mathbb{I}_{e-z}(d(x,y))]$, this makes it easier to show that the influences are small, but adapting this technique to our case is not straight forward.

In general, as we have seen previously for the case of the torus, bounding the influence for an edge is tightly linked to the probability that the edge belong to a geodesic.

In the case of the square, we can still hope for this probability to be $O(\frac{1}{n})$ as the most influential edges are those closes to the left and right bounder of the square. For point to point, instead, we expect the influence to be $O(1)$ near the starting and end point. In [Benjamini et al., 2011] they raise the midpoint question : what is the probability that 0 belong to a geodesic from $(-ne_1, ne_1)$. In [Ahlberg and Hoffman, 2016] they give a qualitative answer to this question, proving that this probability goes to 0 as n goes to ∞ and [Dembin et al., 2022] give a more quantitative answer, as they show that under certain conditions over the limit shape, this probability decays as a power of n, $\frac{\log^3 n}{n^{\frac{1}{16}}}$ to be exact. (This result should be enough to make the dismissed proof of [Benjamini et al., 2011] work.)

5 Proving sharp threshold from bounds on the variance

As seen on the previous section, the study of influences is not straight forward in first passage percolation. We contour this problem by instead deriving sharp threshold directly from the bounds on the variance. This approach works for events of the shape $(f_n > \gamma)$.

We will use this approach in two different cases : the first, a generalization of tribes, where more precise calculations are possible, making it easier to deduce the properties required to deduce sharp threshold; the second is in the case of first passage percolation, where we consider the event $(f_n > \gamma n)$. The verification of the necessary conditions will be carried over sections 6 and 7.

5.1 A first example : Tribes

The first example will be for the case of tribes. Here $\Omega_n = \{0,1\}^n$ and we equip it with $\mathbb{P}_p = \otimes^n (p\delta_1 + (1-p)\delta_0)$. We partition the set $[\![1,n]\!]$ into K blocks $(B_j)_{1 \leq j \leq K}$ of the same size $\lfloor \frac{n}{K} \rfloor$ (and we toss the Euclidean remainder into a $(K+1)$th block B_{K+1}). We consider $f_n : \Omega_n \to \mathbb{N}$ defined as follows,

$$f_n = \max_{B_1,\dots,B_K} \sum_{x_i \in B_j} x_i \quad where\ we\ partition \quad [\![1,n]\!] = \bigsqcup_{1 \leq j \leq K} B_j \sqcup B_{K+1}$$

We fix $\alpha \in (0,1)$ and focus on the case where $\#|B_j| = n^\alpha$ and $K = n^{1-\alpha}$. This may not be an integer in which case we take the closest integer (or the ceiling or the floor function). Because the error we make in $O(1)$ and to avoid burdening the notation, we can obviate this approximation.

We assert that

$$\mathbb{E}_p[f_n] = pn^\alpha + O(\sqrt{n^\alpha \ln n})$$

Notice that the integer approximation in $O(1)$ is absorbed by the $O(\sqrt{n^\alpha \ln n})$ supporting our decision to not worry too much about the integer problem. The proof of this assertion can be found on A.2

We thus have

$$\frac{f_n}{n^\alpha} \xrightarrow{L^1} p$$

Because the function f_n is self-bounding (cf page 60 [Boucheron et al., 2013])

$$\mathrm{Var}_p\left(f_n\right) \leq O(n^\alpha).$$

Proposition.
For every $\gamma \in (0,1)$, the indicator function $\mathbb{1}_{\{f_n \geq \gamma n^\alpha\}}$ has a sharp threshold at $p_c(\gamma) = \gamma$.

Proof. Let $p^* = \gamma$.

We fix $\varepsilon > 0$.

Let $p \leq p^* - \varepsilon$ and n large enough so that $\mathbb{E}_{p^*-\varepsilon}[\frac{f_n}{n^\alpha}] \leq \gamma n^\alpha$

Then we have :

$$\begin{aligned}
\mathbb{P}_p\left(f_n \geq \gamma n^\alpha\right) &= \mathbb{P}_p\left(f_n - \mathbb{E}_p\left[f_n\right] \geq \gamma n^\alpha - \mathbb{E}_p\left[f_n\right]\right) \\
&\leq \mathbb{P}_p\left(\left|f_n - \mathbb{E}_p\left[f_n\right]\right| \geq \gamma n^\alpha - \mathbb{E}_p\left[f_n\right]\right) \\
&\leq \frac{\mathrm{Var}_p\left(f_n\right)}{n^{2\alpha}(\gamma - \mathbb{E}_p\left[\frac{f_n}{n^\alpha}\right])^2} \quad \text{by Markov/Tchebyshev's inequality}
\end{aligned}$$

We need to control the term $(\gamma - \mathbb{E}_p\left[\frac{f_n}{n^\alpha}\right])^2$ showing it does not vanish with n.

We have

$$\begin{aligned}
(\gamma - \mathbb{E}_p\left[\frac{f_n}{n^\alpha}\right])^2 &\geq (\gamma - \mathbb{E}_{p^*-\varepsilon}\left[\frac{f_n}{n^\alpha}\right])^2 \\
&= \left(p * -(p^* - \varepsilon) + O(\sqrt{\frac{\ln n}{n^\alpha}})\right)^2 \\
&\geq \varepsilon^2 - \varepsilon O(\sqrt{\frac{\ln n}{n^\alpha}})
\end{aligned}$$

For n larger than some N, we have this $O(\sqrt{\frac{\ln n}{n^\alpha}}) \leq \frac{\varepsilon}{2}$ and thus

$$\left(\gamma - \mathbb{E}_p\left[\frac{f_n}{n^\alpha}\right]\right)^2 \geq \frac{\varepsilon^2}{2}.$$

Finally, we have for $n \geq N$,

$$\mathbb{P}_p\left(f_n \geq \gamma n^\alpha\right) \leq \frac{2\mathrm{Var}_p\left(f_n\right)}{n^{2\alpha}\varepsilon^2}$$

We recall that $\mathrm{Var}_p\left(f_n\right) \leq O(n^\alpha)$ there exists some $N_1 \geq N$ such that for all $n \geq N_1$ and for all $p \leq p^* - \varepsilon$:

$$\mathbb{P}_p\left(f_n \geq \gamma n^\alpha\right) \leq \varepsilon$$

.

Symmetrically, we have for $p \geq p^* + \varepsilon$:

$$\begin{aligned}\mathbb{P}_p\left(f_n < \gamma n^\alpha\right) &= \mathbb{P}_p\left(\mathbb{E}_p\left[f_n\right] - f_n > \mathbb{E}_p\left[f_n\right] - \gamma n^\alpha\right)\\ &\leq \mathbb{P}_p\left(\left|f_n - \mathbb{E}_p\left[f_n\right]\right| > \mathbb{E}_p\left[f_n\right] - \gamma n^\alpha\right)\\ &\leq \frac{\mathrm{Var}_p\left(f_n\right)}{n^{2\alpha}(\mathbb{E}_p\left[\frac{f_n}{n^\alpha}\right] - \gamma)^2} \quad \text{by Tchebyshev inequality}\end{aligned}$$

And thus

$$\begin{aligned}\mathbb{P}_p\left(f_n \geq \gamma n^\alpha\right) &= 1 - \mathbb{P}_p\left(f_n < \gamma n^\alpha\right)\\ &\geq 1 - \frac{\mathrm{Var}_p\left(f_n\right)}{n^{2\alpha}(\mathbb{E}_p\left[\frac{f_n}{n^\alpha}\right] - \gamma)^2}\end{aligned}$$

We conclude the same way. ■

5.2 Existence of sharp threshold in first passage percolation

On section 4 we decided to take interest on edge weights that follow a Bernoulli distribution of type $(1-p)\delta_a + p\delta_b$. Here, we consider more general setting :

Definition 5.1 (Distributions along increasing line segments)**.**
Let F and G be two distributions such that G stochastically dominates F, that means that for all $x, F(x) \geq G(x)$. We consider the family of distributions

$$H(p) = (1-p)F + pG \; for \; p \in [0,1]$$

and focus our on the case where the weights $\tau_e \sim H(p)$.
We will keep using the notation $\mathbb{E}_p$ and $\mathbb{P}_p$ to highlight the dependence on the parameter p of any $\sigma((\tau_e)_e)$-measurable function.

This generalizes our boolean setting that correspond to $F = \delta_a$ and $G = \delta_b$.

Remark.
If we use first passage percolation to model for example the propagation of an epidemic, where vertices are individuals and edges are their interactions : this extra variable λ_e can be interpreted as whether the interaction was "safe" or "unsafe" and having different time of contagion distribution for each case. This could help governments know which percentage of the population they need to follow safe interaction protocols for the spread speed to stay below the current handling capabilities of the disease. The shape of the time constant as a function of p can also be very interesting : we make the conjuncture that this function is convex, meaning that the few individuals that do not follow safety protocols have a larger impact on the spread of the disease that what a linear model would predict.

Let $\mu(p) = \lim_{n\to\infty} \frac{\mathbb{E}_p[T_n]}{n}$

We will assume some results

- $\mu(p)$ establish a bijection between $(p_1, 1)$ and (μ_F, μ_G).
- $\mathbb{E}_p[\frac{T_n}{n}] - \mu(p) \leq o(1)$
- for all p: $\mathrm{Var}_p\left(T_n\right) \leq O(n)$

Let $\gamma \in (\mu_F, \mu_G)$. We would like to know if the function $p \mapsto \mathbb{P}_p\left(T_n \geq \gamma n\right)$ has sharp threshold.

Proposition.
If $F, G \in L^2$ and $F(0) < p_c$, $F(infF) < \vec{p_c}$ the previous assumptions hold, and we have for all $\gamma \in (\mu_F, \mu_G)$, there exist a unique p^* such that $\forall \varepsilon > 0$ there exist $N_1 \geq 0$ such that for all $n \geq N_1$,

$$\begin{aligned} &\textit{for all } p \leq p^* - \varepsilon, \quad \mathbb{P}_p\left(T_n \geq n\gamma\right) \leq \varepsilon \\ \textit{and} \quad &\textit{for all } p \geq p^* + \varepsilon, \quad \mathbb{P}_p\left(T_n \geq n\gamma\right) \geq 1 - \varepsilon \end{aligned}$$

Also p^* is given by the formula $p^* = \mu^{-1}(\gamma)$

Proof. Let $p^* = \mu^{-1}(\gamma)$ i.e. $\mu(p^*) = \gamma$ with $p^* \in (p_1, 1)$.

We fix $\varepsilon > 0$.

By strict inequality of $\mu(p)$, there exist δ such that for $p \leq p^* - \varepsilon : \mu(p) \leq \mu(p^*) - 2\delta$.

For n large enough we have

$$\mathbb{E}_p\left[\frac{T_n}{n}\right] \leq \mu(p) + \delta \leq \mu(p^*) - \delta \qquad \textit{thus} \quad \gamma - \mathbb{E}_p\left[\frac{T_n}{n}\right] \geq \delta$$

We have :

$$\begin{aligned} \mathbb{P}_p(T_n \geq n\gamma) &= \mathbb{P}_p\left(T_n - \mathbb{E}_p[T_n] \geq n\gamma - E[T_n]\right) \\ &\leq \mathbb{P}_p\left(|T_n - \mathbb{E}_p[T_n]| \geq n\delta\right) \end{aligned}$$

And Tchebychev inequality gives us

$$\mathbb{P}_p(|T_n - \mathbb{E}_p[T_n]| \geq n\delta) \leq \frac{\mathrm{Var}_p\left(T_n\right)}{n^2\delta^2} \leq \frac{1}{\delta^2} O(\frac{1}{n})$$

So by choosing n large enough so that $\frac{1}{\delta^2} O(\frac{1}{n}) \leq \varepsilon$ we obtain the desired result.

The case $p \geq p^* + \varepsilon$ is symmetrical, except for the fact that because of the subadditive theorem, we know that for all n, $\mathbb{E}_p\left[\frac{T_n}{n}\right] \geq \mu(p)$, so it suffices to fix δ such that for all $p \geq p^* + \varepsilon,\ \mu(p) \geq \mu(p^*) + \delta$ which allows writing :

$$\mathbb{P}_p(T_n \leq n\gamma) \leq \mathbb{P}_p(|T_n - \mathbb{E}[T_n]| \geq \delta).$$

We conclude the same way. ∎

5.3 Some finer control : The study of $N(\varepsilon)$ and the sharpness of the threshold.

All the previous demonstration, we argued that we could choose n large enough for the ε given. Now, one could ask the reverse problem : given n, what is the best ε that one can choose, that means that for given n, we would like to know how sharp of a threshold we have.

Proposition.
Under the assumptions of the previous theorem, the function $p \mapsto \mathbb{P}_p\left(T_n \geq \gamma\right)$ has a threshold of size $O(n^{-\frac{1}{3}})$.

Let $p(n) \in (0,1)$ be such that $\mathbb{E}_{p(n)}\left[\frac{T_n}{n}\right] = \gamma$. This can always be done in the Bernoulli setting as $\mathbb{E}_{p(n)}\left[\frac{T_n}{n}\right] \geq \mu(p)$ and have they both are continuous functions with the same end points. In the general case, this may be impossible for small values of n and γ very close to μ_F. Let's skip this complication for the moment.

Because of Corollary 7.6 we have

$$|\mathbb{E}_{p(n) - \frac{d}{\sqrt[3]{n}}}\left[\frac{T_n}{n}\right] - \mathbb{E}_{p(n)}\left[\frac{T_n}{n}\right] \geq c\frac{d}{\sqrt[3]{n}}$$

By the same Tchebychev inequality as previously (for $p = p(n) - \frac{d}{\sqrt[3]{n}}$) and $\delta = c\frac{d}{\sqrt[3]{n}}$ we have

$$\mathbb{P}_{p(n)-\frac{d}{\sqrt[3]{n}}}(T_n \geq \gamma n) \leq \frac{1}{c^2 d^2 \sqrt[3]{n}} = O(\frac{1}{\sqrt[3]{n}})$$

The other side is completely symmetrical. This shows that we go from a $O(\frac{1}{\sqrt[3]{n}})$ to $1 - O(\frac{1}{\sqrt[3]{n}})$ over a window of size $O(\frac{1}{\sqrt[3]{n}})$.

Now back to the complication : we can on a first approximation, think of $\mathbb{E}_p\left[\frac{T_n}{n}\right]$ as a vertical shift of $\mu(p)$. The size of this vertical shift will determine for which γ fixing $p(n)$ as before may be impossible.

For this, we use the result from [Damron and Kubota, 2013] which states that

$$\mathbb{E}_p\left[\frac{T_n}{n}\right] - \mu(p) = O(\sqrt{\frac{\ln n}{n}}).$$

Because this bound is smaller than the size of the threshold we have deduced, this means that replacing $p(n)$ by p^* in the proof will not have a significant effect.(If we use the conjecture that the bound on the variance is $O(n^{\frac{2}{3}})$ the bound on the size of the threshold improves to $O(n^{-\frac{4}{9}})$).

Remark.
As we have seen, the proof for both tribes and first passage percolation is quite similar, and we could state a more general result : We suppose the existence of a collection of measures $\mathbb{P}_p$ for $p \in [0, 1]$.

Let $(f_n)_n$ be a collection of random variables $\mathbb{P}_p-$measurable.

Suppose the existence of $\kappa \in \mathbb{R}$ such that

$$\frac{f_n}{n^\kappa} \xrightarrow{L^2(\mathbb{P}_p)} \mu(p)$$

and $\mu : [0, 1] \to \mathbb{R}$ is a strictly increasing function, then the event $(f_n \geq \gamma n^\kappa)$ has a sharp threshold behavior.

6 Bound of the variance in FFP

On section 5.2 we made the assumption that $\mathrm{Var}_p(T_n) \leq O(n^\beta)$. Kesten proved in [Kesten, 1993] that the variance was a big $O(n)$, thus $\beta = 1$. The most accepted conjecture is that $\beta = \frac{2}{3}$ in dimension 2. We will present the proof of Kesten. This proof can be adapted for the different graph setting presented in 4.1. It will be written for the case of $T_n = T(0, ne_1)$.

Proposition.
Assume the weight of the edges are sampled following a distribution $F \in L^2$ then :

$$Var(T_n) \leq C_1 n$$

We enumerate $E = \{e_1, \ldots, e_n, \ldots\}$. For this proof, we will take Ω as follows :

$$\Omega = \prod_{s=1}^{\infty} \mathbb{R}_+ \text{ and } \tau_{e_i}(\omega) = \omega_i$$

that we equip with the measure $\nu = \bigotimes_{s=1}^{\infty} F$

We will also note

$$\Omega_k = \prod_{s=k}^{\infty} \mathbb{R}_+ \text{ and } \nu_k = \bigotimes_{s=k}^{\infty} F$$

We build the following filtration :

$$\mathcal{F}_k = \sigma(\omega_1, \ldots, \omega_k) \quad k \geq 0.$$

We use the Martingale representation of $T_n - \mathbb{E}[T_n]$:

$$T_n - \mathbb{E}[T_n] = \sum_{k=1}^{\infty} \mathbb{E}[T_n|\mathcal{F}_k] - \mathbb{E}[T_n|\mathcal{F}_{k-1}]$$

This representation is valid because

$$\begin{aligned} M_l := & \sum_{k=1}^{l} \mathbb{E}[T_n|\mathcal{F}_k] - \mathbb{E}[T_n|\mathcal{F}_{k-1}] \\ &= \mathbb{E}[T_n|\mathcal{F}_l] - \mathbb{E}[T_n] \\ &= \mathbb{E}\left[T_n - \mathbb{E}(T_n) \mid \mathcal{F}_l\right] \end{aligned}$$

is a (closed) $(\mathcal{F}_l)$-martingale and converges almost surely and in L^1 to $T_n - \mathbb{E}[T_n]$ by Doob's convergence theorem.

The increments of (M_l) are denoted by

$$\Delta_k = \mathbb{E}[T_n|\mathcal{F}_k] - \mathbb{E}[T_n|\mathcal{F}_{k-1}]$$

We want to control

$$\mathbb{E}\left[(T_n - \mathbb{E}[T_n])^2\right] = \mathbb{E}\left[\left(\sum_{k=1}^{\infty} \mathbb{E}[T_n|\mathcal{F}_k] - \mathbb{E}[T_n|\mathcal{F}_{k-1}]\right)^2\right] = \sum_{k=1}^{\infty} \mathbb{E}[\Delta_k^2]$$

The last equality comes from the fact that the increments are orthogonal, we can thus write $\mathbb{E}[M_l] = \sum_{k=1}^{l} \mathbb{E}[\Delta_k^2]$ for all l. The martingale converging in L^1 thus in L^2 the limit when l goes to infinity holds.

To control this decomposition, we will try to have a finer control over

$$\mathbb{E}[\Delta_k^2|\mathcal{F}_{k-1}].$$

We define for $\omega, \sigma \in \Omega$ their merge at level k as follows:

$$[\,\omega, \sigma\,]_k = [\,\omega, \sigma\,]_k = (\omega_1, \dots, \omega_k, \sigma_{k+1}, \dots)$$

We thus have

$$\mathbb{E}[T_n|\mathcal{F}_k](\omega) = \int_{\Omega_{k+1}} f([\,\omega, \sigma\,]_k)\nu_{k+1}(d\sigma)$$

Because $[\,\omega, \sigma\,]_k$ depends on only in the first k coordinates of ω, we can add an extra layer of integration :

$$\mathbb{E}[T_n|\mathcal{F}_k](\omega) = \int_{\Omega_k} f([\,\omega, \sigma\,]_k)\nu_k(d\sigma)$$

Thus

$$\Delta_k = \int_{\Omega_k} f([\,\omega, \sigma\,]_k) - f([\,\omega, \sigma\,]_k - 1)\nu_k(d\sigma)$$

We thus want to bound $|f([\,\omega, \sigma\,]_k - f([\,\omega, \sigma\,]_k - 1))|$

This is the change in the value of T_n when τ_{e_k} goes from σ_k to ω_k. We have a trivial bound that is $|f([\,\omega, \sigma\,]_k - f([\,\omega, \sigma\,]_k - 1))| \leq |\sigma_k - \omega_k|$

We can improve this bound by arguing that for this value to change, e_k must be on a geodesic. So :

$$|f([\,\omega, \sigma\,]_k - f([\,\omega, \sigma\,]_k - 1))| \leq |\sigma_k - \omega_k| \mathbb{1}_{e_k \in \pi_n([\omega,\sigma]_{k-1}) \cup \pi_n([\omega,\sigma]_k)}$$

We write $I_k(\omega, \sigma) := \mathbb{1}_{e_k \in \pi_n([\omega,\sigma]_{k-1}) \cup \pi_n([\omega,\sigma]_k)}$ to lighten the notation. We thus have :

$$\begin{aligned} \mathbb{E}[\Delta_k^2|\mathcal{F}_{k-1}] &= \mathbb{E}\left[\left(\int_{\Omega_k} f([\,\omega, \sigma\,]_k) - f([\,\omega, \sigma\,]_k - 1)\nu_k(d\sigma)\right)^2 \middle| \mathcal{F}_{k-1}\right] \\ &\leq \mathbb{E}\left[\left(\int_{\Omega_k} |\sigma_k - \omega_k| I_k(\omega, \sigma)\nu_k(d\sigma)\right)^2 \middle| \mathcal{F}_{k-1}\right] \\ \text{Jensen inequality} \qquad &\leq \mathbb{E}\left[\int_{\Omega_k} |\sigma_k - \omega_k|^2 I_k(\omega, \sigma)\nu_k(d\sigma) \middle| \mathcal{F}_{k-1}\right] \end{aligned}$$

Inside the conditional expectation we have a function of $(\omega_1, \dots, \omega_k)$. Thus, the conditional expectation is just the integration with respect to ω_k :

$$
\begin{aligned}
\mathbb{E}\left[\int_{\Omega_k} |\sigma_k - \omega_k|^2 I_k(\omega,\sigma)\nu_k(d\sigma)\middle|\mathcal{F}_{k-1}\right] &= \int_{\mathbb{R}_+}\int_{\Omega_k} |\sigma_k - \omega_k|^2 I_k(\omega,\sigma)\nu_k(d\sigma)F(d\omega_k) \\
\text{using } \nu_k = F\times\nu_{k+1} \qquad &= \int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\int_{\Omega_{k+1}} |\sigma_k - \omega_k|^2 I_k(\omega,\sigma)\nu_{k+1}(d\sigma)F(d\sigma_k)F(d\omega_k) \\
&= \int_{\Omega_{k+1}}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} |\sigma_k - \omega_k|^2 I_k(\omega,\sigma)F(d\sigma_k)F(d\omega_k)\nu_{k+1}(d\sigma) \\
\text{by noticing the symmetry of the integrand } &= 2\int_{\Omega_{k+1}}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} |\sigma_k - \omega_k|^2 I_k(\omega,\sigma)\mathbb{1}_{\sigma_k\le\omega_k}F(d\sigma_k)F(d\omega_k)\nu_{k+1}(d\sigma)
\end{aligned}
$$

By exploiting the symmetry, we can focus on the case where the time passage for edge k lowers from ω_k to σ_k. So if the edge e_k was already in a geodesic, it will remain in a geodesic. This allows to simplify

$$\forall \sigma_k \le \omega_k \quad I_k(\omega,\sigma) = \mathbb{1}_{e_k\in\pi_n([\,\omega,\sigma\,]_k-1)}$$

We can also bound $(\sigma_k - \omega_k)^2 \le \omega_k^2$ (this is like bounding $\mathrm{Var}\,(\omega) \le \mathbb{E}[\omega^2]$, so we lose in the constant, I think). We get :

$$
\begin{aligned}
\mathbb{E}[\Delta_k^2|\mathcal{F}_{k-1}] &\le 2\int_{\Omega_{k+1}}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \omega_k^2 \mathbb{1}_{e_k\in\pi_n([\,\omega,\sigma\,]_k-1)}\mathbb{1}_{\sigma_k\le\omega_k}F(d\sigma_k)F(d\omega_k)\nu_{k+1}(d\sigma) \\
&\le 2\int_{\Omega_k} \mathbb{1}_{e_k\in\pi_n([\,\omega,\sigma\,]_k-1)}\nu_k(d\sigma)\int_{\mathbb{R}_+}\omega_k^2 F(d\omega_k) \\
&= \mathbb{P}\left(e_k\in\pi_n|\mathcal{F}_{k-1}\right)\int_{\mathbb{R}_+} x^2F(dx)
\end{aligned}
$$

Finally, by writing :

$$\mathrm{Var}\,(T_n) = \mathbb{E}\left[(\sum_{k=1}^{\infty}\Delta_k)^2\right]$$

We have :

$$
\begin{aligned}
\mathrm{Var}\,(T_n) &= \mathbb{E}\left[(\sum_{k=1}^{\infty}\Delta_k)^2\right] \\
&= \sum_{k=1}^{\infty}\mathbb{E}[\Delta_k^2] \quad \text{by orthogonality of the increments and continuity of } \mathbb{E} \\
&= \sum_{k=1}^{\infty}\mathbb{E}[\mathbb{E}[\Delta_k^2 \mid \mathcal{F}_{k-1}] \\
&\le \mathbb{E}\left[\sum_{k=1}^{\infty}\mathbb{P}\left(e_k\in\pi_n|\mathcal{F}_{k-1}\right)\right]\int_{\mathbb{R}_+}x^2F(dx) \\
&\le C\mathbb{E}[|\pi_n|] = O(n)
\end{aligned}
$$

Which concludes the proof.

This justifies our hypothesis that $\mathrm{Var}\,(T_n) = O(n)$.

7 Strict monotonicity of the time constant[van den Berg and Kesten, 1993]

Our second hypothesis was that μ was strictly increasing. This is a harder result to prove. This was first proved by Kesten and Van den Berg in [van den Berg and Kesten, 1993].

The proof they use will be very similar to ours. They prove this monotonicity for a richer order in the set of distribution that is second order stochastic domination (also known as being less variable). For our purpose, first order stochastic domination will suffice.

7.1 Definition of an order over the distributions

Definition 7.1.
We say G dominates F if for all $x, F(x) \geq G(x)$. This means that the distribution of G has its weight closer to $+\infty$ than F.

Proposition.
If G dominates F then there exists (X, Y) with $X \sim F$ and $Y \sim G$ such that $X \leq Y$ a.s.

Proof. Let F^{-1} and G^{-1} be the inverse of the cumulative distribution. We have $G^{-1} \geq F^{-1}$ for all x. Thus, if we take $\xi \sim \mathcal{U}([0,1])$, we have $G^{-1}(\xi) \geq F^{-1}(\xi)$ a.s. and $X := F^{-1}(\xi) \sim F$, $Y := G^{-1}(\xi) \sim G$. ■

Definition 7.2.
We say a distribution F is more variable than a distribution G if for all increasing concave function $\varphi \in L^1(F) \cap L^1(F)$ we have :

$$\mathbb{E}_F[\varphi(X)] \leq \mathbb{E}_G[\varphi(X)]$$

Proposition.
We have a nice characterization for this in the case both distribution have finite mean : F is more variable than G if and only if for all x

$$\int_{-\infty}^{x} G(y)dy \leq \int_{-\infty}^{x} F(y)dy$$

The proof can be found on the Appendix A.3.

The reason why second order domination is of interest in first passage percolation is because T_n is a concave increasing function from $E \to \mathbb{R}_+$. Moreover, we can approximate T_n by

$$T_n^K := \text{The distance of the shortest path from 0 to n inside a compact K}$$

which is also a collection of concave increasing functions. Using both facts, it becomes immediate that if G dominates (second order) F, then $\mu(G) \geq \mu(F)$.

A harder result, is that if $F \neq G$ then this inequality becomes strict. This result is not that intuitive, for instance if some distribution F is unbounded, one may think that larger weights will be avoided, and thus the time passage for X and $X \wedge M$ were M is some constant large enough will be of the same order of magnitude. But this theorem states the opposite as $\mu(F \wedge M)$ is more variable than $\mu(F)$ implying that a non-negligible fraction of edges of weight larger than M are used. In fact, the proof of this theorem relies on a more powerful result :

Theorem 7.3.
We denote π_n a geodesic for T_n. Let $X \sim F$ be the marginal distribution of the edges. Assume $F(0) < p_c$ (and $F(\inf F) < \vec{p_c}$). Let $I \in \mathcal{F}$ such that $\mathbb{P}(X \in I) > 0$.
Then

$$\liminf \frac{1}{n} \mathbb{E}\left[\sum_e \mathbb{1}_{e \in \pi_n} \mathbb{1}_{X_e \in I}\right] > 0$$

We can also introduce the empirical distribution of weights along the optimal path :

$$\Pi_n = \frac{1}{|\pi_n|} \sum_{e \in \pi_n} \delta_{X_e}$$

The limiting distribution (assuming it exists) will be absolutely continuous with respect to F.

7.2 Proof of increasing time

This proof is inspired on [van den Berg and Kesten, 1993, Gorski, 2022].

Theorem 7.4.
Let F,G be two non-negative distributions with $G \in L^1$ and $G(0) < p_c$ and $G(r) < \vec{p_c}$ where $r = \inf Supp(G)$ and $p_c, \vec{p_c}$ denote the critical probability for percolation and oriented percolation respectively.
If G strictly dominates F and G then

$$\mu(F, x) < \mu(G, x).$$

This theorem can be directly deduced from Theorem 7.3 (it is in fact a corollary of the previous theorem, but because it is an important result, we will present it as a theorem).

Assume Theorem 7.3 is true for the moment. Assume we have two distributions such that G dominates F. We can thus build $X \sim F$ and $Y \sim G$ such that $X \leq Y$ almost surely and $\mathbb{P}(X < Y) > 0$.

We note T_n the time passage from 0 to nx when the edges have weights issued from the distribution G and $\widetilde{T_n}$ when the weights are issued from F.

Corollary 7.5 (of Theorem 7.3)**.**
Under the previous theorem hypothesis (and the assumption of the existence of geodesics), there exists a constant $c > 0$ such that for n large enough :

$$\mathbb{E}[T_n - \widetilde{T_n}] \geq cn$$

Let I, δ denote a measurable set and a positive constant such that $Y - X > \delta$ for all $\omega \in I$ and such that $\mathbb{P}(I) > 0$. Such pairing I, δ must exist because

$$(Y > X) = \bigcup_n (Y \geq X + \frac{1}{n}).$$

If for all $n \in \mathbb{N}$ the right set has null measure, then the left one is also of null measure. Thus, $X = Y$ almost surely, which is excluded.

We then have

$$\begin{aligned}
\mathbb{E}[T_n - \widetilde{T_n}] &\geq \mathbb{E}[Y \cdot \pi_n - X \cdot \pi_n] \quad \text{where } \pi_n \text{is a geodesic for } Y. \\
&= \mathbb{E}[(Y - X) \cdot \pi_n] \\
&= \mathbb{E}[\mathbb{E}[Y - X | Y] \cdot \pi_n] \\
&\geq \mathbb{E}[\delta \mathbb{1}_{X_e \in I} \cdot \pi_n] \\
&= \delta \mathbb{E}[\# |\{e \in E \mid e \in \pi_n, \ X_e \in I\}|] \\
&\geq \delta \varepsilon n \quad \text{by Theorem 7.3 for } n \text{ large enough} \\
&\geq cn \quad \text{for some } c > 0 \text{ that does not depend on } n.
\end{aligned}$$

Which proves the corollary, and by dividing by n and making n tend to $+\infty$ we obtain,

$$\mu(G) - \mu(F) \geq c > 0$$

which proves theorem 7.4.

□

Proof of Theorem 7.3.
We will prove the theorem in the case where T_n is the point to point passage time from 0 to nx. The square and torus case can be handled with minors adjustments.

Let $B(z, L)$ denote the L^1 ball centered in z. We tile $V(\mathbb{Z}^d)$ with balls of radius $L \in \mathbb{N}$. (we consider an edge is inside the ball if both extremities are inside the ball). [a] This gives us a partition of

$$E = \bigsqcup_{q \in \mathcal{Q}} E \cap B(z_q, L) \tag{2}$$

Let $k \in 2\mathbb{N}+1$ be an integer that we will fix later. We choose k to be odd so that

$$B(z_q, kL) = \bigsqcup_{\|z'_q - z_q\| \leq (k-1)L} B(z'_q, L). \tag{3}$$

Let $A_L(z)$ denote the event that any path from $B(z,L)$ to $\partial B(z,kL)$ picks up weight at least $(k-1)L(r+\varepsilon)$ where $r = \inf F$.

Thanks to classical percolation theory[b], the distance from $B(z,L)$ to $\partial B(z,kL) \sim C(k-1)L$ with $C > r$ when L goes to infinity. We have for $\varepsilon > 0$ such that $r + \varepsilon < C$

$$\mathbb{P}\left(A_L(z)\right) \to 1 \quad \text{as} \quad L \to \infty.$$

We will now prove that when this probability is close enough to 1, the amount of tiles where A_L doesn't happen will become very small. More exactly :

Lemma (1)**.**
For all fixed $k \in \mathbb{N}$ and for all L large enough, there exist $\theta(L)$ and $\eta(L)$ such that

$$\mathbb{P}\left(\exists \gamma : x \to y \ \textit{ visiting at most } \theta \|x-y\|_1 \textit{ distinct } B(z_q, L) \textit{ such that } A_L(z) \textit{ holds}\right) \leq e^{-\eta\|y-x\|_1}$$

Proof of Lemma 1
Let's fix γ to be a path from x to y that visits M different tiles $B(z_q, L)$ and never visits a tile twice. We denote $i \in \mathcal{I} \subset \mathcal{Q}$ the index of the tiles that are visited by γ. We can fix a subindex $\mathcal{J} \subset \mathcal{I}$ of at least $m = \frac{M}{Ck^2}$ elements such that the boxes $B(z_j, kL)$ are disjoint.[c]

$$\begin{aligned}
&\mathbb{P}\left(\text{there are at most } \theta\|x-y\|_1 \, i \in \mathcal{I} \text{ such that } A_L(z_i) \textit{ holds}\right) \\
\leq&\mathbb{P}\left(\text{there are at most } \theta\|x-y\|_1 \, j \in \mathcal{J} \text{ such that } A_L(z_j) \textit{ holds}\right) \\
=&\mathbb{P}\left(\exists \mathcal{K} \subset \mathcal{J} \text{ of cardinal } m - \theta\|x-y\|_1 \text{ such that for all } j \in \mathcal{K}, A_L(z_j) \text{ doesn't hold}\right) \\
\leq&\mathbb{P}\left(\bigcup_{\mathcal{K} \subset \mathcal{J} \text{ of cardinal } m-\theta\|x-y\|_1} \text{for all } j \in \mathcal{K}, A_L(z_j) \text{ doesn't hold}\right) \\
\leq& \sum_{\mathcal{K} \subset \mathcal{J} \text{ of cardinal } m-\theta\|x-y\|_1} \mathbb{P}\left(\bigcap_{j \in \mathcal{K}} \overline{A_L(z_j)}\right) \\
=& \sum_{\mathcal{K} \subset \mathcal{J} \text{ of cardinal } m-\theta\|x-y\|_1} \prod_{j \in \mathcal{K}} \mathbb{P}\left(\overline{A_L(z_j)}\right) \quad \text{by independence} \\
=&\binom{m}{\theta\|x-y\|_1} \mathbb{P}\left(\overline{A_L(z)}\right)^{m-\theta\|x-y\|_1}
\end{aligned}$$

We say that a subindex $\mathcal{I} \subset \mathcal{Q}$ is path compatible (from x to y) if there exists a path (from x to y) γ that visits exactly $B(z_i, L), i \in \mathcal{I}$ and never visits a tile twice. There are at most D^M subindex of cardinal M that are path compatible from x to y where $D = 8$ is the number of neighboring tiles a tile $B(z,L)$ has. Indeed, let γ be a path from x to y, let $\mathcal{I}$ be the subindex of the tiles that this path visits. We sort $\mathcal{I}$ in the order that γ visit the tiles. There are D possible choices for the first second element and at most $D-1$ possible choices for the other elements. If we consider $\mathbb{G}/\sim$ where $\sim$ is the equivalence relationship induced by the partition $\mathbb{G} = \bigsqcup B(z_q, L)$, this corresponds to studying the path from $B(z_x, L)$ to $B(z_y, L)$ with no loops.

For any path γ, we have the induced path $\bar{\gamma}$ of the tiles that he visits (each tile seen as a vertex). If $\tilde{\gamma}$ uses at most $\theta \|x-y\|_1$ tiles $B(z_q, L)$ such that $A_L(z_q)$ holds, then $\tilde{\bar{\gamma}}$ that is the path obtained from $\bar{\gamma}$ where loops have been removed will also satisfy the property. Thus,

$$\begin{aligned}
&\exists\gamma : x \to y \ \text{ visiting at most } \theta\|x-y\|_1 \textit{ distinct } B(z_q, L) \text{ such that } A_L(z_q) \textit{ holds}\\
=&\exists\gamma : x \to y \text{ visiting at most once each } B(z_q, L) \text{ and visiting at most } \theta\|x-y\|_1 \textit{ distinct } B(z_q, L)\\
&\qquad\qquad \text{such that } A_L(z_q) \textit{ holds}
\end{aligned}$$

We also have that any path from x to y has to visit at least $\frac{\|x-y\|_1}{2L} - 2$ tiles $B(z_q, L)$. We omit the -2 as what's important is that the number of tiles that are visited is an $\Omega(\|x-y\|_1)$.

$$\begin{aligned}
&\mathbb{P}\left(\exists\gamma : x \to y \ \text{ visiting at most } \theta\|x-y\|_1 \textit{ distinct } B(z_q, L) \text{ where } A_L(z_q) \textit{ holds}\right)\\
=&\mathbb{P}\left(\exists\gamma : x \to y \text{ visiting at most once each } B(z_q, L) \text{ and at most } \theta\|x-y\|_1 \textit{ distinct } B(z_q, L) \text{ where } A_L(z_q) \textit{ holds}\right)\\
\leq&\mathbb{P}\left(\bigcup_{M\geq\|x-y\|/2L} \exists\gamma \text{ visiting } M \text{ different } B(z_q, L) \text{ each once, such that at most for } \theta\|x-y\|_1\, A_L(z_q) \textit{ holds}\right)\\
\leq&\sum_{M\geq\|x-y\|/2L} \mathbb{P}\left(\exists\mathcal{I} \subset \mathcal{Q} \text{ of cardinal } M \text{ that is path compatible such that } A_L(z_i) \text{ holds for at most ...}\right)\\
\leq&\sum_{M\geq\|x-y\|/2L} \mathbb{P}\left(\bigcup_{\mathcal{I}\subset\mathcal{Q} \text{ of cardinal } M \text{ that is path compatible}} A_L(z_i) \text{ holds for at most } \theta\|x-y\|_1\ i \in \mathcal{I}\right)\\
\leq&\sum_{M\geq\|x-y\|/2L} D^M \binom{m}{\theta\|x-y\|_1} \mathbb{P}\left(\overline{A_L(z)}\right)^{m-\theta\|x-y\|_1}\\
\leq&\mathbb{P}\left(\overline{A_L(z)}\right)^{-\theta\|x-y\|_1} \sum_{M\geq\|x-y\|/2L} ((2\mathbb{P}\left(\overline{A_L(z)}\right))^{\frac{1}{Ck^2}} D)^M
\end{aligned}$$

We fix L large enough such that $\mathbb{P}\left(\overline{A_L(z)}\right)$ is small enough for $(2\mathbb{P}\left(\overline{A_L(z)}\right))^{\frac{1}{2Ck^2}} D < 1$.

$$\begin{aligned}
&= \mathbb{P}\left(\overline{A_L(z)}\right)^{-\theta\|x-y\|_1} ((2\mathbb{P}\left(\overline{A_L(z)}\right))^{\frac{1}{Ck^2}} D)^{\|x-y\|_1/2L} \frac{1}{1-(2\mathbb{P}\left(\overline{A_L(z)}\right))^{\frac{1}{Ck^2}} D}\\
&= \mathbb{P}\left(\overline{A_L(z)}\right)^{(\frac{1}{4CLk^2}-\theta)\|x-y\|_1} \left(2^{\frac{1}{Ck^2}} \mathbb{P}\left(\overline{A_L(z)}\right)^{\frac{1}{2Ck^2}} D\right)^{\frac{1}{2L}} \frac{1}{1-(2\mathbb{P}\left(\overline{A_L(z)}\right))^{\frac{1}{Ck^2}} D}\\
&= C\mathbb{P}\left(\overline{A_L(z)}\right)^{(\frac{1}{4CLk^2}-\theta)\|x-y\|_1} \qquad \text{with } C > 0
\end{aligned}$$

We now fix θ such that $\frac{1}{4CLk^2} - \theta > 0$ and because $\mathbb{P}\left(\overline{A_L(z)}\right) < 1$ we have indeed

$$\mathbb{P}\left(\exists\gamma : x \to y \ \text{ visiting at most } \theta\|x-y\|_1 \textit{ distinct } B(z_i, L) \text{ such that } A_L(z) \textit{ holds}\right) \leq e^{-\eta\|x-y\|_1}$$

□

Remark.
Because the decay is exponential, if instead of considering point to point percolation we consider another model (like the square) were this event can be seen as the reunion over all the x on the left side and y on the right side, which is a polynomial reunion the decay will still be exponential for other models.

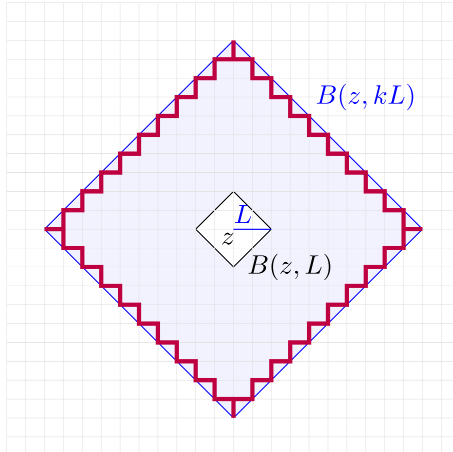


Figure 3: Tile used to partition $\mathbb{Z}^d$

Let $C_L(z)$ be the event that $\omega_e \in I$ for all $e \in B(z, kL)$ with one vertex on the boundary $\partial B(z, kL)$ (thick purple edges on Figure3) and $\omega_e < r + \frac{\varepsilon}{2}$ for all others $e \in B(z, kL) \setminus B(z, L)$

Note that if a geodesic visits $B(z, L)$ and $A_L(z)$ occurs, then, if we resample only the weights inside $B(z, kL) \setminus B(z, L)$ such that $C_L(z)$ occurs, then the geodesic will pass through $B(z, kL)$. (We say pass by, means enter and exit). Indeed, consider the geodesic π_ω before resampling, the weight she picks up crossing through $B(z, kL)$ is at least $2d(B(z, L), \partial B(z, kL)) + \pi \mathbb{1}_B(z, L) \cdot \omega$. After the resampling, this same path cost $2(I + (r + \frac{\varepsilon}{2})(k-1))L + \pi \mathbb{1}_B(z, L) \cdot \omega$. We thus fix $k(\varepsilon)$ such that $(k-1)(r+\varepsilon)L > 2 \max I$ to ensure that $\pi_\omega \cdot \tilde{\omega} \leq \pi_\omega \cdot \omega$. Because the paths that don't go through $B(z, kL)$ have the same time passage as before, we have

$$\pi_\omega \cdot \tilde{\omega} \leq \pi_\omega \cdot \omega \leq \gamma \cdot \omega = \gamma \cdot \tilde{\omega}$$

for all γ not going through $B(z, kL)$. This proves that the optimal path goes through $B(z, kL)$.

□

[a]For people who would want an explicit tilling : we consider in $\mathbb{Z}^2$ the following tilling $z_{u,v} = (uL, vL)$ for $u, v \in \mathbb{Z}^2$ such that $u + v \in 2\mathbb{Z}$

[b]Here is where the hypothesis of $G(r) \leq \vec{p_c}$ is used.

[c]To be exact $m = \frac{M}{(2k-1)^2}$ but saying that we have $m = O(\frac{M}{k^d})$ is enough and easier to prove.

Resampling argument

Claim 1 : Let π_n be a geodesic.

$$\mathbb{P}\left(\pi_n \text{ visits } B_L(z) \text{ and } \omega_e \in I \text{ for some } e \in \pi_n \cap B(z, kL)\right) \geq c\mathbb{P}\left(\pi_n \text{ visits } B(z, L) \text{ and } A_L(z) \text{ occurs.}\right)$$

Proof. Let $\omega^\star$ be a configuration identical to ω outside $B(z, kL) \setminus B(z, L)$ and independent of ω on $B(z, kL) \setminus B(z, L)$. Due to the observation above,

$$\begin{aligned} &\{\omega \in A_L(z)\} \cap \{\pi_n(\omega) \text{ visits } BL(z)\} \cap \{\omega^* \in C_L(z)\} \\ &\subset \{\pi_n(\omega^*) \text{ visits } B(z, kL)\} \cup \{\exists e \in \pi_n(\omega^*) \cap B(z, kL) : \omega^*_e \in I\} \end{aligned}$$

Due to independence we obtain :

$$\begin{aligned} &\mathbb{P}\left(\pi_n \text{ visits } B(z, kL) \text{ and } \omega_e \in I \text{ for some } e \in \pi_n \cap B(z, kL)\right) \\ &\geq \mathbb{P}\left(\omega \in A_L(z) \cap \pi_n(\omega) \text{ visits } B(z, L)\right) \mathbb{P}\left(\omega^* \in C_L(z)\right) \\ &:= c\mathbb{P}\left(\omega \in A_L(z) \cap \pi_n(\omega) \text{ visits } B(z, L)\right) \end{aligned}$$

□

Claim 2:

$$\mathbb{E}[\# \left|\{e \in \pi_n : \omega_e \in I\}\right|] \geq cn$$

Proof. We start by noticing that

$$\sum_q \sum_{e \in E \cap B(z_q, kL)} f(e) = k^2 \sum_{e \in E} f(e).$$

This comes from the partitions 2 and 3. We have :

$$\begin{aligned}
\mathbb{E}[\#\{e \in \pi_n : \omega_e \in I\}] &= \mathbb{E}\left[\sum_{e \in E} \mathbb{1}_{e \in \pi_n} \mathbb{1}_{\omega_e \in I}\right] \\
&= \frac{1}{k^2} \mathbb{E}\left[\sum_q \sum_{e \in E \cap B(z, kL)} \mathbb{1}_{e \in \pi_n} \mathbb{1}_{\omega_e \in I}\right] \\
&\geq \frac{1}{k^2} \sum_q \mathbb{P}\left(\pi_n \text{ visits } B(z_q, kL) \text{ and } \omega_e \in I \text{ for some } e \in \pi_n \cap B(z_q, kL)\right) \\
&\geq \frac{c}{k^2} \sum_q \mathbb{P}\left(\omega \in A_L(z_q) \cap \pi_n(\omega) \text{ visits } B(z_q, L)\right) \qquad \text{by claim 1} \\
&= \frac{c}{k^2} \mathbb{E}\left[\# \left|\{q \in \mathcal{Q} | \pi_n(\omega) \text{ visits } B(z_q, L) \text{ and } A_L(z_q) \textit{ holds}\}\right|\right] \\
&\geq \frac{c}{k^2} \mathbb{E}\left[\min_{\gamma : 0 \to nx} \# \left|\{q \in \mathcal{Q} | \gamma \text{ visits } B(z_q, L) \text{ and } A_L(z_q) \textit{ holds}\}\right|\right] \\
&\geq \frac{c}{k^2} \theta n (1 - e^{-\eta n \|x\|_1}) \quad \text{by the lemma 2 for } x = 0 \text{ and } y = nx
\end{aligned}$$

■

Remark.
This theorem makes some supplementary assumptions, some of them are necessary while others are just convenient for the proof.

- $F, G \in L^1$ *is a commodity assumption : it was shown that the time constant can be defined when* $F \notin L^1$ *(insert ref here): the convergence is weaker, as it only occurs in probability and makes the study of the time constant harder.*
- $G(0) < p_c$ *is mandatory as if* $G(0) \geq p_c$ *we have* $\mu(F) = \mu(G) = 0$. *(insert reference).*
- $G(r) < \vec{p_c}$ *is partially mandatory if* $\inf supp(F) = \inf supp(G)$ *inside the cone where directed percolation happens one has* $\mu(F, x) = \mu(G, x) = r\|x\|$. *Outside the cone, it was proved in dimension 2 by [Marchand, 2002] that the strict inequality still holds. In higher dimension, the problem is to my knowledge still open, but we can conjecture that they are only two possible scenarios : either the geodesic only uses edges with weight* $\inf F$ *or Theorem 7.3 holds.*

7.3 Back to the limiting function

Corollary 7.6 (Theorem 7.3)**.**
Let $0 < p < q < 1$. *Then*

$$\mathbb{E}_q[T_n] - \mathbb{E}_p[T_n] \geq nc(q - p)$$

As for the proof of Theorem 7.4, we use a coupling $X \sim F$ and $Y \sim G$ with $X \leq Y$. We also consider for each edge an independent random variable $U_e \sim \mathcal{U}([0, 1])$. For $p \in [0, 1]$, we set $\lambda_e(p) = \mathbb{1}_{(U_e \leq p)}$ and $\tau_e(p) = \lambda_e(p) Y_e + (1 - \lambda_e(p)) X_e \sim pG + (1 - p)F$.

We then have

$$\begin{aligned}
\mathbb{E}_q[T_n] - \mathbb{E}_p[T_n] &\geq \mathbb{E}[(\lambda(q)Y + (1 - \lambda(q))X) \cdot \pi_n(q) - (\lambda(p)Y + (1 - \lambda(p))X) \cdot \pi_n(q)] \\
&= \mathbb{E}[(\lambda(q) - \lambda(p))(Y - X) \cdot \pi_n(q)] \\
&\geq \delta \sum_e \mathbb{E}[\mathbb{1}_{(p < U_e < q)} \mathbb{1}_{(e \in \pi(q))} \mathbb{1}_{(X_e \in I)}] \quad \text{with the same notation as in 7.5} \\
&\geq (q - p)\varepsilon \delta n
\end{aligned}$$

In our proof, we used the fact that $F(\inf F) < \vec{p_c}$, yet this will fail for y the Bernoulli setting. This is why we choose γ in the open interval (μ_F, μ_G) as $\mu(p)$ may be flat around 0. The set of $p \in [0,1]$ for which $(1-p)F + pG$ cause the existence of percolation of some sort (that are exactly the p where μ is flat) is lower closed, so it is of the shape $[0, p_1]$ in the case it is non-empty. Thus, in all the cases, μ establish a bijection between $(p_1, 1)$ and (μ_F, μ_G).

8 Conclusion

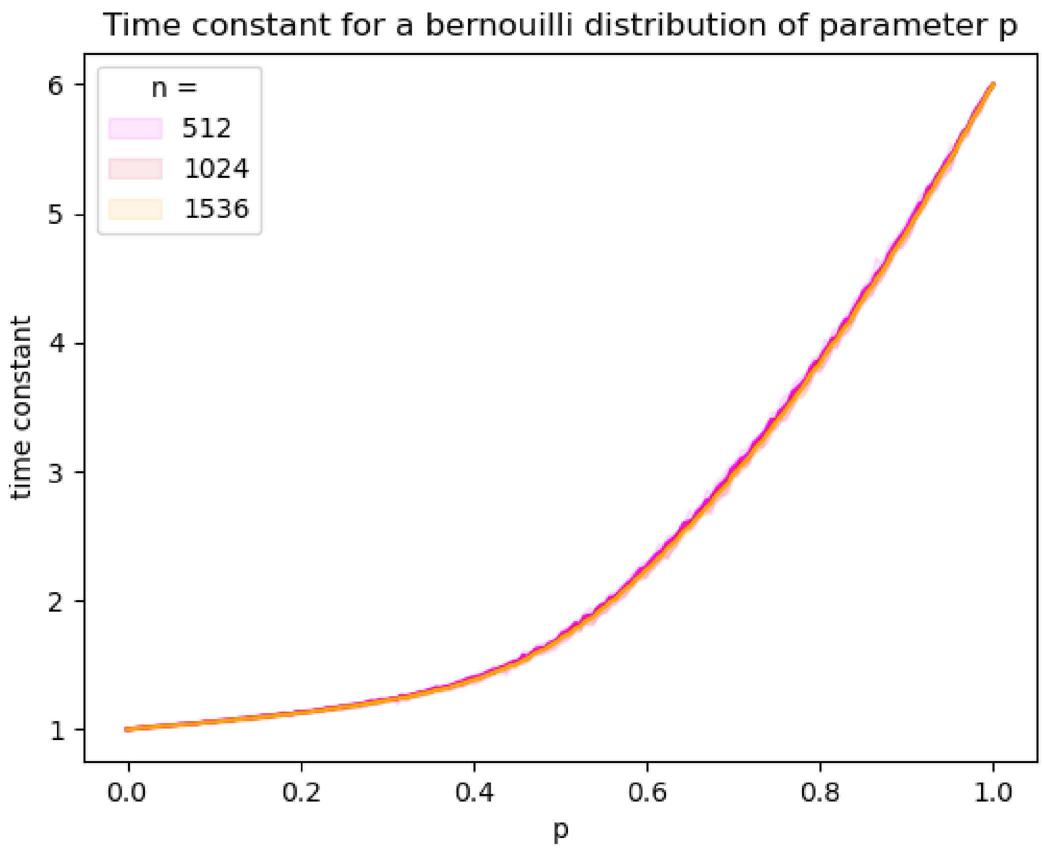


Figure 4: Simulation of the function $\mathbb{E}_p(T_n/n)$

We study the first passage percolation model when the edge's distribution is of the sort $(1-p)F + pG$ and G stochastically dominates F. This gives raise to a series of functions $\mathbb{E}_p[T_n]$, dependent on this parameter p that converge pointwise towards a limit function $\mu(p)$. This function (which are simply segments of the function μ defined over the space of probability distributions) is a natural approach to try to have a better understanding of the behavior of μ. It can also be on itself an interesting model, where p represents the frequency of a slowing factor, thus understanding the behavior of $\mu(p)$ helps us understand the effectiveness of this slowing factor.

The function $\mu(p)$ is also of interest as it is quite rich in structure. [Cerf and Dembin, 2022] proved that this function is Lipschitz continuous in the case where $G = \delta_\infty$ and $F = \delta_1$. In the general case we know that this function is continuous, increasing, very often strictly increasing. We can wonder if the functions $\mathbb{E}_p[T_n]$ converge uniformly in p. We can also ask if this function is derivable and finally :

Question 1 - Is the function $\mu(p)$ defined previously, a convex function?

A first element to answer is the following simulation Figure4.

A collection of open questions can be found in [Auffinger et al., 2017]. Here are some of the questions I've been asking myself :

Question 2 - Find a distribution such that $\frac{1}{\mu(e_1)} B \not\subset B_{\|\cdot\|_2}$

Acknowledgements

I would like to thank my supervisor, Daniel Ahlberg for his guidance, patience, time and support and without whom this thesis would not have existed. Also, I thank Annemarie Luger for giving me the idea to apply for the course and helping me in the process. I'm also grateful to anonymous people who answer mathematical/English/latex/... question on forums on the internet. Finally, thanks to my family for their support.

In memory of Alejandro Cuevas.

References

[Ahlberg et al., 2023] Ahlberg, D., Deijfen, M., and Sfragara, M. (2023). Chaos, concentration and multiple valleys in first-passage percolation.

[Ahlberg and Hoffman, 2016] Ahlberg, D. and Hoffman, C. (2016). Random coalescing geodesics in first-passage percolation. *arXiv preprint arXiv:1609.02447*.

[Auffinger et al., 2017] Auffinger, A., Damron, M., and Hanson, J. (2017). *50 years of first-passage percolation*, volume 68. American Mathematical Soc.

[Benjamini et al., 2011] Benjamini, I., Kalai, G., and Schramm, O. (2011). First passage percolation has sublinear distance variance. *Selected Works of Oded Schramm*, pages 779–787.

[Boucheron et al., 2013] Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press.

[Bourgain et al., 1992] Bourgain, J., Kahn, J., Kalai, G., Katznelson, Y., and Linial, N. (1992). The influence of variables in product spaces. *Israel Journal of Mathematics*, 77:55–64.

[Cerf and Dembin, 2022] Cerf, R. and Dembin, B. (2022). The time constant for bernoulli percolation is lipschitz continuous strictly above pc. *The Annals of Probability*, 50(5):1781–1812.

[Cox and Durrett, 1981] Cox, J. T. and Durrett, R. (1981). Some limit theorems for percolation processes with necessary and sufficient conditions. *The Annals of Probability*, pages 583–603.

[Cox and Kesten, 1981] Cox, J. T. and Kesten, H. (1981). On the continuity of the time constant of first-passage percolation. *Journal of Applied Probability*, 18(4):809–819.

[Damron and Kubota, 2013] Damron, M. and Kubota, N. (2013). Gaussian concentration for the lower tail in first-passage percolation under low moments. *arXiv preprint arXiv:1406.3105*.

[Dembin et al., 2022] Dembin, B., Elboim, D., and Peled, R. (2022). Coalescence of geodesics and the bks midpoint problem in planar first-passage percolation. *arXiv preprint arXiv:2204.02332*.

[Erdős et al., 1960] Erdős, P., Rényi, A., et al. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60.

[Garban and Steif, 2012] Garban, C. and Steif, J. E. (2012). Noise sensitivity of boolean functions and percolation.

[Gorski, 2022] Gorski, C. (2022). Strict monotonicity for first passage percolation on graphs of polynomial growth and quasi-trees. *arXiv preprint arXiv:2208.13922*.

[Hammersley and Welsh, 1965] Hammersley, J. M. and Welsh, D. J. (1965). First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Bernoulli 1713, Bayes 1763, Laplace 1813: Anniversary Volume. Proceedings of an International Research Seminar Statistical Laboratory University of California, Berkeley 1963*, pages 61–110. Springer.

[Kahn et al., 1989] Kahn, J., Kalai, G., and Linial, N. (1989). *The influence of variables on Boolean functions*. Citeseer.

[Kesten et al., 1986] Kesten, C., René, H., Walsh, J. B., and Kesten, H. (1986). Aspects of first passage percolation. In *École d'été de probabilités de Saint Flour XIV-1984*, pages 125–264. Springer.

[Kesten, 1993] Kesten, H. (1993). On the speed of convergence in first-passage percolation. *The Annals of Applied Probability*, pages 296–338.

[Marchand, 2002] Marchand, R. (2002). Strict inequalities for the time constant in first passage percolation. *The Annals of Applied Probability*, 12(3):1001–1038.

[Russo, 1982] Russo, L. (1982). An approximate zero-one law. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 61(1):129–139.

[van den Berg and Kesten, 1993] van den Berg, J. and Kesten, H. (1993). Inequalities for the time constant in first-passage percolation. *The Annals of Applied Probability*, pages 56–80.

A Proofs

A.1 Russo theorem

Proposition ([Russo, 1982]).
For all $\varepsilon > 0$, there exist $\delta > 0$ such that if A is an increasing event satisfying $\mathbb{I}_i^p(A) < \delta$ for all p and all i, then there exists p_0 such that:

$$\mathbb{P}_p(A) \leq \varepsilon \quad \text{for all } p \leq p_0 - \varepsilon$$

$$\mathbb{P}_p(A) \geq 1 - \varepsilon \quad \text{for all } p \geq p_0 + \varepsilon$$

Proof. Let's fix $\varepsilon > 0$. Let $\delta > 0$ be a variable that we will adjust on a later stage.*

Assume $\mathbb{I}_i^p(A) < \delta$. Then by theorem 3.10 we have

$$\sum_i \mathbb{I}_i^p(\mathbb{1}_A) \geq c\mathrm{Var}_p(\mathbb{1}_A)\log(\frac{1}{\delta})$$

Which becomes by theorem 3.6

$$\frac{d}{dp}\mathbb{E}_p[\mathbb{1}_A] \geq c\mathrm{Var}_p(\mathbb{1}_A)\log(\frac{1}{\delta})$$

This inequality is almost the same we obtained for the equally distributed influences case. So we do exactly the same steps:

It now yields

$$p_2 - p_1 \leq \frac{4}{c}\frac{\log(1/2\varepsilon)}{\log(1/\delta)}$$

.

We now choose δ small enough such that this inequality becomes

$$p_2 - p_1 \leq 2\varepsilon$$

and $p_0 = \frac{p_1+p_2}{2}$ will yield the desired result. ∎

A.2 Limiting behavior of tribes

Proposition.
Let f_n be the tribes function defined in 5.1 then

$$\mathbb{E}_p[f_n] = pn^\alpha + O(\sqrt{n^\alpha \ln n})$$

Proof.
Recall that n^α denotes in reality some integer approximation of n^α, let's say the floor of n^α. We will add the floor brackets when applying the logarithm to avoid a rushed exponent simplification.

The idea to prove this is that $\sum_{x_i \in B_j} x_i \sim B(n^\alpha, p)$ thus by the Central Limit Theorem,

$$\frac{\sum_{x_i \in B_j} x_i - pn^\alpha}{n^{\alpha/2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, pq)$$

and we have $\mathbb{E}[\max_{i \in [\![1,n]\!]} N_i] \leq C\sqrt{\log n}$ when $N_i \sim \mathcal{N}(0,1)$ thus

$$\begin{aligned}
\mathbb{E}_p[f_n] &= pn^\alpha + \frac{n^\alpha}{n^{\frac{\alpha}{2}}} \max_{B_1, \ldots, B_K} \frac{\sum_{i \in B_j} x_i - pn^\alpha}{n^{\frac{\alpha}{2}}} \\
&= pn^\alpha + O(n^{\frac{\alpha}{2}} \sqrt{\ln n})
\end{aligned}$$

Let's denote

$X_j = \sum_{x_i \in B_j} x_i \sim B(n^\alpha, p)$,

$Y_j = X_j - pn^\alpha$ and

$Z = \max_j Y_j$

By Jensen inequality, we have

$$\begin{aligned}
e^{t\mathbb{E}[Z]} \leq \mathbb{E}[e^{tZ}] &= \mathbb{E}[\max_j e^{tY_j}] \\
&\leq \sum_j \mathbb{E}[e^{tY_j}] \\
&= \sum_j e^{-tpn^\alpha} \mathbb{E} e^{tX_j} \\
&= \frac{n}{\lfloor n^\alpha \rfloor} e^{-tpn^\alpha} (q + pe^t)^{n^\alpha} \\
&= n^{1-\alpha} \frac{n^\alpha}{\lfloor n^\alpha \rfloor} e^{-tpn^\alpha} (q + pe^t)^{n^\alpha}
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}[Z] &\leq \frac{1}{t}((1-\alpha)\ln n + \ln(1 + \frac{\{n^\alpha\}}{\lfloor n^\alpha \rfloor}) - tpn^\alpha + n^\alpha \ln(1 + p(e^t - 1))) \\
&= \frac{(1-\alpha)\ln n}{t} + \frac{n^\alpha}{t}(-tp + \ln(1 + p(e^t - 1))) + \frac{1}{t} O(\frac{1}{n^\alpha})
\end{aligned}$$

By doing a series expansion in t at order 2 (around 0) we have

$$\begin{aligned}
-tp + \ln(1 + p(e^t - 1)) &= -tp + \ln(1 + p(t + \frac{t^2}{2} + o(t^2))) \\
&= -tp + tp + \frac{t^2}{2} pq + o(t^2)
\end{aligned}$$

Which shows that the right term goes to 0 when t goes to 0. On the other hand, the left term goes to infinity.

We have

$$\mathbb{E}[Z] \leq \frac{(1-\alpha)\ln n}{t} + n^\alpha \frac{t}{2} pq + o(t) + \frac{1}{t} O(\frac{1}{n^\alpha}).$$

If we ignore the $o(t)$ and the $O(\frac{1}{n^\alpha})$ in this expression, the minimum is obtained for $t = \sqrt{2\frac{(1-\alpha)\ln n}{pqn^\alpha}}$ at which the value is

$$\mathbb{E}[Z] \leq n^{\frac{\alpha}{2}} \sqrt{2(1-\alpha)pq\ln(n)} + o(\frac{\ln n}{n^\alpha}) + O(\sqrt{\frac{n^\alpha}{\ln n}})$$

On the other hand we have

$$E[Z] \geq \mathbb{E}[Y_j] = 0$$

By writing $f_n = Z + pn^\alpha$ we get

$$\mathbb{E}_p[f_n] = pn^\alpha + O(\sqrt{n^\alpha \ln n})$$

■

A.3 Characterization of second order stochastic domination

Proposition.
We have a nice characterization for this in the case both distribution have finite mean : F is more variable than G if and only if for all x

$$\int_{-\infty}^{x} G(y)dy \leq \int_{-\infty}^{x} F(y)dy$$

Proof. Assume

$$\forall x \quad \int_{-\infty}^{x} G(y)dy \leq \int_{-\infty}^{x} F(y)dy.$$

Then for $\varphi_x(t) = (t-x)\mathbb{1}_{(-\infty,x]}(t)$ the statement holds. Indeed,

$$\begin{aligned}
\int_{-\infty}^{x} F(y)dy &= \int\int dF(t)dy\mathbb{1}_{(-\infty,x]}(y)\mathbb{1}_{(-\infty,y](t)} \quad \text{Fubini-Tonelli} \\
&= \int dF(t)\int_{t}^{x} dy \\
&= \int (x-t)dF(t)\mathbb{1}_{(-\infty,x)}(t) \\
&= -\mathbb{E}_F[\varphi_x(X)]
\end{aligned}$$

This expectations are finite if F have finite mean.

Let's write V the set of increasing concave functions for which the inequality holds. We have shown that $\varphi_x \in V$. We also have $1 \in V$. We can also show that is a closed convex cone of $(L^1(F) \cap L^1(G), \|.\|_{L^1(F)} + \|.\|_{L^1(G)})$.

So it is just left to shown that we can approximate (in L^1) any integrable increasing concave function by a positive linear combination of φ_x and 1.

We fix $\varepsilon > 0$. Because $\varphi \in L^1$, we can fix $A > 0$ such that

$$\int_{(-\infty,-A)\cup(A,+\infty)} |\varphi(x)|d(F+G)(x) \leq \varepsilon$$

We can furthermore assume that φ is not constant, and because it is non-decreasing and concave, the limit at $-\infty$ is $-\infty$, we can thus increase A to ensure that $\varphi \leq 0$ on $(-\infty, -A]$. Finally, we will enlarge A so that $|\varphi(A)| \int_A^\infty d(F+G)(x) \leq \varepsilon$. (To show this is possible, we consider two cases : if the limit of φ at $+\infty$ is $+\infty$, we will enlarge A so that $\varphi(A) \geq 0$ and thus $\int_A^{+\infty} |\varphi(A)|d(F+G) \leq \int_A^{+\infty} |\varphi(x)|d(F+G)(x) \leq \varepsilon$ else, $\varphi(x)$ stays bounded and $\int_A^{+\infty} d(F+G) = 2-(F+G)(A)$ goes to 0 as A goes to $+\infty$).

The idea is to use the trapezoid rule to approximate the integral in the segment $[-A, A]$. The affine by parts function that results from this approximation can be written as a linear combination of φ_x and 1.

Explicitly : Let $-A = x_0 < x_1 < \cdots < x_{n-1} < x_n = A$ be a subdivision of $[-A, A]$, we approximate φ by the function

$$f_n(x) = \varphi(x_{i+1})\frac{x - x_i}{x_{i+1} - x_i} + \varphi(x_i)\frac{x_{i+1} - x}{x_{i+1} - x_i} \qquad \text{for } x \in [x_i, x_{i+1})$$

We define S_k for $k \in [\![1, n+1]\!]$ as

$$S_k = \frac{\varphi(x_k) - \varphi(x_{k-1})}{x_k - x_{k-1}} \qquad \text{and } S_{n+1} = 0$$

which are the slopes of f in the interval (x_{k-1}, x_k)

We set

$$g_n := \varphi(x_n) + \sum_{k=1}^{n} (S_k - S_{k+1}) \varphi_{x_k}$$

We assert that for all $x \in [-A, A]$, $f(x) = g(x)$. This can be proved using the fact that $\varphi'_{x_k} = \mathbb{1}_{(-\infty, x_k]}$ to show that $f'_n = g'_n$ so $f_n = g_n + C$ by and then evaluate in x_n to show that $C = 0$.

Using the fact that $x_k = (1-t)x_{k-1} + t x_{k+1}$ with $t = \frac{x_k - x_{k-1}}{x_{k+1} - x_{k-1}}$ and that φ is concave we get $\varphi(x_k) \geq (1-t)\varphi(x_{k-1}) + t\varphi(x_{k+1})$, thus

$$\begin{aligned}
S_k - S_{k+1} &= \frac{\varphi(x_k) - \varphi(x_{k-1})}{x_k - x_{k-1}} + \frac{\varphi(x_k) - \varphi(x_{k+1})}{x_{k+1} - x_k} \\
&\geq \frac{(1-t)\varphi(x_{k-1}) + t\varphi(x_{k+1}) - \varphi(x_{k-1})}{x_k - x_{k-1}} + \frac{(1-t)\varphi(x_{k-1}) + t\varphi(x_{k+1}) - \varphi(x_{k+1})}{x_{k+1} - x_k} \\
&= \frac{\varphi(x_{k+1}) - \varphi(x_{k-1})}{x_{k+1} - x_{k-1}} + \frac{\varphi(x_{k-1}) - \varphi(x_{k+1})}{x_{k+1} - x_{k-1}} = 0.
\end{aligned}$$

We proved that for $k \in [\![1, n-1]\!]$ $S_k - S_{k+1} \geq 0$. Lastly $S_n - S_{n+1} = S_n \geq 0$ because φ is non-decreasing.

Finally, we show that for $x \leq x_0$, $\varphi(x) \leq g_n(x) \leq 0$. Indeed, let $x_{-1} < x_0$, we can define $S_0 = \frac{\varphi(x_0) - \varphi(x_{-1})}{x_0 - x_{-1}}$ and $h_n = g_n + (S_0 - S_1)\varphi_{x_0}$. By the same argument as before, h_n verifies : $h_n(x_{-1}) = \varphi(x_{-1})$. Because $(S_0 - S_1) \geq 0$ and $\varphi_{x_0} \leq 0$ we deduce $\varphi(x_{-1}) \leq g_n(x_{-1})$. On the other hand, we have $g_n(x) \leq g_n(x_0) = g_n(-A) \leq 0$ by our choice of A. We deduce form this that

$$|\varphi| \geq |g_n| \quad \text{in the interval } (-\infty, -A].$$

Because φ is absolutely continuous on $[-A, A]$, we can fix a partition of $[-A, A]$ fine enough so that

$$\sup_{x \in [-A, A]} |\varphi(x) - f_n(x)| \leq \frac{\varepsilon}{2A}$$

. We thus have

$$\begin{aligned}
\int |\varphi(x) - g_n(x)| d(F+G)(x) &\leq \int_{-\infty}^{-A} |\varphi(x) - g_n(x)| d(F+G)(x) + \int_{-A}^{A} |\varphi(x) - g_n(x)| d(F+G)(x) \\
&\quad + \int_{A}^{+\infty} |\varphi(x) - g_n(x)| d(F+G)(x) \\
&\leq 2 \int_{-\infty}^{-A} |\varphi(x)| d(F+G)(x) + \int_{-A}^{A} \frac{\varepsilon}{2A} d(F+G)(x) + \int_{A}^{+\infty} |\varphi(x)| + |g_n(x)| d(F+G)(x) \\
&\leq 2\varepsilon + \varepsilon + \varepsilon + \int_{A}^{+\infty} |\varphi(A)| d(F+G)(x) \quad \text{using the fact that } g_n(x) = \varphi(A) \text{ for } x \geq A \\
&\leq 5\varepsilon
\end{aligned}$$

■

B Fourier Theory

Most of what will follow is a retake of [Garban and Steif, 2012]. Most of the results presented in this section come from the more general theory of Fourier analysis on finite groups. Here we will restrain ourselves to the group $\Omega_n \cong (\mathcal{P}([\![1, n]\!]), \Delta) \cong (\mathbb{Z}/2\mathbb{Z}, +)^n \cong (\mathbb{U}_2, \times)^n$. We will prefer this last representation as the characters are more natural. So unless otherwise specified, for all this section we set $\Omega_n = \{-1, 1\}$.

We consider the space $L^2(\Omega_n)$ of complex valued functions endowed with the inner product :

$$\begin{aligned}\langle f|g\rangle &= \sum_{x_1,\dots,x_n} f(x_1,\dots,x_n)\overline{g(x_1,\dots,x_n)}2^{-n}\\ &= \mathbb{E}[f\overline{g}]\end{aligned}$$

where $\mathbb{E}$ denotes the expectation with respect to the uniform measure $\mathbb{P}$ on Ω_n.

For any subset $S \subset [\![1,n]\!]$, let χ_S be the function on Ω_n defined by :

$$\chi_S(x) := \prod_{i\in S} x_i$$

We have $L^2(\Omega)$ is a $\mathbb{C}$-vectorial space of dimension $|\Omega^n| = 2^n$. The $(\chi_S)_{S\subset[\![1,n]\!]}$ form an orthonormal basis.

Proof.

- $\mathbb{E}[\chi_S^2] = \mathbb{E}[1] = 1$
- $\mathbb{E}[\chi_S\chi_P] = \mathbb{E}[\chi_{S\Delta P}]$ It suffices then to show that for every $S \neq \emptyset$ we have $\mathbb{E}[\chi_S] = 0$. Suppose S is not empty, then we can fix $i \in S$ and by conditioning on the value of the i component we get

$$\begin{aligned}\mathbb{E}[\chi_S] &= \mathbb{E}[\chi_S|x_i = 1]\mathbb{P}(x_i = 1) + \mathbb{E}[\chi_S|x_i = -1]\mathbb{P}(x_i = -1)\\ &= \mathbb{E}[\chi_S|x_i = 1]\frac{1}{2} - \mathbb{E}[\chi_S|x_i = 1]\frac{1}{2}\\ &= 0\end{aligned}$$

■

Thus, we have for any $f \in L^2(\Omega_n)$:

$$f = \sum_{S\subset[\![1,n]\!]} \langle f|\chi_S\rangle \chi_S$$

We denote $\hat{f}(S) := \langle f|\chi_S\rangle = \mathbb{E}[f\chi_S]$ that are called the Fourier-Walsh coefficients of f.

By identifying $\Omega_n \approx \mathcal{P}([\![1,n]\!])$ we get an isometry from $L^2(\Omega_n)$ to $L^2(\mathcal{P}([\![1,n]\!]))$

$$\langle f|g\rangle = \sum_{S\subset[\![1,n]\!]} \hat{f}(S)\overline{\hat{g}(S)} = \left\langle \hat{f}\middle|\hat{g}\right\rangle$$

Definition B.1 (Energy Spectrum)**.**
The energy spectrum E_f of a function $f \in L^2(\Omega_n)$ is defined by

$$E_f(m) := \sum_{S\in\mathcal{P}_m([\![1,n]\!])} \hat{f}(S)^2, \quad m \in [\![1,n]\!].$$

Of course, we have $\sum_{m=1}^n E_f(m) = \|f\|_2^2$. So this corresponds to the L^2 norm of $\sum_{\#|S|=m} \hat{f}(S)\chi_S$.

Definition B.2 (Convolution operator)**.**
We define

$$* : \left| \begin{array}{rcl} L^2(\Omega_n)\times L^2(\Omega_n) & \longrightarrow & L^2(\Omega_n) \\ f,g & \longmapsto & f*g \end{array}\right.$$

by

$$\begin{aligned}f*g(u) :&= \sum_{\omega\in\Omega_n} f(\omega)g(\omega u)\mathbb{P}(\omega)\\ &= \mathbb{E}[f\overline{(g\circ\tau_u)}]\end{aligned}$$

Where $\tau_u : \left| \begin{array}{rcl} \Omega_n & \longrightarrow & \Omega_n \\ \omega & \longmapsto & \omega s := (\omega_1 u_1,\dots,\omega_n u_n) \end{array}\right.$

Theorem B.3 (Convolution theorem).
This result is more general (see Fourier analysis in finite groups) but we only prove it for this case.

$$\widehat{f * g} = \hat{f}\hat{g}$$

Proof.

$$\begin{aligned}
\langle f * g | \chi_S \rangle &= \sum_{u \in \Omega_n} f * g(u) \chi_S(u) \mathbb{P}(u) \\
&= \sum_{u \in \Omega_n} \sum_{\omega \in \Omega_n} f(\omega) g(u\omega) \mathbb{P}(\omega) \chi_S(u) \mathbb{P}(u) \\
&= \sum_{\omega \in \Omega_n} f(\omega) \mathbb{P}(\omega) \sum_{u \in \Omega_n} g(u\omega) \chi_S(u) \mathbb{P}(u) \\
&= \sum_{\omega \in \Omega_n} f(\omega) \mathbb{P}(\omega) \sum_{u \in \Omega_n} g(u) \chi_S(u\omega) \mathbb{P}(u\omega) \quad \tau_\omega \text{ is a bijection} \\
&= \sum_{\omega \in \Omega_n} f(\omega) \chi_S(\omega) \mathbb{P}(\omega) \sum_{u \in \Omega_n} g(u) \chi_S(u) \mathbb{P}(u) \quad \text{because } \chi_S(u\omega) = \chi_S(u)\chi(\omega) \\
&= \hat{f}(S)\hat{g}(S)
\end{aligned}$$

■

As any Fourier transformation, this one yields very interesting results. We will focus on the one's closer to our area of study:

Proposition.
If $f : \Omega_n \to \{0, 1\}$ then for any $k \in [\![1, n]\!]$ we have :

$$\mathbb{I}_k(f) = 4 \sum_{S \subset [\![1,n]\!]} \hat{f}(S)^2 \mathbb{1}_{k \in S}$$

and

$$\mathbb{I}(f) = 4 \sum_{S \subset [\![1,n]\!]} |S| \hat{f}(S)^2$$

Proof. We introduce the discrete derivative operator (along the kth coordinate):

$$\nabla_k f : \left| \begin{array}{ccc} \Omega_n & \longrightarrow & \mathbb{C} \\ \omega & \longmapsto & f(\omega) - f(\sigma_k(\omega)) \end{array} \right.$$

Where σ_k maps Ω_n to itself by flipping the kth bit.

By writing $f(\omega) = \sum_{S \subset [\![1,n]\!]} \hat{f}(S) \chi_S(\omega)$ we get

$$\nabla_k f(\omega) = \sum_{S \subset [\![1,n]\!]} \hat{f}(S)(\chi_S(\omega) - \chi_S(\sigma_k(\omega))) = 2 \sum_{S \subset [\![1,n]\!]} \hat{f}(S) \chi_S(\omega) \mathbb{1}_{k \in S}$$

By uniqueness of the decomposition, it follows:

$$\widehat{\nabla_k f}(S) = 2 \mathbb{1}_{k \in S} \hat{f}(S)$$

Because $f : \Omega_n \to \{0, 1\}$ we have $\mathbb{I}_k(f) = \|\nabla_k f\|_2^2$. Applying Parseval's formula

$$\mathbb{I}_k(f) = \sum_{S \subset [\![1,n]\!]} 4 \hat{f}(S)^2 \mathbb{1}_{k \in S}$$

Finally, by summing over k and switching both sums, we get the second result. ■

Remark.

If f maps into $\{-1,1\}$ then $\frac{f+1}{2}$ maps into $\{0,1\}$ and $\mathbb{I}_k(f) = \mathbb{I}_k(\frac{f+1}{2})$ and $\widehat{\frac{f+1}{2}} = \frac{1}{2}\hat{f} + \frac{1}{2}\mathbb{1}_\emptyset$ thus

$$\mathbb{I}_k(f) = \sum_{S\subset [\![1,n]\!]} \hat{f}(S)^2 \mathbb{1}_{k\in S}$$

B.1 Monotone functions and their spectra

Monotone functions enjoy alternative useful spectra description

Proposition.
If $f : \Omega_n \to \{0,1\}$ is monotone, then for all k

$$\mathbb{I}_k(f) = 2\hat{f}(\{k\})$$

Proof.

$$\begin{aligned}\hat{f}(\{k\}) &= \mathbb{E}[f\chi_{\{k\}}] \\ &= \mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\in\Pi}] + \mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\notin\Pi}]\end{aligned}$$

The term $\mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\notin\Pi}] = 0$ is null. Indeed,

$$\begin{aligned}\mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\notin\Pi}] &= \mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\notin\Pi}|x_k = 1]\mathbb{P}(x_k = 1) + \mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\notin\Pi}|x_k = -1]\mathbb{P}(x_k = -1) \\ &= \mathbb{E}[f\mathbb{1}_{k\notin\Pi}|x_k = 1]\mathbb{P}(x_k = 1) - \mathbb{E}[f\mathbb{1}_{k\notin\Pi}|x_k = 1]\mathbb{P}(x_k = -1) \\ &= \mathbb{E}[f\mathbb{1}_{k\notin\Pi}]\mathbb{P}(x_k = 1) - \mathbb{E}[f\mathbb{1}_{k\notin\Pi}]\mathbb{P}(x_k = -1) \quad \text{because } f\mathbb{1}_{k\notin\Pi} \text{ is independent of } x_k \\ &= 0\end{aligned}$$

The second term is equal to $\mathbb{I}_k(f)/2$. Indeed,

$$\begin{aligned}\mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\in\Pi}] &= \mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\in\Pi}|x_k = 1]\mathbb{P}(x_k = 1) + \mathbb{E}[f\chi_{\{k\}}\mathbb{1}_{k\in\Pi}|x_k = -1]\mathbb{P}(x_k = -1) \\ &= \mathbb{E}[\mathbb{1}_{k\in\Pi}]\mathbb{P}(x_k = 1) \\ &= \mathbb{I}_k(f)/2\end{aligned}$$

By noticing that $f\chi_k\mathbb{1}_{k\in\Pi}\mathbb{1}_{x_k=1} \equiv \mathbb{1}_{k\in\Pi}\mathbb{1}_{x_k=1}$ and $f\chi_k\mathbb{1}_{k\in\Pi}\mathbb{1}_{x_k=-1} \equiv 0$. This is where the hypothesis of f being increasing is important. ∎

In fact, we can extend this proof to functions $f : \Omega_n \to \mathbb{R}$ which gives us:

$$\hat{f}(\{k\}) = \frac{1}{2}\mathbb{E}[|\nabla_i f|]$$

B.2 Non-uniform probability

If we replace $\mathbb{P}$ by $\mathbb{P}_p$ we loose the orthogonality of our basis. We will thus consider this change as the multiplication by a weight function

$$W_p(\omega) = \frac{\mathbb{P}_p(\omega)}{\mathbb{P}(\omega)}.$$

We can thus expect the behavior of changing p to some sort of convolution in the Fourier space.

How do we define this convolution?

We start from writing that $f = \sum_S \hat{f}(S)\chi_S$ and $g = \sum_S \hat{g}(S)\chi_S$ and by noticing that $\chi_S\chi_T = \chi_{S\Delta T}$ so

$$\begin{aligned}fg &= \sum_{S,T} \hat{f}(S)\hat{g}(T)\chi_S\chi_T \\ &= \sum_U \sum_{S\Delta T = U} \hat{f}(S)\hat{g}(T)\chi_U \\ &= \sum_U \hat{f} * \hat{g}(U)\chi_U\end{aligned}$$

Where we defined

$$\hat{f} * \hat{g}(S) := \sum_{S_1 \Delta S_2 = S} \hat{f}(S_1)\hat{g}(S_2)$$
$$= \sum_{T} \hat{f}(T)\hat{g}(S\Delta T)$$

Proposition.

$$\widehat{W}_p(S) = (2p-1)^{\#|S|}$$

and

$$E_{W_p}(m) = \binom{n}{m}(2p-1)^{2m}$$

Proof.

$$\begin{aligned}
\langle W_p|\chi_S\rangle &= \mathbb{E}[W_p\chi_S] \\
&= \mathbb{E}_p[\chi_S] \\
&= \mathbb{E}_p[\prod_{k\in S} x_k] \\
&= \prod_{k\in S} \mathbb{E}_p[x_k] \quad \text{by independence of the } x_k \\
&= (2p-1)^{\#|S|}
\end{aligned}$$

■

We can generalize this proof to the case we take n different probabilities for each coordinate, we get

$$\widehat{W}_{p_1,\ldots,p_n}(S) = \prod_{k\in S}(2p_k - 1)$$

We can deduce the energy by just computing the definition. We can notice the fact that we have a symmetry with respect the axis $p = \frac{1}{2}$.

We remark we have the same phenomenon as in other Fourier transform : when p is close to $\frac{1}{2}$, W_p is very flat, close to uniform and in the Fourier space, we will get a very sharp function that decreases very fast as $\#|S|$ grows bigger. In the opposite case, when p goes to either 0 or 1, W_p is very sharp, with all the weight being concentrated close to either 0_n or 1_n and in the Fourier space, the function becomes more flat. It would be interesting to see if there is also an uncertainty principle in this setting.

B.3 Another proof of the Margulis-Russo

Now we will try to prove some results from the previous part:

Margulis-Russo :

We can rewrite the Margulis-Russo formula as follows :

$$\frac{d}{dp}\langle f|W_p\rangle = \langle|\nabla f||W_p\rangle$$

$$
\begin{aligned}
\partial_i \langle f | W_{p_1,\dots,p_n} \rangle &= \partial_i \left\langle \sum_S \hat{f}(S)\chi_S \middle| \sum_S \widehat{W}(S)\chi_S \right\rangle \\
&= \left\langle \sum_S \hat{f}(S)\chi_S \middle| \sum_S \partial_i \widehat{W}(S)\chi_S \right\rangle \\
&= \left\langle \sum_S \hat{f}(S)\chi_S \middle| \sum_S \mathbb{1}_{i\in S} 2\widehat{W}(S \setminus \{i\})\chi_S \right\rangle \\
&= \left\langle \sum_S \mathbb{1}_{i\in S} 2\hat{f}(S)\chi_S \middle| \sum_{i\in S} \widehat{W}(S \setminus \{i\})\chi_S \right\rangle
\end{aligned}
$$

Because f is increasing, we have

$$|\nabla_i f| = \chi_i \nabla_i f$$

Thus

$$
\begin{aligned}
\widehat{|\nabla_i f|}(S) &= \delta_i * \widehat{\nabla_i f}(S) \\
&= \widehat{\nabla_i f}(S \Delta \{i\}) \\
&= \mathbb{1}_{i\notin S} 2\hat{f}(S \sqcup \{i\})
\end{aligned}
$$

Let's continue :

$$
\begin{aligned}
\partial_i \langle f | W_{p_1,\dots,p_n} \rangle &= \left\langle \sum_{i\notin S} 2\hat{f}(S \sqcup \{i\})\chi_S\chi_i \middle| \sum_{i\in S} \widehat{W}(S \setminus \{i\})\chi_S \right\rangle \\
&= \left\langle \sum_{i\notin S} 2\hat{f}(S \sqcup \{i\})\chi_S\chi_i \middle| \sum_{i\notin S} \widehat{W}(S)\chi_S\chi_i \right\rangle \\
&= \left\langle \sum_{i\notin S} \widehat{|\nabla_i f|}(S)\chi_S\chi_i \middle| \sum_{i\notin S} \widehat{W}(S)\chi_S\chi_i \right\rangle \\
&= \sum_S \widehat{|\nabla_i f|}(S)\widehat{W}(S) \\
&= \langle |\nabla_i f| | W_{p_1,\dots,p_n} \rangle
\end{aligned}
$$

We get the more general result for any kind of boolean function:

$$\frac{d}{dp}\mathbb{E}_p[f] = \sum_{i=1}^n \mathbb{E}_p[\chi_i \nabla_i f]$$

C The Margulis-Russo Formula

C.1 What regularity do we have for $\mu(p)$?

Here is another proof.

Theorem C.1 (Generalized Margulis-Russo Formula)**.**
Let $T : \Omega_n \to \mathbb{R}_+$ be a non-decreasing (bounded) random variable. Then we have :

$$\frac{d}{dp}\mathbb{E}_p[T] = \sum_{i=1}^n \mathbb{E}_p[|\nabla_i T|]$$

Proof.

$$
\begin{aligned}
\frac{d}{dp}\mathbb{E}_p[T] &= \frac{d}{dp}\int_0^\infty \mathbb{P}_p\left(T > \gamma\right) d\gamma \\
&= \frac{d}{dp}\int_0^\infty \mathbb{E}_p\left[\mathbb{1}_{T>\gamma}\right] d\gamma \\
&= \int_0^\infty \frac{d}{dp}\mathbb{E}_p\left[\mathbb{1}_{T>\gamma}\right] d\gamma
\end{aligned}
$$

Because T is non-decreasing, then $\mathbb{1}_{T>\gamma}$ is non-decreasing for all γ. We thus have by the Margulis-Russo formula:

$$
\frac{d}{dp}\mathbb{E}_p\left[\mathbb{1}_{T>\gamma}\right] = \sum_{i=1}^n \mathbb{I}_i^p(\mathbb{1}_{T>\gamma})
$$

that we can bound by n for $\gamma \leq \sup T$ and 0 for $\gamma > \sup T$ which justifies the derivation/integration inversion.

$$
\begin{aligned}
\frac{d}{dp}\mathbb{E}_p[T] &= \sum_{i=1}^n \int_0^\infty \mathbb{I}_i^p(\mathbb{1}_{T>\gamma}) d\gamma \\
&= \sum_{i=1}^n \int_0^\infty \sum_{\omega\in\Omega_n} |\mathbb{1}_{T>\gamma}(\omega) - \mathbb{1}_{T>\gamma}(\sigma_i\omega)| \mathbb{P}_p\left(\omega\right) d\gamma \\
&= \sum_{i=1}^n \sum_{\omega\in\Omega_n} \int_0^\infty |\mathbb{1}_{T>\gamma}(\omega) - \mathbb{1}_{T>\gamma}(\sigma_i\omega)| d\gamma \mathbb{P}_p\left(\omega\right) \\
&= \sum_{i=1}^n \sum_{\omega\in\Omega_n} \left|\int_0^\infty \mathbb{1}_{T>\gamma}(\omega) - \mathbb{1}_{T>\gamma}(\sigma_i\omega) d\gamma\right| \mathbb{P}_p\left(\omega\right) \\
&= \sum_{i=1}^n \sum_{\omega\in\Omega_n} |T(\omega) - T(\sigma_i\omega)| \mathbb{P}_p\left(\omega\right) \\
&= \sum_{i=1}^n \mathbb{E}_p\left[|\nabla_i T|\right]
\end{aligned}
$$

∎

Let F and G be two distributions, then the joint measure for $\mathbb{P}_p$ is $\nu = \otimes^E((1-p)F + pG)$.

We then have

$$
\mathbb{E}_p[T] = \int T(\omega)\nu(\omega)
$$

We can derive the measure :

$$
\frac{d}{dp}\bigotimes^{E}((1-p)F + pG) = \sum_{e\in E}\bigotimes^{<e}((1-p)G + pF) \otimes (G - F) \otimes \bigotimes^{>e}((1-p)G + pF)
$$

Assuming the switch between the derivation and integral operator is legit.

$$
\frac{d}{dp}\mathbb{E}_p[T] = \sum_{e\in E}\mathbb{E}_p[T|\lambda_e = 1] - \mathbb{E}_p[T|\lambda_e = 0]
$$

We set

$$
\theta_e = T|\tau_e = \infty - T|\tau_e = 0.
$$

This random variable tells us how important is the edge for the percolation. We can thus decompose

$$
T = (\omega_e + T|\tau_e = 0)\mathbb{1}_{\omega_e < \theta_e} + \mathbb{1}_{\tau_i \geq} T|\tau_e = \infty
$$

We recall the existence of a time constant

$$\mu(F) \overset{\text{a.s.}}{=} \lim_{n\to\infty} \frac{T_n}{n}$$

Theorem C.2 (Continuity of the time constant)**.**

$$\textit{If } F_k \xrightarrow{w} F \textit{ then } \mu(F_k) \to \mu(F)$$

under the assumption of a dominating function.

Theorem C.3 (Continuity of the time constant)**.**

$$\textit{If } F_k \xrightarrow{w} F \textit{ then } \mu(F_k) \to \mu(F)$$

Lemma.
Assume $F_k \to F$. Then $\limsup \mu(F_k) \le \mu(F)$.

Lemma.
Define $F^B(x) = \mathbb{1}_{x<B}U(x) + \mathbb{1}_{x\ge B}$. Then $\mu(F^B) \to \mu(F)$.

Proof. Let's assume both lemmas for the moment. Because of the first lemma, it suffices to show that $\liminf_k \mu(F_k) \ge \mu(F)$. Then, let B a continuity point of F. We have $F_k \le F_k^B$ thus $\liminf \mu(F_k^B) \le \mu(F_k)$ and $F_k^B \xrightarrow[k\infty]{w} F^B$. ∎

Proof. We start by noticing that μ is a monotone operator, in the sense that if $G \le F$ (G dominates F) then $\mu(F) \le \mu(G)$.

Because of this, by setting $\tilde{F_k} = \min(F_k, F)$ and $\overline{F_k} = \max(F_k, F)$ we have $\tilde{F_k} \le F_k \le \overline{F_k}$ thus $\mu(\tilde{F_k}) \ge \mu(F_k) \ge \mu(\overline{F_k})$ and because $\tilde{F_k} \to F$ and $\overline{F_k} \to F$ it shows that proving the result for $F_k \ge F$ and $F_k \le F$ is enough. Of course if $F_k \ge F$ then $\forall k, \mu(F_k) \le \mu(F)$ and thus $\limsup \mu(F_k) \le \mu(F)$.

So we only consider the case where $F_k \le F$. We define $U^{-1}(x) = \inf_t\{t|U(t) > x\}$ the inverse cumulative distribution. We thus have $F_k^{-1}(x) \to F^{-1}(x)$.

We can then fix $(\xi_i)_i \sim \mathcal{U}$ and the pushfoward distribution through F_n^{-1} and F^{-1} verifies $F^{-1}(\xi_i) \sim F$ and $F^{-1}(\xi_i) \le F_n^{-1}(\xi_i)$

We define $X_i = F^{-1}(\xi_i)$ and $X_i^{(k)} = F_k^{-1}(\xi_i)$

Let's fix $\varepsilon > 0$. We can fix a path γ such that $T_n^F \ge \gamma \cdot X + \varepsilon$. (Actually the existence of a minimizing path is assured by various theorems, but I will try to avoid them for the moment).

We then have $T_n^{(k)} \le \gamma \cdot X^{(k)} \le \gamma \cdot X^{(k)} + \gamma \cdot (X^{(k)} - X) \le T_n + \varepsilon + \gamma \cdot (X^{(K)} - X)$. We can then fix K such that $\gamma \cdot (X^{(k)} - X) \le \varepsilon$ for all $k \ge K$.

This proves that $\limsup_k T_n^{(k)} \le T_n$ almost surely. And because $T_n^{(k)} \ge T_n$ a.s. we get

$$\lim_k T_n^{(k)} = T_n \quad \text{a.s.}$$

To get the L^1 convergence, we need an argument of domination (for the moment). We thus make the assumption that there exist some G such that $G \le F_k$ and that G admits a finite expectation. (We can then apply either reverse Fatou Lemma or Lebesgue dominated convergence) we get :

$$\lim_k \mathbb{E}_{F_k}[T_n] \le \mathbb{E}_F[T_n]$$

Now, we fix N such that

$$\mathbb{E}_F\left[\frac{T_N}{N}\right] - \mu(F) \leq \varepsilon.$$

We now fix K such for all $k \geq K$,

$$\mathbb{E}_{F_k}\left[\frac{T_N}{N}\right] - \mathbb{E}_F\left[\frac{T_N}{N}\right] \leq \varepsilon$$

We thus get

$$\begin{aligned}\mu(F_k) &\leq \mathbb{E}_{F_k}\left[\frac{T_N}{N}\right]\\ &\leq \mathbb{E}_F\left[\frac{T_N}{N}\right] + \varepsilon\\ &\leq \mu(F) + 2\varepsilon\end{aligned}$$

The more general proof can be found in [Cox and Kesten, 1981]. The idea is to reuse the proof of the existence of the time constant in the not integrable case [Cox and Durrett, 1981] which is to use circuits.

Let $G(x) \leq F_k(x)$ for all k, x. Such a distribution exists (we can take the inf and verify that the obtained distribution does not put weight on infinity...).

Then, we fix x_0 such that $G(x_0) > \frac{1}{4}$. We can then go from FPP to classical percolation by saying an edge is open if $X_e < x_0$. As a result from classical percolation, because $\mathbb{P}(X_e < x_0) > p_c = \frac{1}{2}$ we have a unique infinite cluster. For each point x we consider ∂x be the smallest circuit (one loop) of open edges (in a way, if we augment each closed vertex by $(-1,1)^2$ this will correspond to the frontier of the adherence) and $\mathring{x}$ be the interior of this circuit. The idea is to study

$$\tilde{d}(x,y) = d(\partial x, \partial y)$$

which happens to have the good properties. ■